

Dyslexi hos barn och ungdomar

Tester och insatser

En systematisk litteraturöversikt

Augusti 2014



SBU • Statens beredning för medicinsk utvärdering
Swedish Council on Health Technology Assessment

SBU utvärderar sjukvårdens metoder

SBU, Statens beredning för medicinsk utvärdering, är en statlig myndighet som utvärderar hälso- och sjukvårdens metoder.

SBU analyserar metodernas nytta, risker och kostnader och jämför vetenskapliga fakta med svensk vårdpraxis. Målet är att ge ett bättre beslutsunderlag för alla som avgör hur vården ska utformas.

SBU ger ut flera rapportserier. I ”SBU Utvärderar” har SBU:s expertgrupper själva gjort den systematiska utvärderingen. Serien omfattar både etablerade metoder (gula rapporter) och nya metoder (Alert). ”SBU Kommenterar” sammanfattar och kommenterar utländska medicinska kunskapsöversikter. SBU svarar också på frågor direkt från beslutsfattare i vården via SBU:s Upplysningstjänst.

Välkommen att läsa mer om SBU:s rapporter och verksamhet på www.sbu.se.

Denna utvärdering publicerades år 2014. Resultat som bygger på ett starkt vetenskapligt underlag fortsätter vanligen att gälla under en lång tid framåt. Andra resultat kan ha hunnit bli inaktuella. Det gäller främst områden där det vetenskapliga underlaget är otillräckligt eller begränsat.

Denna rapport (nr 225) kan beställas från Strömberg distribution
Telefon: 08-779 96 85 • Fax: 08-779 96 10 • E-post: sbu@strd.se

Grafisk produktion av Anna Edling, SBU
Tryckt av Elanders Sverige AB, Mölnlycke, 2014
Rapportnr: 225 • ISBN 978-91-85413-66-9 • ISSN 1400-1403

Citera denna rapport: SBU. Dyslexi hos barn och ungdomar – tester och insatser. En systematisk litteraturöversikt. Stockholm: Statens beredning för medicinsk utvärdering (SBU); 2014. SBU-rapport nr 225. ISBN 978-91-85413-66-9.

Dyslexi hos barn och ungdomar

Tester och insatser

En systematisk litteraturoversikt

Projektgrupp

<i>Sakkunniga</i>	<i>SBU</i>
Stefan Samuelsson (ordförande)	Agneta Brolund (informationsspecialist)
Barbro Bruce	Thomas Davidson
Gunnel Colnerud	(hälsoekonom)
Christina Hellman	Kickan Håkanson
Mona Näreskog	(projektadministratör)
Idor Svensson	Agneta Pettersson
Ulrika Wolff	(biträdande)
Jakob Åsberg	projektledare)
Petter Gustavsson (extern referensexpert)	Karin Stenström (projektledare)
Gunilla Salo (extern referensperson)	

Externa granskare

Carsten Elbro	Pekka Niemi
Bente Eriksen Hagtvet	Eva Wigforss

Innehåll

SBU:s sammanfattning och slutsatser	11
1. Inledning	29
Syfte	30
Målgrupper	30
2. Bakgrund	33
Begrepp och definitioner	33
Dyslexi är en funktionsnedsättning	33
Dyslexi har definierats på olika sätt under åren	34
Dyslexi – en fråga om gränsdragning	35
Samförekomst med andra funktionsnedsättningar	35
Identifiering av dyslexi	36
Förutsättningar och nulägesbeskrivning	36
SBU:s uppdrag och oberoende roll	37
Jävsproblematik	37
3. Metodbeskrivning	39
Frågor	39
Avgränsningar	39
<i>Inklusionskriterier för tester för att förutsäga dyslexi</i>	40
<i>Inklusionskriterier för tester för att upptäcka och utreda dyslexi</i>	41
<i>Inklusionskriterier för insatser</i>	41
Val av gräns för dyslexi	42
Val av tester för granskning	43
Val av studiedesign	44
Urval av studier	44
Litteratursökning	46
Granskningsprocessen	47
Fas 1: Gallring av artiklar mot inklusions- och exklusionskriterier	47
Fas 2: Kvalitetsgranskning av relevanta artiklar	48
Fas 3: Tabellering av relevanta data från studier	48

Metoder för sammanvägning av resultat	48
Det vetenskapliga underlagets styrka	49
Stegen vid evidensgradering	50
4. Resultatsammanställning	51
Bedömning av evidensstyrkan	53
5. Att förutsäga dyslexi – resultat av granskningen	59
Beskrivning av ingående studier	61
Fonologisk medvetenhet	61
Beskrivning av resultatet	61
Bedömning av evidensstyrkan	63
Snabb automatiserad benämningsförmåga (RAN)	63
Beskrivning av resultatet	63
Bedömning av evidensstyrkan	65
Bokstavskännedom	65
Beskrivning av resultatet	65
Bedömning av evidensstyrkan	66
Sammanslagna prediktorer	67
Beskrivning av resultatet	67
6. Tester för att upptäcka och utreda dyslexi – resultat av granskningen	69
Beskrivning av ingående studier	71
Beskrivning av resultaten	71
Bedömning av evidensstyrkan	77
7. Insatser för barn och ungdomar med dyslexi – resultat av granskningen	81
Randomiserade kontrollerade studier	83
Insatser som jämförs med andra insatser	83
Beskrivning av ingående studier	83
Beskrivning av resultaten	84
Bedömning av sammanvägd effekt och evidensstyrka	88
Insatser som jämförs med ordinarie undervisning	88
Beskrivning av ingående studier	88

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet korrekt läsning av ord	89
Beskrivning av resultatet	89
Bedömning av evidensstyrkan	91
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet korrekt läsning av non-ord	92
Beskrivning av resultatet	92
Bedömning av evidensstyrkan	94
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet stavning	95
Beskrivning av resultatet	95
Bedömning av evidensstyrkan	97
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läsförståelse	98
Beskrivning av resultatet	98
Bedömning av evidensstyrkan	100
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet fonologisk medvetenhet	101
Beskrivning av resultatet	101
Bedömning av evidensstyrkan	103
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läshastighet	104
Beskrivning av resultatet	104
Bedömning av evidensstyrkan	106
Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läsförståelse, efter sex månader till ett år	107
Beskrivning av resultatet	107
Bedömning av evidensstyrkan	109
Träning av enbart fonologisk medvetenhet för utfallsmåtten korrekt läsning av ord, stavning, läsförståelse och fonologisk medvetenhet	110
Beskrivning av resultatet	110
Bedömning av evidensstyrkan	112
Träning av enbart läsflyt för utfallsmåtten läsning av ord på tid och läsförståelse	114
Beskrivning av resultatet	114
Bedömning av evidensstyrkan	116

Icke-randomiserade studier	117
Beskrivning av studierna	117
Bedömning av effekter	119
Bedömning av evidensstyrka	119
Systematisk översikt om musik	120
Bedömning av evidensstyrkan	120
8. Alternativa verktyg för barn och ungdomar med dyslexi – resultat av granskningen	121
Beskrivning av ingående studier	122
Alternativa verktyg – beskrivning av resultatet	122
Bedömning av evidensstyrkan	122
9. Etiska och sociala aspekter	123
Inledning	123
Vad säger skollagen?	123
Resultat av litteratursökningen och urval av studier	124
Utredning för barn med dyslexi	124
Insatser för barn med dyslexi	124
Etiska principer	125
Forskningsetik	126
Yrkesetik för professioner involverade vid dyslexi	126
Gemensamma etiska principer	127
Kända etiska risker	128
Sociala aspekter	130
Låg självkänsla	131
Stödinsatser	131
Svårigheter i vuxen ålder	131
10. Hälsoekonomiska aspekter	133
Inledning	133
Frågeställningar	133
Resultat av litteratursökningen och urval av studier	134
Analyser och egna beräkningar	134
Kostnad för test och utredning	134
Kostnad för alternativa verktyg	135

Kostnad för insatser för att träna relationen mellan bokstäver och språkljud	135
Kostnadseffektivitet för insatser för att träna relationen mellan bokstäver och språkljud	137
Tidiga respektive sena insatser	140
Sammanfattningsvis	141
11. Inventering av tester som används i Sverige	143
Tillgängliga svenska tester	143
Användande av de olika testerna	144
Hur är det i Sverige idag?	145
12. Diskussion	153
Sammanfattning av resultaten	153
Litteratursökningen i förhållande till den skiftande terminologin	154
Metodfrågor	156
Förutsättningar, arbetssätt och vägval inom projektet	156
Att förutsäga dyslexi	156
Svenska och internationella tester	157
Interventionsstudierna	157
Övriga avgränsningar	158
Fördjupad resultatdiskussion	160
Att förutsäga dyslexi	160
Tester	161
Insatser	163
Alternativa verktyg	166
Jämförelser med resultat från andra översikter	167
Intervention for Dyslexia	167
Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties	167
13. Konsekvenser av rapportens bedömningar	169
Enhetlig utredning	169
Tester för att förutsäga dyslexi	169
Insatser	170
Resurser och utbildning	170
Framtida utmaningar för forskningen	171

14. Kunskapsluckor och angelägna forskningsområden	173
Tester för att förutsäga dyslexi	173
Tester för att upptäcka och utreda dyslexi	173
Insatser	174
Alternativa verktyg	174
Systematiska översikter saknas	175
Etik	175
Hälsoekonomi	175
15. Ordförklaringar och förkortningar	177
16. Personer som medverkat till rapporten	185
Projektgrupp	185
Extern referensexpert	186
Extern referensperson	186
Kansli	186
Externa granskare	186
Bindningar och jäv	187
17. Studier som ligger till grund för resultat och slutsatser	189
18. Referenser	319

SBU:s sammanfattning och slutsatser



SBU • Statens beredning för medicinsk utvärdering
Swedish Council on Health Technology Assessment

SBU:s sammanfattning och slutsatser

Denna rapport har inriktats på att utvärdera tester och insatser för barn och ungdomar med dyslexi. Då det inte finns en absolut gräns för läsförmåga (ordavkodning) hos personer med dyslexi sattes gränsen till den 10:e percentilen. Rapporten har gjorts på förfrågan från Socialstyrelsen och är en del av ett regeringsuppdrag.

SBU:s slutsatser

- ▶ Om barn med dyslexi får öva kopplingen mellan språkljud (fonem) och bokstäver (grafem) på ett strukturerat sätt, förbättras deras läsförmåga, stavning, läsförståelse, läshastighet och förmåga att uppmärksamma språkets ljudmässiga uppbyggnad (fonologisk medvetenhet).
- ▶ Det går inte att uttala sig om nyttan av andra former av läs- och skrivträning eller av alternativa verktyg (hjälpmedel för att stödja, kompensera och utveckla läsförmåga som t ex appar i mobiltelefonen). Metoderna är otillräckligt utvärderade.
- ▶ Det finns tester som kan förutsäga dyslexi redan innan barnen har fått undervisning i att läsa och skriva i skolan. Brister i fonologisk medvetenhet, snabb automatiserad benämningsförmåga samt bokstavskännedom har ett samband med dyslexi. Nyttan och eventuella risker med sådana tidiga testförfaranden har inte utvärderats i denna rapport. Insatser som riktas mot barn med sådana bristande förmågor innan de får lästräning har inte utvärderats i rapporten.
- ▶ I Sverige används mer än 50 olika tester för att upptäcka och utreda barn med dyslexi. Inget av testerna är vetenskapligt utvärderat dvs det saknas studier som undersöker om de är tillförlitliga och mäter det som avses.

Kunskapsluckor

- För att stödja barn och ungdomar med dyslexi behövs följande forskning:
 - Studier som utvärderar tillförlitligheten hos tester och etablerar en referensstandard för tester.
 - Studier som undersöker om läs- och skrivträning och alternativa verktyg förbättrar kunskapsutveckling och livskvalitet (självförtroende, självkompetens (self-efficacy), självständighet, självbild).
 - Randomiserade kontrollerade studier som utvärderar effekten av olika former av läs- och skrivträning och alternativa verktyg som används i Sverige.
 - Studier som följer barn under ett antal år och som undersöker sambandet mellan språkliga förmågor i tidig ålder och dyslexi.
 - Hälsoekonomiska studier som undersöker om det finns tester och insatser som är kostnadseffektiva.

Bakgrund och syfte

Dyslexi, även kallad specifika läs- och skrivsvårigheter/dyslexi, är en i regel medfödd funktionsnedsättning. Dyslexi yttrar sig framför allt som svårigheter med ordavkodning vid läsning och problem med stavning men den kan också, som en sekundär konsekvens, visa sig som problem med läsförståelse. En uppskattning är att 5–8 procent av befolkningen har dyslexi. Idag används ett stort antal metoder för att identifiera och stödja barn med dyslexi. Syftet med detta projekt har varit att systematiskt granska det vetenskapliga underlaget för metoder för att förutsäga, upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi hos barn och ungdomar samt att utvärdera vilka insatser som har vetenskapligt underlag. Vi har också belyst ämnesområdet ur en ekonomisk och etisk synvinkel samt identifierat kunskapsluckor för att stimulera till ny forskning. Vi har gjort en inventering av de tillgängliga

svenska testerna som avser att upptäcka och utreda barn och ungdomar med dyslexi.

Uppdraget

Efter en förfrågan från Socialstyrelsen har SBU genomfört denna systematiska litteraturöversikt. Översikten är en del av ett regeringsuppdrag för att uppdatera vägledningsdokumenten för barnhälsovård och elevhälsa. Rapporten berör beslutsfattare, myndigheter, brukarföreningar och andra som är engagerade i frågor som rör barn och ungdomar med dyslexi.

Metod

Vid granskningen och sammanställningen av den vetenskapliga litteraturen har samma metodik använts som i övriga SBU-projekt, för att göra urval av studier, bedöma studiernas kvalitet, väga samman resultaten och bedöma det vetenskapliga underlagets styrka.

Den systematiska litteratursökningen gjordes i fyra olika internationella databaser som innehåller originalartiklar inom bl a utbildning, pedagogik, psykologi, psykiatri, medicin, logopedi och lingvistik. Litteratursökningen gjordes t o m september år 2013. Varje inkluderad studie har kvalitetsgranskats och tabellerats enligt en särskilt utarbetad metodik. Styrkan på det vetenskapliga underlaget (evidensgraderingen) bedömdes med hjälp av det internationellt utarbetade systemet Grading of Recommendations Assessment Development and Evaluation (GRADE). Evidensstyrkan anger hur starkt det sammanlagda vetenskapliga underlaget är för att besvara en fråga på ett tillförlitligt sätt, dvs i vilken grad vi är säkra på vårt resultat. Ju högre slutlig evidensstyrka, desto större sannolikhet att det sammanvägda resultatet är stabilt över tid. I projektgruppen har åtta sakkunniga ingått med erfarenheter inom bl a specialpedagogik, psykologi, logopedi, lingvistik, medicin och etik. Projektet har också haft tillgång till två referensexperter som bidragit med erfarenheter om tester och hur de utvärderas. Därutöver har SBU:s kansli bidragit med hälsoekonomisk expertis. Projektet har dessutom anlitat fyra externa granskare med relevant kompetens.

Faktaruta 1 Studiekvalitet, evidensstyrka och slutsatser.

Studiekvalitet avser den vetenskapliga kvaliteten hos en enskild studie och dess förmåga att besvara en viss fråga på ett tillförlitligt sätt.

Evidensstyrka är en bedömning av hur starkt det sammanlagda vetenskapliga underlaget är för att besvara en viss fråga på ett tillförlitligt sätt. SBU tillämpar det internationellt utarbetade evidensgraderingssystemet GRADE. För varje effektmått utgår man i den sammanlagda bedömningen från studiernas design. Därefter kan evidensstyrkan påverkas av förekomsten av försvagande eller förstärkande faktorer (påverkansfaktorer), dvs studiekvalitet, samstämmighet, överförbarhet, effektstorlek, precision i data, risk för publikationsbias och andra aspekter, t ex dos-responssamband.

Evidensstyrkan graderas i fyra nivåer:

- **Starkt vetenskapligt underlag** (⊕⊕⊕⊕). Bygger på studier av god kvalitet som vid en samlad bedömning av påverkansfaktorer har starkt vetenskapligt stöd.
- **Måttligt starkt vetenskapligt underlag** (⊕⊕⊕○). Bygger på studier av god kvalitet som vid en samlad bedömning av påverkansfaktorer bedöms ha måttligt starkt vetenskapligt stöd.
- **Begränsat vetenskapligt underlag** (⊕⊕○○). Bygger på studier av god kvalitet som vid en samlad bedömning av påverkansfaktorer har begränsat vetenskapligt stöd.
- **Otillräckligt vetenskapligt underlag** (⊕○○○). När vetenskapligt underlag saknas, tillgängliga studier har låg kvalitet eller när studier av likartad kvalitet visar motsägande resultat, anges det vetenskapliga underlaget som otillräckligt.

Ju starkare evidens, desto mindre sannolikt är det att redovisade resultat kommer att påverkas av nya forskningsrön inom en överblickbar framtid.

Slutsatser innebär att man gör en sammanfattande bedömning av nytta, risker och kostnadseffektivitet.

Avgränsningar och vägval inom projektet

Vi studerade inte orsakerna till dyslexi och vi exkluderade studier om förvärvad dyslexi efter t ex skada eller sjukdom. Studier där någon annan problematik än dyslexi var det primära funktionshindret, såsom t ex ADHD, autismspektrumstörning, språkstörning eller utvecklingsstörning, exkluderades också.

Översikten kan delas in i tre huvudfrågor. Den första är om det finns tester som kan förutsäga dyslexi redan innan barnet har fått formell läs- och skrivträning (upp till cirka sex års ålder). Här gjordes en förut-sättningslös litteratursökning över alla tänkbara metoder och tester som kunde förutsäga dyslexi. Den andra frågan är hur tillförlitliga de tester är som används för att upptäcka och utreda dyslexi hos barn och ungdomar (6 till 20 år). Den tredje frågan är om det finns effektiva insatser för barn och ungdomar med dyslexi. Värdet av insatser före formell läs-träning i skolan till barn med risk för dyslexi ingick inte.

Studierna om att förutsäga dyslexi skulle vara publicerade åren 1980–2013, studierna om tester 1990–2013, studierna om insatser, etik, hälso-ekonomi 1980–2013 och studierna om alternativa verktyg (så kallade assistive technologies) 2000–2013. Studierna skulle vara publicerade på engelska, svenska, danska eller norska.

Tester

I rapporten ingår en inventering över vilka tester som förekommer i Sverige idag för att upptäcka och utreda dyslexi och som används i större eller mindre utsträckning. Totalt identifierade vi 52 tester. Även ett begränsat antal internationellt etablerade tester inkluderades i granskningen i samråd med internationellt välkända forskare inom dyslexi-området. En litteratursökning genomfördes på samtliga inkluderade tester. De studier som undersökte testerna granskades för att ta reda på deras vetenskapliga stöd. Dessutom gjorde vi en granskning av de svenska testernas manualer (några saknade manualer) för att samla information om testernas tillförlitlighet.

Insatser

Enbart kontrollerade studier inkluderades (randomiserade eller icke-randomiserade), som undersökte effekten av insatser. De grupper som jämfördes skulle bestå av minst 30 barn eller ungdomar med dyslexi. Någon exakt gräns för när en person har lässvårigheter av sådan art att det kan karaktäriseras som dyslexi finns inte. Därför var vi tvungna att sätta en gräns för vad som ansågs vara dyslexi för att göra projektet möjligt. Vi satte gränsen till de tio procenten av de mest lässvaga barnen och ungdomarna. Att vi satte just denna gräns styrdes av hur forskarna vanligen avgränsar gruppen personer med dyslexi. Vidare skulle studierna utvärdera effekterna av korrekt läsning av ord, korrekt läsning av non-ord (ord utan betydelse), stavning, fonologisk medvetenhet, läshastighet, läsförståelse, självkompetens, självbild, självförtroende, självständighet och kunskapsutveckling. Studierna skulle ha en kontrollgrupp bestående av elever med dyslexi, som antingen fick enbart ordinarie undervisning eller annan definierad åtgärd riktad mot läs- och skrivsvårigheterna.

Resultat

Vi gick igenom 12 743 artikelsammanfattningar. Av de 1 225 artiklar som sedan beställdes i fulltext var 93 relevanta för någon av projektets tre frågeställningar och 59 hade tillräckligt hög kvalitet för att ingå i det vetenskapliga underlaget för evidensgradering. Evidensgraderingen är sammanfattad i tabellerna 1–3.

Metoder för att förutsäga dyslexi

Metoder som mäter brister i fonologisk medvetenhet, snabb automatiserad benämningsförmåga (RAN) och bokstavskänedom verkar kunna förutsäga dyslexi i tidig ålder. Det finns begränsat vetenskapligt underlag för att brister i fonologisk medvetenhet och i RAN kan förutsäga dyslexi. Det finns måttligt vetenskapligt underlag för att brister i bokstavskänedom kan förutsäga dyslexi. De metoder som mäter brister i andra förmågor som är karaktäristiska för dyslexi är otillräckligt studerade.

Tester

Det går inte att avgöra vilka svenska testmetoder som är tillräckligt bra för att upptäcka och utreda dyslexi bland barn och ungdomar då det saknas studier som utvärderar testernas tillförlitlighet. Många av de etablerade internationella testerna är också otillräckligt utvärderade. Undantaget är enstaka deltester i testbatterierna Dynamic Indicators of Basic Early Literacy Skills (DIBELS) och Woodcock-Johnson (WJ).

Insatser

Det finns en form av insats för barn och ungdomar med dyslexi som har vetenskapligt underlag. Den kallas phonics på engelska. Vi har inte motsvarande term på svenska, men phonics innebär bl a att barnet tränar sambandet mellan fonem (språkljud) och grafem (bokstäver) på ett strukturerat sätt. Ofta börjar träningen på en mycket grundläggande nivå med fonemisk medvetenhet och en enkel fonem–grafem-koppling för att successivt övergå till mer avancerade principer om skriftspråket. Insatsen har visat sig stärka barnets förmåga att läsa korrekt och snabbt samt förbättra stavning, läsförståelse och fonologisk medvetenhet. Hädanefter kommer vi att kalla denna typ av insats för strukturerad fonem–grafem-koppling. För övriga insatser (andra typer av träning, musikterapi, alternativa verktyg) är det vetenskapliga underlaget otillräckligt. Det saknas studier där elevernas livskvalitet (självförtroende, självkompetens, självständighet, självbild) och kunskapsutveckling utvärderas.

Det är viktigt att poängtera att rapportens resultat gäller för grupper av barn och ungdomar. Forskningsresultat på gruppnivå kan ge viktig information för den enskilde, men kan aldrig ersätta en individuell bedömning eftersom svårigheterna vid dyslexi kan variera.

Evidensgraderande resultat

- Det finns begränsat vetenskapligt underlag för att brister i fonologisk medvetenhet kan förutsäga dyslexi (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att brister i RAN kan förutsäga dyslexi (⊕⊕○○).
- Det finns måttligt starkt vetenskapligt underlag för att brister i bokstavskänedom kan förutsäga dyslexi (⊕⊕⊕○).

Tabell 1 Evidensgraderade resultat; metoder för att förutsäga dyslexi hos barn och ungdomar.

Metoder för att förutsäga dyslexi	Brister i fonologisk medvetenhet	Brister i snabb automatiserad benämningsförmåga (RAN)	Brister i bokstavskänedom
Vetenskapligt underlag	Begränsat (⊕⊕○○)	Begränsat (⊕⊕○○)	Måttligt starkt (⊕⊕⊕○)

- Det finns otillräckligt vetenskapligt underlag för att avgöra vilka svenska tester som är tillförlitliga för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi bland barn och ungdomar, då det saknas relevanta studier (⊕○○○).
- Det finns begränsat vetenskapligt underlag för enstaka deltester inom Dynamic Indicators of Basic Early Literacy Skills (DIBELS) och Woodcock-Johnson (WJ) (⊕⊕○○).

Tabell 2 Evidensgraderade resultat; tester för att upptäcka och utreda dyslexi hos barn och ungdomar.

Tester	Tillgängliga svenska tester ¹	Etablerade internationella tester ²
Vetenskapligt underlag	Otillräckligt (⊕○○○)	Otillräckligt för testerna som helhet (⊕○○○) Begränsat för enstaka deltester (⊕⊕○○)

¹ 52 tester som förekommer i Sverige utvärderades i översikten

² 18 internationella tester utvärderades i översikten

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra korrekt läsning av ord (⊕⊕○○).
- Det finns måttligt starkt vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en stor effekt på att förbättra korrekt läsning av non-ord (⊕⊕⊕○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra stavning (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra läsförståelse (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en liten effekt på att förbättra fonologisk medvetenhet (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en liten effekt på att förbättra läshastighet (⊕⊕○○).

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning efter sex till tolv månader efter avslutad insats har en liten effekt på att förbättra läsförståelse (⊕⊕○○).
- Det finns otillräckligt vetenskapligt underlag för att träning av enbart fonologisk medvetenhet (utan annan koppling till skriftspråket), jämfört med ordinarie undervisning har effekt på att förbättra korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att träning av enbart läsflyt jämfört med ordinarie undervisning har effekt på att förbättra läsning av ord på tid samt läsförståelse (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att avgöra om musik, jämfört med kontrollgrupp (väntelista eller ingen insats), förbättrar utfallsmåtten (läsning, läsförståelse, läsflyt, fonologisk medvetenhet och stavning) då det saknas relevanta studier (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att avgöra om det finns effekt av alternativa verktyg då det saknas studier (⊕○○○).

Tabell 3 Evidensgraderade resultat; insatser för barn och ungdomar med dyslexi.

Insatser → Utfallsmått ↓	Strukturerad träning av sambandet mellan språkljud (fonem) och bokstäver (grafem)	Träning av enbart fonologisk medvetenhet	Träning av enbart läsflyt (fluency)	Musik	Alternativa verktyg t ex mobilappar mm
Korrekt läsning av ord	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Korrekt läsning av non-ord (ord utan betydelse)	Stor förbättring (⊕⊕⊕○)	Bristande tillförlitlighet (⊕○○○)	Bristande tillförlitlighet (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Stavning	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läsförståelse	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Bristande tillförlitlighet (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läshastighet	Liten förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Fonologisk medvetenhet	Liten förbättring (⊕⊕○○)	Bristande tillförlitlighet (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läsförståelse efter uppföljning (6–12 månader) efter avslutad insats	Liten förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)

Tabellen fortsätter på nästa sida

Tabell 3 fortsättning

Insatser → Utfallsmått ↓	Strukturerad träning av sambandet mellan språk- ljud (fonem) och bokstäver (grafem)	Träning av enbart fonologisk medveten- het	Träning av enbart läsflyt (fluency)	Musik	Alternativa verktyg t ex mobil- appar mm
Livskvalitets- mått t ex själv- kompetens, självständighet, självbild, själv- förtroende, kunskaps- utveckling	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)

(⊕⊕⊕○): Måttligt starkt vetenskapligt underlag

(⊕⊕○○): Begränsat vetenskapligt underlag

(⊕○○○): Otillräckligt vetenskapligt underlag

Etiska och sociala aspekter

Såväl utredningar av barn med läs- och skrivsvårigheter som insatser för barn som fått diagnosen dyslexi är förknippade med etiska risker. En dyslexiutredning är ofta påfrestande och omfattningen av testerna måste anpassas till det enskilda barnets reaktioner. Barnet kan utsättas för intrång i sin integritet om han eller hon genomgår många tester trots känslor av obehag eller olust. En utredning väcker förhoppning om snar hjälp som inte alltid infrias beroende på t ex resurs- eller kompetensbrist. Även när en utredning är genomförd och barnet har fått ett intyg kan insatserna dröja och ibland t o m helt utebli. Föräldrar och lärare kan också ha olika uppfattning om en utredning ska påbörjas tidigt eller vänta. I sådana fall har utredaren ett etiskt ansvar att ge vägledning till ett beslut. Utredarens kompetens är viktig och avgörande. Bristande kunskap och felanvända metoder utgör skaderisker. Kravet på kompetens gäller också för de som är ansvariga för att barnet sedan får adekvat pedagogisk hjälp.

Rätten till integritet innebär att det är elever och föräldrar som avgör i vilken grad diagnosen ska vara privat eller offentlig. Här måste man ta hänsyn till att barn och föräldrar kan ha olika uppfattning om t ex risken för stigmatisering. Det ställer krav på flexibilitet hos lärare och annan skolpersonal. Rätten till självbestämmande ställer krav på personalen i samband med såväl utredning som åtgärder. Elever och föräldrar kan ha olika önskemål om insatserna och det kan vara svårt att avgöra vems självbestämmande som ska gälla.

Skollagen anger *Om en utredning visar att en elev är i behov av särskilt stöd ska han eller hon ges sådant stöd* (3 kap 8 §). I skolor med resurs- eller kompetensbrist finns risk att elever med stora och outredda svårigheter får mindre hjälp eller ingen hjälp alls. Tillgången till kompetens är en rättvisefråga. Bristande eller varierad utbildning kan leda till ojämlikhet i diagnostik och behandling för elever med förmodad dyslexi. Olika kommuner prioriterar olika enligt det kommunala självstyret.

Etiska och sociala aspekter är sammanflätade och överlappande. De sociala aspekterna i samband med dyslexi rör elevernas sociala liv både i skolan och på fritiden. Det gäller kamratrelationer och tillgång till fritidsaktiviteter, där deras svårigheter inte ska vara hindrande. På längre sikt gäller det vuxenliv och deltagande i t ex utbildning, arbetsliv och samhällsliv.

Risken för att barn med dyslexi utvecklar lägre självkänsla och i högre grad ängslan än andra barn har uppmärksammats och studerats. En officiell diagnos tycks ha positiv betydelse för självkänslan och minskar ängslan.

Upplevelsen av otillräcklighet och att betraktas som okunnig och inkompetent följer ofta med från skolåren in i vuxen ålder och i arbetslivet. Att låg självkänsla riskerar att finnas kvar, bekräftas i longitudinella studier. En tidig och väl utförd utredning och diagnos samt pedagogiska insatser tycks ge förutsättningar för högre självkänsla som vuxen.

Hälsoekonomi

Det finns otillräckligt vetenskapligt underlag (studier saknas) för att bedöma kostnadseffektiviteten för tester och insatser som ges till barn och ungdomar med dyslexi.

Konsekvenser av rapportens resultat

Dyslexi medför konsekvenser såväl för den enskilde individen som för samhället. Denna rapport visar att det finns ett vetenskapligt underlag för att brister i fonologisk medvetenhet, automatiserad benämningsförmåga och bokstavskänneteknik kan förutsäga dyslexi innan barnen fått möjlighet att lära sig läsa och skriva i samband med skolstarten. Rapporten ger därmed kunskap om hur man tidigt kan uppmärksamma de barn som skulle ha extra nytta av strukturerad träning i sambandet mellan fonem och grafem.

I dag saknas tydliga riktlinjer för hur barn och ungdomar ska utredas vid misstanke om dyslexi.

Vi har gjort en inventering över vilka tester som går att få tag på i Sverige idag och som förekommer i större eller mindre utsträckning för att identifiera och utreda dyslexi hos barn och ungdomar. Sammanställningen har gjorts tillsammans med bl a Specialpedagogiska skolmyndigheten som bidragit med sin förteckning över svenska tester. Inventeringen visar att det idag förekommer över 50 olika kartläggnings- och utredningstester runt om i landet. Vissa av dessa har en teoretisk förankring men är inte vetenskapligt utvärderade. För att kunna prioritera mellan testerna behövs forskning som undersöker testernas tillförlitlighet. Då kan man i framtiden säkerställa att utredningarna görs med relevanta tester med god tillförlitlighet som mäter det som avses. Det skulle leda till ett mer enhetligt utredningsförfarande i landet och sannolikt vara kostnadseffektivt för samhället om alla barn och ungdomar kartlades och utreddes på ett likartat sätt.

Rapporten ger vetenskapligt stöd för att strukturerad träning av fonem–grafem-koppling har positiv effekt på att förbättra läsförmåga, stavning, läsförståelse, läshastighet och fonologisk medvetenhet jämfört med ordi-

narie undervisning. Resultatet av vår granskning visar på vikten av insatser där strukturerad träning av fonem–grafem-koppling ingår i skolundervisningen för barn med dyslexi. Det är dock viktigt att beakta att nyttan av enbart en träningsinsats oftast inte hjälper alla barn och ungdomar med dyslexi.

Kunskapsluckor och angelägna forskningsområden

Vi har identifierat ett antal kunskapsluckor som beror på avsaknad av studier, eller på att det vetenskapliga underlaget i studierna var bristfälligt. Några av de mest angelägna forskningsområdena är följande:

Metoder för att förutsäga dyslexi

- Det behövs fler studier av tillräcklig kvalitet som följer barnen på lång sikt för att identifiera metoder som kan förutsäga dyslexi. Det är viktigt att framtida studier följer ett stort antal barn över en längre tid, gärna flera år, och där sambandet mellan testmetoder och dyslexi redovisas på ett tydligt sätt.

Tester

- Det saknas vetenskapligt underlag för att avgöra vilka svenska tester som är tillräckligt bra för att upptäcka och utreda dyslexi bland barn och ungdomar. Det behövs studier som undersöker de enskilda testernas tillförlitlighet. En orsak till denna kunskapslucka kan vara att det inte finns en tradition inom fältet att utvärdera och publicera forskning om testernas prestanda. Dessutom saknas ett bästa möjliga test (referenstest eller så kallad gold standard) att jämföra testerna med. Det är därför angeläget att etablera en referensstandard för tester. Det behövs även mer forskning för att utvärdera de internationella testerna.

Insatser

- Det behövs mer forskning med hög vetenskaplig kvalitet (randomiserade kontrollerade studier) för att undersöka nyttan av tänkbara insatser för barn och ungdomar med dyslexi. Det behövs även forskning som studerar barns och ungdomars livskvalitet (självförtroende, självkompetens, självständighet, självbild) och kunskapsutveckling.
- Det behövs fler studier som undersöker nyttan av alternativa verktyg, t ex tekniska hjälpmedel såsom appar/program för mobiltelefoner och surfplattor med avsikt att främst kompensera och stödja läs- och skrivförmågan för barn och ungdomar med dyslexi.

Etiska och sociala aspekter

- För att bistå de professionella som utreder och ställer diagnos behövs studier som utvärderar hur mötena mellan professionella och elever och föräldrar går till, samt att de tillfrågas om sina erfarenheter.

Hälsoekonomi

- Det behövs hälsoekonomiska studier för att utvärdera vilka tester och insatser som är kostnadseffektiva för barn och ungdomar med dyslexi.

1. Inledning

Dyslexi, även kallad specifika läs- och skrivsvårigheter, är en i regel medfödd funktionsnedsättning som innebär stora svårigheter att lära sig avkoda skrivna ord. Det betyder att kärnproblemet vid dyslexi är dålig ordavkodning som orsakas av en långsam omvandling av bokstäver och bokstavskombinationer till språkljud.

Ord utan betydelse (så kallade non-ord) och nya ord är särskilt svåra att läsa för elever med dyslexi. Stavningsproblem är också vanliga. Dyslexin kan sekundärt yttra sig som läsförståelseproblem. En uppskattning som vanligen görs i forskning och i kliniska sammanhang är att 5–8 procent av befolkningen har svårigheter av sådan art att de uppfyller kriterierna för dyslexi [1]. Att tidigt upptäcka och sätta in insatser för barn som löper risk att få dyslexi kan ha stor betydelse för deras kunskapsinhämtning och livskvalitet. Idag används ett stort antal metoder för att identifiera och stödja barn med dyslexi, men det är oklart om det finns tester och insatser som har vetenskapligt underlag och i så fall vilka de är. Det finns behov av en systematisk granskning av det vetenskapliga underlaget för sådana tester och insatser.

Efter en förfrågan från Socialstyrelsen har SBU genomfört en systematisk litteraturöversikt över de tester och insatser som används och som är vetenskapligt dokumenterade inom dyslexiområdet. Litteraturöversikten har gjorts i enlighet med SBU:s metodbok [2]. Rapporten riktar sig till beslutsfattare, myndigheter, brukarföreningar och andra som är engagerade i frågor som rör barn och ungdomar med dyslexi. I rapporten används genomgående benämningen dyslexi för att beteckna specifika läs- och skrivsvårigheter/dyslexi.

Syfte

Syftet med projektet har varit att granska det vetenskapliga underlaget för att förutsäga, upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi samt att utvärdera vilka insatser för barn och ungdomar med dyslexi som har vetenskapligt underlag. Vi har också belyst ämnesområdet ur en ekonomisk och etisk synvinkel och identifierat kunskapsluckor för att stimulera till ny forskning. Vi har gjort en inventering av tillgängliga svenska tester som används för att upptäcka och utreda barn och ungdomar med dyslexi.

Målgrupper

Rapporten kommer att vara ett underlag för Socialstyrelsens vägledningsdokument för barnhälsovård och elevhälsa. Resultaten är av intresse för Social- och Utbildningsdepartementen och för många andra aktörer i samhället som är i behov av kunskapsstöd för barn med dyslexi.

Exempel på sådana målgrupper är:

- Myndigheter som bedriver verksamhet inom utbildningsområdet, såsom Skolverket, Specialpedagogiska skolmyndigheten, Skolinspektionen, Universitetskanslersämbetet (UKÄ) och Universitets- och högskolerådet (fd Högskoleverket)
- Kommuner (skolförvaltning samt skolpolitiker)
- Sveriges Kommuner och Landsting (SKL), Hälso- och sjukvårdsförvaltningen – avdelningen för närsjukvård
- Lärare och studenter vid lärarutbildningar inklusive påbyggnadsutbildningar till speciallärare och specialpedagog
- Logopedier och logopedutbildningar
- Psykologer och psykologutbildningar

- Personal inom skolväsendet, barnhälsovården (BHV) och skolhälsovården (SHV)
- Professions- och intresseföreningar som t ex Skolläkarföreningen, Riksföreningen för skolsköterskor, specialpedagoger, skolpsykologer och kuratorer inom skolhälsovård, Svenska Dyslexiföreningen, Sveriges Skolledarförbund, Svenska Förbundet för Specialpedagogik, Svenska logopedförbundet (SLOF) inom Dokumentation, Information, Kultur (DIK), Svenska Föreningen för Foniatri och Logopedi (SFFL), Psykologförbundet, Läraförbundet, Lärarnas Riksförbund, Svenska Skolläkarföreningen
- Brukarorganisationer som t ex Dyslexiförbundet FMLS, Föräldraföreningen för Dyslektiska Barn (FDB)
- Trafikverket
- Kriminalvården
- Arbetsförmedlingen.

2. Bakgrund

Begrepp och definitioner

Dyslexi är en funktionsnedsättning

I denna rapport har vi valt att följa den definition som föreslogs av Høien & Lundberg och som sedan kom att användas av Svenska Dyslexistiftelsen, då med tillägget om svårigheter med läsförståelse som en sekundär konsekvens [3]. Här har vi valt att byta ut ordet *störning* till *svikt*:

Dyslexi är en svikt i vissa språkliga funktioner, särskilt de fonologiska (fonologi avser språkets ljudmässiga form), som är viktiga för att kunna utnyttja skriftens principer för kodning av språket. Svikten ger sig först och främst tillkänna som svårigheter att uppnå en automatiserad ordavkodning vid läsning. Men den kommer också tydligt fram genom dålig stavning. Sekundära konsekvenser kan innefatta svårigheter med läsförståelse och begränsad läserfarenhet, vilket kan hämma tillväxten av ordförråd och kunskap om omvärlden. Den dyslektiska svikten går i regel igen i släkten och man har anledning att anta en genetisk disposition som kan medföra neurobiologiska avvikelser. Karakteristiskt för dyslexi är att svikten är varaktig och svårbehandlad. Även om läsningen efterhand kan bli acceptabel, kvarstår oftast stavningsproblemen. Vid en mer grundläggande kartläggning av fonologisk förmåga finner man att svagheter på detta område ofta kvarstår upp i vuxen ålder [1].

Dyslexi orsakas vanligtvis av svårigheter att hantera språkljud (fonem), dvs fonologiska problem. Den primära manifestationen av dyslexi är bristande ordavkodning [4,5]. Dyslexi är i regel ärftlig och barn till föräldrar med dyslexi löper en större risk att få dyslexi. Resultatet från tvillingstudier pekar mot att den genetiska risken är cirka 50–55 procent [6,7]. Svårigheterna visar sig oftast i samband med den formella läs- och skrivundervisningen vid skolstart.

Dyslexi har definierats på olika sätt under åren

Redan i slutet av 1800-talet publicerade Morgan de första, mer detaljerade beskrivningarna av personer med dyslexi [8]. I rapporterna beskrevs en pojke i skolåldern som i de flesta andra avseenden var som sina andra jämnåriga, men som hade oväntade svårigheter med att lära sig läsa och skriva. Läs- och skrivsvårigheterna uppmärksammades först av ögonläkare som noterade läs- respektive stavfel. Sådana svårigheter innebar ofta spegelvändningar av korta ord vid läsning eller att bokstäver utelämnades vid stavning. Detta bidrog till att tidiga teorier om orsaker till dyslexi handlade om brister i syn eller hur hjärnan lagrade ordbilder. Långt in på 1900-talet var det därför vanligt att dyslexi kallades för ordblindhet [9,10]. Under det kommande halvsekleet bedrevs forskning kring dyslexi endast i blygsam omfattning. I mitten på 1970-talet intensifierades forskningen och från att ha förklarat dyslexi som ett visuellt problem stod det alltmer klart att dyslexi förefaller bero på brister i språklig medvetenhet och bearbetning, dvs förmågan att reflektera över det egna språket och framför allt beträffande språkets minsta betydelseskiljande delar, fonem [11,12].

Den första definitionen av dyslexi gjorde World Federation of Neurology år 1968. I den betonades att dyslexi främst yttrar sig som svårigheter att lära sig läsa även om personen fått ordinarie läs- och skrivundervisning och har en åldersadekvat intellektuell förmåga samt rika möjligheter att praktisera läsning. Liknande definitioner dominerade dyslexiforskningen under 1980- och 1990-talen.

I takt med en allt större samstämmighet bland forskare om att brister i fonologisk medvetenhet och bearbetning skulle vara den huvudsakliga orsaken till dyslexi, uppstod dock ett behov av att ompröva definitionen från 1968. En anledning var att World Federation of Neurologys definition inriktades på vad som inte var dyslexi, en annan var att det i definitionen skulle finnas en diskrepans mellan läsförmåga och intelligens (det så kallade diskrepanskriteriet). I praktiken betydde det att dyslexi förutsatte en påtaglig skillnad mellan läsförmåga (uppmätt med ett standardiserat avkodningstest) och intellektuell nivå (bestämd med hjälp av ett intelligenstest). Med andra ord skulle det finnas en skillnad mellan läsålder och intelligens. Forskning från mitten av 1990-talet visade att

barnets intelligens inte kunde förutsäga hur läs- och skrivförmågan skulle utvecklas under de första åren i skolan. En allvarlig invändning mot diskrepanskriteriet var att läs- och skrivinterventioner hade lika stor effekt oavsett om elever med läsproblem uppfyllde diskrepanskriterierna eller inte [13,14].

Dyslexi – en fråga om gränsdragning

En till två elever per skolklass beräknas i genomsnitt ha dyslexi. Det är dock viktigt att konstatera att fonologisk förmåga å ena sidan och läs- och stavningsförmåga å andra sidan, är färdigheter som är normalfördelade i befolkningen. Det betyder att det inte finns någon exakt gräns för vad som anses vara dyslexi. Vidare förekommer individuella variationer hos de barn och ungdomar som har dyslexi. Det skulle alltså kunna vara så att en person har en inte alltför dålig ordavkodningsförmåga i förhållande till jämnåriga. Personen kan dock likväl ha en underliggande fonologisk nedsättning som förhindrar en full automatisering av läsningen. Denna nedsättning kan vara ett hinder för just denna person i t ex skolsituationen, utan att personen för den skull tillhör de allra sämst presterande. I en skolsituation kan därför någon med sådana problem behöva stöd av olika slag för att kunna nå sin fulla potential. I definitionen av dyslexi som vi tillämpar i denna rapport är svårigheter i ordavkodning och/eller ordigenkänning de viktigaste kriterierna. För att läsa om detta projekts val av gräns för dyslexi, se Kapitel 3.

Samförekomst med andra funktionsnedsättningar

Personer med dyslexi har inte sällan även andra funktionsnedsättningar som försvårar den tidiga läs- och skrivutvecklingen. En studie av Willcutt och medarbetare visade att mellan 15 och 40 procent av alla elever med dyslexi också hade ADHD-liknande problem [15]. Av dessa barn och ungdomar med ADHD-diagnos hade cirka 25 procent i varierande grad även matematiska svårigheter [15–17]. Dessa siffror måste dock tolkas med viss försiktighet. Det kan t ex förhålla sig så att elever med dyslexi, pga sina svårigheter, utvecklar ett beteende som kan påminna om beteendet hos en person med ADHD. Man kan också före-

ställa sig att en elev med ADHD utvecklar svårigheter med skriftspråket pga svårigheter att ta till sig läs- och skrivundervisning.

Identifiering av dyslexi

Många studier har påvisat att det finns goda möjligheter att förutsäga dyslexi och att preventiva åtgärder väsentligt minskar risken för framtida svårigheter med läs- och skrivinläringen. Dyslexi manifesterar sig vanligen i samband med den första läs- och skrivundervisningen, även om den ibland kan bli uppenbar långt senare. Den tidiga läs- och skrivundervisningen syftar främst till att eleverna lär sig att koda av bokstäver och koppla samman dem med språkljud vid läsning och att segmentera ord i språkljud vid stavning. Det är viktigt att skapa förutsättningar för eleverna att automatisera dessa grundläggande färdigheter så att avkodningen inte bara blir korrekt utan också kan utföras flytande och snabbt. I vår rapport har vi utgått från att en elev måste ha fått minst ett års formell läs- och skrivundervisning innan det är rimligt att fastställa om det föreligger dyslektiska läs- och skrivsvårigheter. Detta innebär inte att skolan bör avvakta tills en elev misslyckas med den tidiga läsningen, utan stöd bör sättas in i ett så tidigt skede som möjligt.

Sammanfattningsvis kan vi konstatera att det är svårt att på ett exakt och precist sätt definiera dyslexi. Vi har utgått från en preliminär definition som har vuxit fram under decennier och som det idag råder en stor konsensus om, både nationellt och internationellt. Att vi definitionsmässigt lägger stor vikt vid fonologisk förmåga i rapporten utesluter inte att ytterligare faktorer ofta finns med i den dyslektiska problembilden [18–20]. Det är förhållandevis vanligt att barn och ungdomar med dyslexi också har neuropsykiatriska, språkliga och kognitiva funktionsnedsättningar som kan påverka läs- och skrivutvecklingen.

Förutsättningar och nulägesbeskrivning

Det råder oenighet om hur begreppet dyslexi ska tillämpas i praktiken världen över. Forskare och praktiker från olika länder använder sig av en skiftande terminologi där olika avgränsningar görs för vad som anses vara dyslexi. I fråga om differentialdiagnostik råder betydande oenighet

inom området läs- och skrivsvårigheter. Olika utredningar, såsom den av Skolinspektionen (2011:08), har visat att det också i Sverige råder stor variation av metoder för kartläggning av och insatser vid dyslexi [21]. Mot denna bakgrund har SBU fått i uppdrag att sammanställa den kunskap som finns idag om tester och insatser för barn och ungdomar med dyslexi.

SBU:s uppdrag och oberoende roll

Behovet av aktuell kunskap om dyslexi hos barn och ungdomar är stort. Brist på kunskap kan äventyra barns och ungdomars möjligheter att få en adekvat diagnostisering och tillgång till evidensbaserade insatser. Skollagen från år 2010 uttrycker tydligt att skolans undervisning ska vila på vetenskaplig grund. Det betyder att det ska finnas stöd i forskningen för de insatser skolan använder och de kunskaper som lärs ut.

SBU har en oberoende roll och sammanställer kunskap systematiskt samt granskar kritiskt den forskning som finns på området. Det är sedan Socialstyrelsen som tillsammans med Skolverket kommer att använda denna nya kunskap för att uppdatera de vägledningsdokument som ligger till grund för implementering.

Jävsproblematik

SBU:s kunskapsöversikter ska vara objektiva och trovärdiga. Det ska inte råda något tvivel om kunskapsunderlagets vetenskapliga trovärdighet. Projektets experter har, i enlighet med SBU:s krav, lämnat deklARATIONER om bindningar och jäv. I dessa framgår bl a om experterna har något kommersiellt intresse i de utvärderade testerna och insatserna. Dokumenten finns tillgängliga på SBU:s kansli. I projektgruppen finns experter som är medförfattare till svenska tester och träningsprogram och som har publicerat artiklar inom dyslexiområdet. De har inte deltagit i bedömningarna av underlagen för de svenska testerna och inte heller har de bedömt artiklar där de själva varit medförfattare eller på annat sätt varit jäviga.

3. Metodbeskrivning

Rapporten är en systematisk genomgång av den vetenskapliga litteraturen för att utvärdera tester och insatser för barn och ungdomar med dyslexi. Vi har använt den metodik som beskrivs i SBU:s metodbok för urval av studier, bedömning av studiernas kvalitet, sammanvägning av resultat och bedömning av det vetenskapliga underlagets styrka [2].

Frågor

Projektet omfattade fem övergripande frågor:

- Vilka testmetoder finns som kan förutsäga dyslexi i tidig ålder?
- Vilka testmetoder är tillförlitliga, dvs har vetenskapligt underlag, för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi bland barn och ungdomar?
- Vilka insatser förbättrar läsförmåga och livskvalitet för barn och ungdomar med dyslexi?
- Vilka etiska och sociala aspekter finns för tester och insatser för barn och ungdomar med dyslexi?
- Vilka tester och insatser för barn och ungdomar med dyslexi är kostnadseffektiva?

Avgränsningar

Projektet har haft många inklusions- och exklusionskriterier, bl a avseende population och utfallsmått (se detaljerad beskrivning Bilaga 3.1–3.4).

Originalstudier och systematiska översikter skulle vara på engelska eller skandinaviska språk. Studier om förvärvad dyslexi t ex efter skada eller sjukdom exkluderades. Studier med annan problematik än en primär dyslexi, såsom ADHD, autismspektrumstörningar, språkstörningar eller utvecklingsstörningar exkluderades. Värdet av insatser före formell läs- och skrivträning i skolan till barn med risk för dyslexi har inte ingått.

Inklusionskriterier för tester för att förutsäga dyslexi

- Studier som undersökte om ett test/deltest var tillförlitligt för att förutsäga läs- och skrivsvårigheter, karaktäristiska för dyslexi, bland barn och ungdomar
- Studier som definierade dyslexi enligt en gräns på 1,0 standardavvikelse, vilket motsvarar cirka 16 procent av en ålderskohort (se stycket om ”Val av gräns för dyslexi”)
- Studier inriktade på barn mellan 0 och 6 år i länder som brukar det alfabetiska skriftspråket
- Studier av longitudinell karaktär, för att ge barnen en chans att lära sig läsa och skriva under ett års tid innan läs- och skrivförmågan mättes
- Studier som undersökte samband mellan prediktor och dyslexi, uttryckt som oddskvot, sensitivitet, specificitet eller korrekt klassificerade barn
- Studier publicerade åren 1990–2013.

Inklusionskriterier för tester för att upptäcka och utreda dyslexi

- Studier som undersökte om ett test var tillförlitligt för att upptäcka och utreda dyslexi hos barn och ungdomar
- Studier inriktade på personer mellan 6 och 20 år som brukar det alfabetiska skriftspråket
- Studier som utvärderade svenska tester (Tabell 11.1, Kapitel 11) och internationella tester (Tabell 6.1, Kapitel 6)
- Studier som använde ett referenstest som inkluderades bland de svenska eller internationella testerna
- Studier som i första hand angav sensitivitet och specificitet som utfallsmått och i andra hand validitet eller reliabilitet
- Studier publicerade åren 1990–2013.

Inklusionskriterier för insatser

- Studier som undersökte insatser för barn och ungdomar med dyslexi
- Studier som definierade dyslexi enligt en gräns på 1,25 standardavvikelser, vilket motsvarar cirka 10 procent av en ålderskohort (se stycket om ”Val av gräns för dyslexi”)
- Studier inriktade på personer med dyslexi mellan 7 och 20 år och som brukar det alfabetiska skriftspråket
- Enbart kontrollerade studier (randomiserade eller icke-randomiserade) där de jämförda grupperna bestod av minst 30 barn eller ungdomar med dyslexi
- Studier som mätte något av följande utfallsmått: antal korrekt lästa ord, antal korrekt lästa non-ord (ord utan betydelse), stavning, fonologisk medvetenhet, läshastighet, läsförståelse, kunskapsutveckling

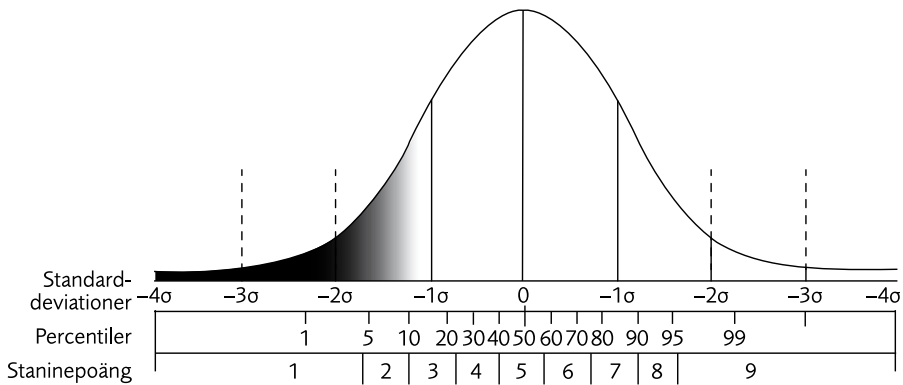
och livskvalitet (självbild, självkompetens, självförtroende, självständighet)

- Studier med en kontrollgrupp bestående av elever med dyslexi som inte fått någon åtgärd alternativt någon annan definierad åtgärd
- Studier som undersökte nyttan av alternativa verktyg, så kallade assistive technologies, såsom läspennor, talsynteser och appar i mobiltelefonen
- Studier publicerade åren 1980–2013 (för alternativa verktyg 2000–2013).

Val av gräns för dyslexi

Det finns inte någon etablerad gräns för hur stora fonologiska problem som krävs för att de ska karaktäriseras som dyslexi. Interventionsstudier utgår sällan från graden av fonologiska problem för att definiera barn med dyslexi. För att kunna avgöra vilka studier som skulle uppfylla våra inklusionskriterier beslutade vi att istället utgå från läsförmåga som mått på dyslexi. Vi satte därför en gräns för vad som kunde anses vara normal läsförmåga. Om den genomsnittliga läsförmågan avvek mer än 1,25 standardavvikelse (dvs 10:e percentilen) i förhållande till jämförelsegruppen eller normen vid baslinjen, ansågs barnen ha dyslexi (Figur 3.1). Läsförmågan bedömdes med hjälp av tester som mätte ordavkodning, non-ordläsning och/eller läshastighet av enstaka ord/non-ord alternativt sammanhängande texter samt stavning. Läsförståelsetesterna, utan krav på hastighet, tilläts avvika betydligt mindre från medelvärdet eller inte avvika alls. Om studierna redovisade stavningsförmåga eller läsförståelse, förutom läsförmåga, bidrog de till att säkerställa definitionen av dyslexi.

Konsekvensen av avgränsningen blev att studierna inkluderade barn och ungdomar med allvarliga svårigheter att koda av skrivna ord. I de studier där det var svårt att avgöra om alla barn uppfyllde våra kriterier, beslutade projektgruppen gemensamt om de skulle gå vidare i granskningsprocessen eller exkluderas. Att det blev just denna gräns styrdes av hur forskarna vanligen avgränsar gruppen personer med dyslexi.



Figur 3.1 Spridning av läsförståelse för hela befolkningen kan beskrivas med hjälp av normalfördelningskurvan [22].

Val av tester för granskning

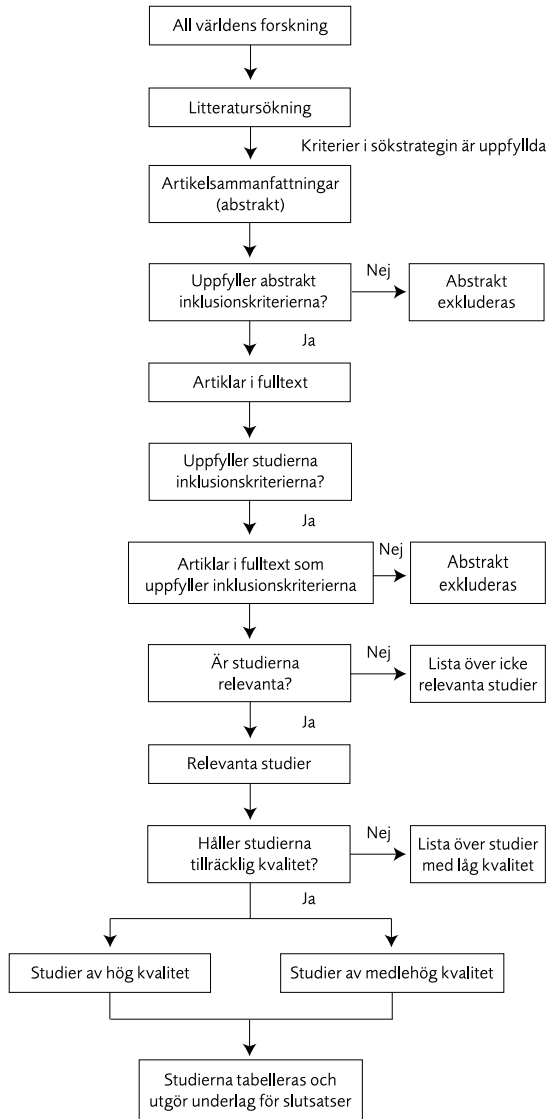
I rapporten ingår en inventering över vilka tester som förekommer i Sverige för att upptäcka och utreda dyslexi, och som används i större eller mindre utsträckning. Totalt identifierade vi 52 tester (Kapitel 11, Tabell 11.1). Även ett begränsat antal internationella tester inkluderades i granskningen (Kapitel 6, Tabell 6.1). Valet gjordes efter samråd med internationellt välkända forskare inom dyslexiområdet. Därefter genomfördes en litteratursökning på de svenska och internationella testerna. Studier som undersökte testerna granskades för att ta reda på testernas vetenskapliga stöd. Förutom det gjordes en granskning av de svenska testernas manualer (några saknade manualer) för att samla information om deras tillförlitlighet (Bilaga 4 – Mall: Utvärdering av svenska tester). För testmetoder som syftar till att förutsäga dyslexi gjordes en förutsättningslös sökning.

Val av studiedesign

I rapporten ingår diagnostiska studier (kohortstudier eller tvärsnittstudier) för att undersöka testernas tillförlitlighet. Det var enbart longitudinella studier som var av intresse för att förutsäga dyslexi, dvs sådana studier som undersökte barn över lång tid (år). För interventionsstudier accepterades kontrollerade studier (randomiserade eller icke-randomiserade) där båda grupperna skulle bestå av barn och ungdomar med dyslexi.

Urval av studier

Urvalet har följt en strukturerad process enligt SBU:s metodbok [2] (Figur 3.2).

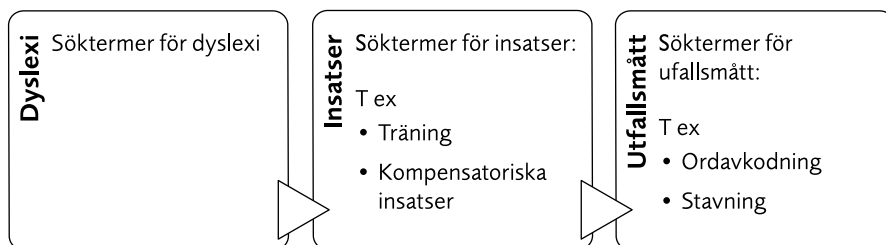


Figur 3.2 Illustration av den process som tillämpats vid urval och bedömning av studier.

Litteratursökning

För att identifiera så många relevanta test- och insatsstudier som möjligt utfördes sökningar i fyra internationella databaser; PsycInfo, ERIC (Education Resources Information Center), PubMed och LLBA (Linguistics and Language Behavior Abstracts). Databaserna täcker ämnesområdena beteendevetenskap, psykologi, psykiatri, utbildning, pedagogik, medicin, logopedi och lingvistik. Sökningar gjordes också i databaserna Cochrane Library, DARE, Prospero och i POP Database samt för hälsoekonomiska studier i NHS EED. Ytterligare studier söktes via de vetenskapliga artiklarnas referenslistor. Gallring av dubletter gjordes successivt så att endast unika referenser lades till från sökningarna i respektive databas. Som komplement till sökningarna granskades referenslistor för att identifiera ytterligare studier av relevans.

Sökstrategin utformades med utgångspunkt från att frågeställningarna och sökningarna skulle vara breda och förutsättningslösa. Den skiftande terminologin inom dyslexiområdet beaktades särskilt och i diskussionskapitlet diskuteras den mer utförligt.



Figur 3.3 Sökstrategin för insatser. Tre olika block av söktermer kombineras till en sökstrategi. Det första blocket bestod av söktermer för dyslexi. Det andra innehöll termer för de olika insatserna. Det tredje blocket omfattade söktermer för olika utfallsmått.

SBU:s informationsspecialist genomförde litteratursökningarna i samråd med experterna i projektgruppen och med projektledaren. Experterna bidrog även med centrala artiklar inom ämnesområdet. Därefter granskades artiklarnas titlar, sammanfattningar och särskilda kontrollerade ämnesord och i samråd med experterna beslutades om en sökstrategi. Sökningarna utfördes under september år 2012.

Litteratursökningarna om *insatser* vid dyslexi begränsades till åren 1980–2013 och för *tester* till 1990–2013. Två kompletterande sökningar gjordes. Den ena sökningen på *att förutsäga* dyslexi utfördes i databaserna PsycInfo och PubMed och begränsades till åren 1990–2013. Den andra sökningen på *kompensatoriska insatser* utfördes i databaserna PsycInfo och ERIC och begränsades till åren 2000–2013. I slutet av september år 2013 uppdaterades sökningarna i databaserna PsycInfo, ERIC och PubMed. En kontroll gjordes också för att bekräfta att inte någon av de studier som använts för slutsatserna hade återtagits, t ex på grund av oegentligheter eller forskningsfusk. Sökstrategin redovisas i Bilaga 1.

Granskningsprocessen

Fas 1: Gallring av artiklar mot inklusions- och exklusionskriterier

Två projektledare vid SBU:s kansli gallrade bort uppenbart ovidkommande artiklar från litteratursökningarnas abstraktlistor. Det fortsatta arbetet gjordes i par av experterna som fick ansvar för olika delar av projektet. I ett första skede gallrades studier bort som bedömdes som irrelevanta. Övriga beställdes i fulltext. Därefter bedömde experterna i varje par, oberoende av varandra, studiernas relevans och fattade sedan ett gemensamt beslut. Om experterna i ett par inte kunde komma överens, diskuterades studien i hela projektgruppen. Experterna bedömde inte artiklar där de själva varit medförfattare eller på annat sätt jäviga (kommersiella intressen för vissa tester). Studierna bedömdes med stöd av relevansmallar (Bilaga 3.5–3.7). De studier som bedömdes vara relevanta, dvs uppfyllde projektets inklusionskriterier, gick sedan vidare till kvalitetsgranskningen.

Fas 2: Kvalitetsgranskning av relevanta artiklar

En kvalitetsgranskning syftar till att bedöma risken för om studiens resultat har kunnat påverkas av systematiska fel, så kallad bias. Studiens kvalitet bedöms med stöd av checklistor. SBU strävar efter att checklistor och mallar ska vara likartade. De behöver dock inte vara identiska i de olika projekten.

Insatsstudierna granskades med hjälp av SBU:s standardmall för kontrollerade studier (randomiserade och icke-randomiserade). Mallen hade anpassats för detta projekt (Bilaga 3.10–3.11). Studier om prognostiska tester och kartläggningstester bedömdes med checklistan QUADAS, som hade modifierats något (Bilaga 3.9–3.11) [23]. Systematiska översikter bedömdes med stöd av checklistan AMSTAR (Bilaga 3.12) [24].

Beroende på hur välgjord en studie är kan den få hög studiekvalitet (låg risk för bias), medelhög studiekvalitet (medelhög risk för bias) eller låg studiekvalitet (hög risk för bias). Data från de studier som bedömdes ha hög eller medelhög kvalitet användes för att väga samman resultaten och bedöma evidensstyrkan. Det är ett mått på hur starkt det sammanlagda vetenskapliga underlaget är för att besvara en fråga på ett tillförlitligt sätt.

Fas 3: Tabellering av relevanta data från studier

Väsentliga uppgifter från studier som bedömdes vara av hög eller medelhög kvalitet sammanställdes i tabeller, se Kapitel 17. Syftet med tabellerna är att läsarna på ett enkelt sätt ska kunna få en överblick över de inkluderade studierna och hur de har bedömts. Tabellerna är skrivna på engelska för att underlätta för forskare och praktiker i andra länder att tillgodogöra sig en del av SBU:s grundläggande arbete.

Metoder för sammanvägning av resultat

Ambitionen var att föra samman jämförbara uppgifter för att kunna få en översiktlig och en mer allmängiltig tolkning av resultaten.

För tester och insatser där det fanns två eller flera studier av medelhög eller hög kvalitet för utfallsmåttet, gjordes en presentation av data i en

tabell inriktad på testet/insatsen i förhållande till utfallsmåttet. När det fanns två eller flera studier, där testet/insatsen och utfallsmåttet var tillräckligt lika, gjordes metaanalyser med hjälp av programmet Review Manager (version 5.2). Metaanalyserna gjordes med en så kallad fixed effect model. Det underliggande antagandet är att samtliga resultat är slumpmässiga urval från samma population som kan beskrivas som följande: barn och ungdomar som alla utsatts för tester/insatser och som undersökts i analysen för vissa utvalda utfallsmått, se vidare Bilaga 5, Bakgrund och övervägande angående statistik.

Metaanalyserna krävde ingående diskussioner och avvägningar i projektgruppen. För att illustrationen av studieresultaten i metaanalyserna skulle bli jämförbara transformerades de från rådata (standard score) till standardiserade medelvärdeskillnader (Cohen's d) i Review Manager. Det bör betonas att de sammanvägda resultaten bygger på originaldata i tabellerna och att metaanalyserna enbart är gjorda för att underlätta överblicken, eftersom data ofta har angetts i olika form i studierna.

Följande principer beaktades då resultaten illustrerades i metaanalyser:

- Varje person i en studie kan bara ingå en gång i en metaanalys
- En metaanalys kan innehålla samma utfallsmått, men det kan vara mätt med olika typer av tester. Dock ansåg vi testerna vara tillräckligt lika för att det skulle vara meningsfullt att väga samman data.

Baserat på tabeller med stöd av illustration i metaanalyser gjordes till sist en sammanvägning av resultaten.

Det vetenskapliga underlagets styrka

SBU använder det internationellt utarbetade systemet Grading of Recommendations Assessment Development and Evaluation (GRADE) för att bedöma hur tillförlitliga de sammanvägda resultaten är (Faktaruta 3.1) [25]. Tillförlitligheten uttrycks med hjälp av evidensstyrka. Evidensgraderingen görs i två steg, se SBU:s metodbok [2]. Ju högre slutlig evidens-

styrka, desto större sannolikhet att det sammanvägda resultatet är stabilt över tid och inte påverkas av nya forskningsrön.

Stegen vid evidensgradering

I ett första steg ges varje sammanvägt resultat en preliminär evidensstyrka som enbart bestäms av vilken design de ingående studierna har. Om resultatet bygger på randomiserade studier blir den preliminära evidensstyrkan hög (⊕⊕⊕⊕). För studier om diagnostisk tillförlitlighet kan även välgjorda tvärsnittsstudier utgå från en hög preliminär evidensstyrka. I nästa steg bedöms i vilken utsträckning som underlaget har svagheter beroende på metodologiska brister, heterogenitet, bristande relevans för svenska förhållanden, bristande precision i det sammanvägda resultatet och risk för publikationsbias (Faktaruta 3.1).

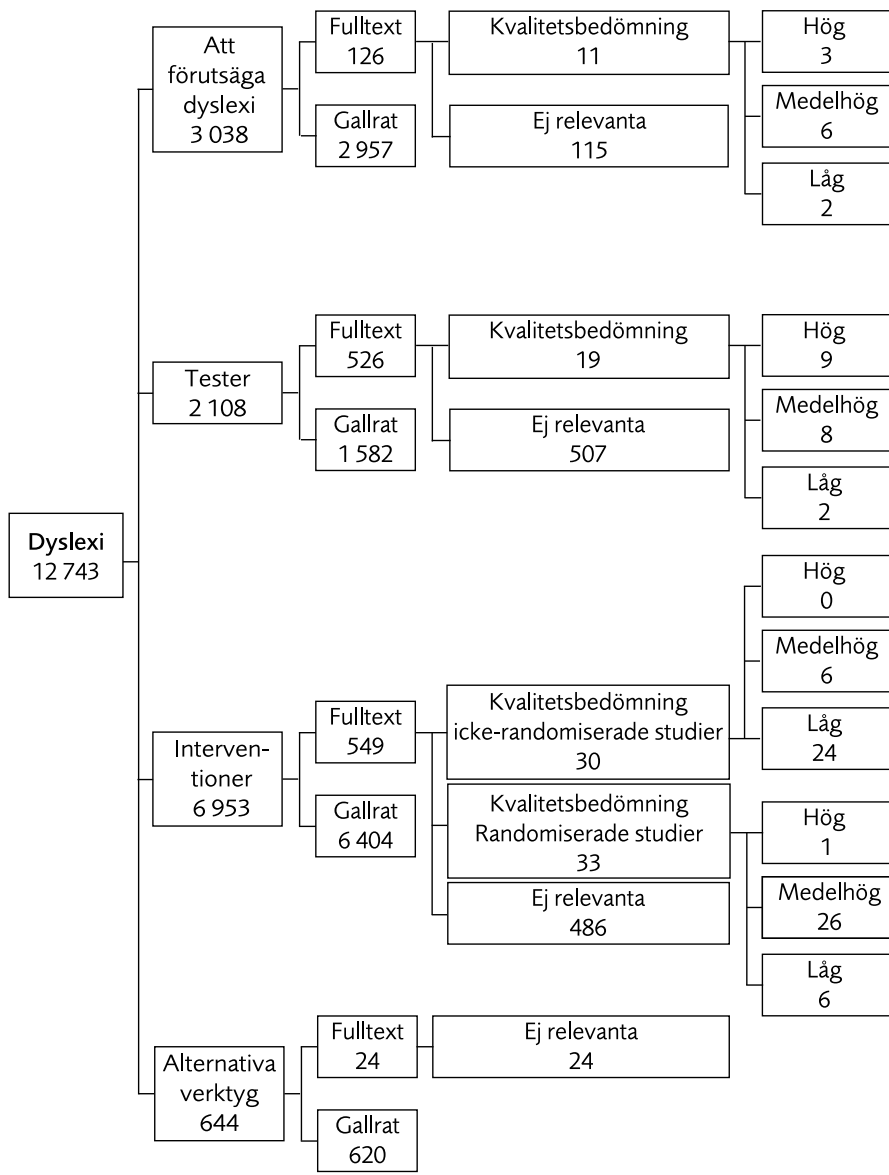
Faktaruta 3.1 Preliminär evidensstyrka baserad på studiedesign samt omständigheter som motiverar sänkning eller höjning av evidensstyrkan enligt GRADE.

Preliminär evidensstyrka	Symbol	Studiedesign
Stark	⊕⊕⊕⊕	Randomiserade studier
Måttligt stark	⊕⊕⊕○	
Begränsad	⊕⊕○○	Observationsstudier
Otillräcklig	⊕○○○	Fallstudier m m
Graderingen sänks om nedanstående försvagande omständigheter förekommer		
Brister i studiekvalitet	max -2	
Brister i överensstämmelse mellan studierna	max -2	
Brister i överförbarhet eller relevans	max -2	
Brister i precision	max -2	
Brist i form av hög sannolikhet för publikationsbias	max -1	

4. Resultatsammanställning

I detta kapitel redovisas en resultatsammanställning från litteratursökningen, gallringen och relevans- och kvalitetsbedömningen. I Tabell 4.1–4.3 sammanfattas de evidensgraderade resultaten. I Kapitel 5–8 beskrivs resultaten för; testmetoder för att förutsäga dyslexi, tester för att upptäcka och utreda dyslexi samt för insatser vid dyslexi. Varje kapitel inleds med en aggregerad beskrivning av de studier som utgör det vetenskapliga underlaget. Därefter följer en beskrivning av det sammanvägda resultatet från studierna följt av en bedömning av evidensstyrkan.

Vi gick igenom 12 743 artikelsammanfattningar. Av de 1 225 artiklar som sedan beställdes i fulltext var 93 artiklar relevanta för projektets tre frågeställningar och 59 hade tillräckligt hög kvalitet för att ingå i det vetenskapliga underlaget för evidensgradering, se flödesschema Figur 4.1.



Figur 4.1 Flödesschema över ingående studier.

Bedömning av evidensstyrkan

- Det finns begränsat vetenskapligt underlag för att brister i fonologisk medvetenhet kan förutsäga dyslexi (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att brister i snabb automatiserad benämningsförmåga (RAN) kan förutsäga dyslexi (⊕⊕○○).
- Det finns måttligt starkt vetenskapligt underlag för att brister i bokstavskännedom kan förutsäga dyslexi (⊕⊕⊕○).

Tabell 4.1 Evidensgraderade resultat; metoder för att förutsäga dyslexi hos barn och ungdomar.

Metoder för att förutsäga dyslexi	Brister i fonologisk medvetenhet	Brister i snabb automatiserad benämningsförmåga (RAN)	Brister i bokstavskännedom
Vetenskapligt underlag	Begränsat (⊕⊕○○)	Begränsat (⊕⊕○○)	Måttligt starkt (⊕⊕⊕○)

- Det finns otillräckligt vetenskapligt underlag för att avgöra vilka svenska tester som är tillförlitliga för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi bland barn och ungdomar, då det saknas relevanta studier (⊕○○○).
- Det finns begränsat vetenskapligt underlag för tillräcklig test-retest-reliabilitet för Dynamic Indicators of Basic Early Literacy Skills (DIBELS) deltest Letter-naming fluency (LNF) (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för tillräcklig test-retest-reliabilitet för Woodcock-Johnsons (WJ) deltester Word Identification subtest/Letter-Word Identification (⊕⊕○○).

- Det finns begränsat vetenskapligt underlag för validiteten för DIBELS deltest Nonsense Word Fluency och Test of Word Reading Efficiency (TOWRE) deltest Phonemic Decoding Efficacy (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för validiteten för WJ-Revised och The Peabody Individual Achievement Test-Revised (PIAT-R) med avseende på ordavkodning respektive läsförståelse (⊕⊕○○).

Tabell 4.2 Evidensgraderade resultat; tester för att upptäcka och utreda dyslexi hos barn och ungdomar.

Tester	Tillgängliga svenska tester¹	Etablerade internationella tester²
Vetenskapligt underlag	Otillräckligt (⊕○○○)	Otillräckligt för testerna som helhet (⊕○○○) Begränsat för enstaka deltester (⊕⊕○○)

¹ 52 tester som förekommer i Sverige utvärderades i översikten

² 18 internationella tester utvärderades i översikten

Tabell 4.3 Evidensgraderade resultat; insatser för barn och ungdomar med dyslexi.

Insatser → Utfallsmått ↓	Strukturerad träning av sambandet mellan språkljud (fonem) och bokstäver (grafem)	Träning av enbart fonologisk medvetenhet	Träning av enbart läsflyt (fluency)	Musik	Alternativa verktyg t ex mobilappar mm
Korrekt läsning av ord	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Korrekt läsning av non-ord (ord utan betydelse)	Stor förbättring (⊕⊕⊕○)	Bristande tillförlitlighet (⊕○○○)	Bristande tillförlitlighet (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Stavning	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läsförståelse	Måttlig förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Bristande tillförlitlighet (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läshastighet	Liten förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Fonologisk medvetenhet	Liten förbättring (⊕⊕○○)	Bristande tillförlitlighet (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)
Läsförståelse efter uppföljning (6–12 månader) efter avslutad insats	Liten förbättring (⊕⊕○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)

Tabellen fortsätter på nästa sida

Tabell 3 fortsättning

Insatser →	Strukturerad träning av sambandet mellan språkljud (fonem) och bokstäver (grafem)	Träning av enbart fonologisk medvetenhet	Träning av enbart läsflyt (fluency)	Musik	Alternativa verktyg t ex mobilappar mm
Utfallsmått ↓					
Livskvalitetsmått t ex självkompetens, självständighet, självbild, självförtroende, kunskapsutveckling	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas forskning (⊕○○○)	Saknas relevanta studier (⊕○○○)	Saknas forskning (⊕○○○)

(⊕⊕⊕○): Måttligt starkt vetenskapligt underlag

(⊕⊕○○): Begränsat vetenskapligt underlag

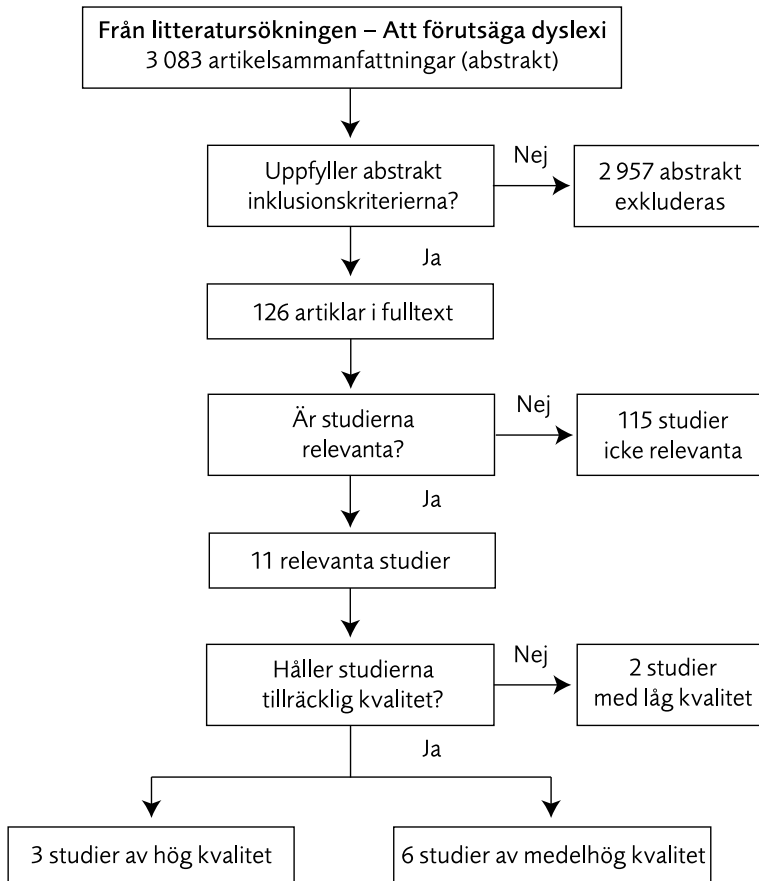
(⊕○○○): Otillräckligt vetenskapligt underlag

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra korrekt läsning av ord (⊕⊕○○).
- Det finns måttligt starkt vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en stor effekt på att förbättra korrekt läsning av non-ord (⊕⊕⊕○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra stavning (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt på att förbättra läsförståelse (⊕⊕○○).

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en liten effekt på att förbättra fonologisk medvetenhet (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en liten effekt på att förbättra läshastighet (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning efter sex till tolv månader efter avslutad insats har en liten effekt på att förbättra läsförståelse (⊕⊕○○).
- Det finns otillräckligt vetenskapligt underlag för att träning av enbart fonologisk medvetenhet jämfört med ordinarie undervisning har effekt på att förbättra korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att träning av enbart läsflyt jämfört med ordinarie undervisning har effekt på att förbättra läsning av ord på tid samt läsförståelse (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att avgöra om musik jämfört med kontrollgrupp (väntelista eller ingen insats) förbättrar utfallsmåtten (läsning, läsförståelse, läsflyt, fonologisk medvetenhet och stavning), då det saknas relevanta studier (⊕○○○).
- Det finns otillräckligt vetenskapligt underlag för att avgöra om det finns effekt av alternativa verktyg då det saknas studier (⊕○○○).

5. Att förutsäga dyslexi – resultat av granskningen

Litteratursökningen identifierade 3 083 artikelsammanfattningar av studier för testmetoder för att förutsäga dyslexi. Av dessa beställdes 126 artiklar i fulltext och 11 artiklar uppfyllde inklusionskriterierna. Tre studier hade hög kvalitet, sex hade medelhög kvalitet och två hade låg kvalitet. Studierna av hög och medelhög kvalitet, dvs de som ligger till grund för resultat och slutsatser, finns sammanfattade i Tabell 17.1 i Kapitel 17.



Figur 5.1 Flödesschema över gallring och bedömning av artiklar för testmetoder för att förutsäga dyslexi.

Beskrivning av ingående studier

Sex studier av medelhög kvalitet och tre studier av hög kvalitet undersökte sambandet mellan underliggande förmågor som kan användas som prediktorer för dyslexi hos barn före den formella läsinläringen [19,26–33]. Tre av studierna var från USA, två från Norge, två från Finland, en från Danmark och en från Israel. De vanligaste prediktorerna som studerades var fonologisk medvetenhet, snabb automatiserad benämningsförmåga (RAN) samt bokstavskänedom. Fem av de nio studierna studerade fonologisk medvetenhet [19,26,28,31,33], fyra studier undersökte RAN [19,28,31,33], medan tre studier studerade bokstavskänedom [28,31,33].

Fem av studierna undersökte sammansatta mått, så kallade kompositmått, bestående av flera möjliga underliggande förmågor [27,29,30,32,33]. Studien av Puolakanaho och medarbetare presenterade dock både sammansatta mått samt mått på enskilda prediktorers förmåga att predicera dyslexi [33].

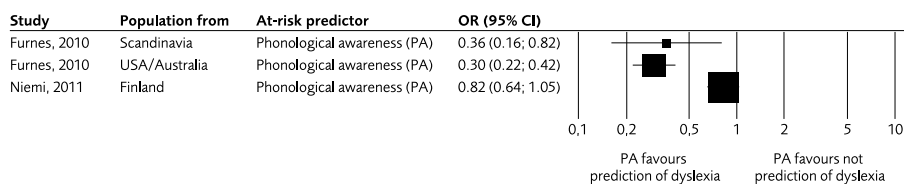
Fonologisk medvetenhet

Beskrivning av resultatet

Tabell 5.1 sammanfattar de studier som undersökte sambandet mellan fonologisk medvetenhet och manifest dyslexi. Resultatet var samstämt mellan de två studier som redovisade sambandet med hjälp av oddskvoter (95 % KI) [28,31]. Den enda studie som redovisade resultaten med måtten på sensitivitet och specificitet fann att sensitiviteten var låg medan specificiteten var hög [19]. Det innebär att fonologisk medvetenhet identifierar personer utan dyslexi med hög träffsäkerhet (dvs hög specificitet) men de identifierar inte med tillräcklig känslighet de som senare visar tecken på dyslexi (dvs låg sensitivitet). En studie redovisade enbart korrekt klassificerade personer utan att ange mått på sensitivitet och specificitet [26]. Puolakanaho och medarbetare redovisade sambandet med hjälp av oddskvot för fonologisk medvetenhet för barn som var fyra och ett halvt år gamla och som följdes upp vid nio års ålder [33].

Tabell 5.1 Sammanställning av resultaten i de studier som undersökte fonologisk medvetenhet (phonological awareness, PA) som prediktor för att förutsäga dyslexi hos barn.

Författare År Referens	Pre- diktor	Utfall	Population	Samband oddskvot (95 % KI)	Sensitivitet/ specificitet	Korrekt klassi- ficerad (%)
Catts et al 1991 [26]	PA	Dyslexi	USA			89
Furnes et al 2010 [28]	PA	Dyslexi	Skandinavien, USA, Australien	0,30 (0,22; 0,42) 0,36 (0,16; 0,82)		
Niemi et al 2011 [31]	PA	Dyslexi	Finland	0,822 (0,643; 1,050)		
Pennington et al 2012 [19]	PA	Dyslexi	USA, Australien, Norge, Sverige		0,415/ 0,935	
Puolakanaho et al 2007 [33]	PA	Dyslexi	Finland	0,667 (p-värde: 0,045)		



Figur 5.2 Fonologisk medvetenhet som prediktor för att förutsäga dyslexi. Jämförbara data från respektive studie ingår i grafen. Det fullständiga materialet återges i Tabell 5.1, grafen ska enbart ses som en illustration. Principer för det som illustreras i grafen beskrivs i metodavsnittet. Notera att den slutliga bedömningen av sambandet mellan fonologisk medvetenhet och dyslexi hos barn har baserats på samtliga data i Tabell 5.1.

Bedömning av evidensstyrkan

I Tabell 5.2 sammanfattas det vetenskapliga underlaget för fonologisk medvetenhet som prediktor för att förutsäga dyslexi hos barn.

- Det finns ett begränsat vetenskapligt underlag för att brister i fonologisk medvetenhet kan förutsäga dyslexi (⊕⊕○○).

Tabell 5.2 Fonologisk medvetenhet som prediktor för att förutsäga dyslexi hos barn.

Prediktor	Antal deltagare (studier)	Mått	Vetenskapligt underlag	Kommentar
Fonologisk medvetenhet	4 339 (5 studier) ^a	Samband (oddskvot) Sensitivitet Specificitet Korrekt klassificering	Begränsat ⊕⊕○○	Avdrag –2 pga bristande samstämmighet mellan studierna samt att olika statistiska mått använts

^a Catts, Furnes, Niemi, Pennington, Puolakanaho [19,26,28,31,33]

Snabb automatiserad benämningsförmåga (RAN)

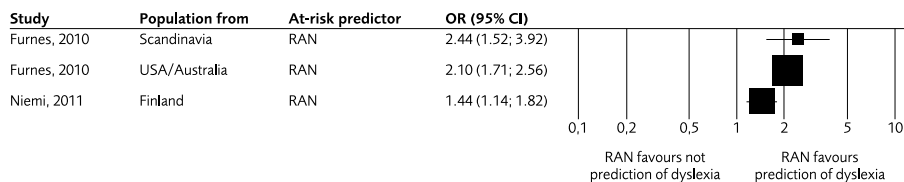
Beskrivning av resultatet

Tabell 5.3 sammanfattar de studier som undersökte sambandet mellan snabb automatiserad benämningsförmåga (RAN) och senare dyslexi. Resultatet var samstämmigt mellan studier där oddskvoten redovisats [28,31]. Pennington och medarbetare redovisade låg sensitivitet för RAN som prediktor för senare dyslexi, medan specificiteten var hög [19]. Det innebär att RAN förmår att utesluta personer utan dyslexi, men brister i förmågan att korrekt identifiera de som har dyslexi. Puolakanaho och medarbetare redovisade sambandet med hjälp av oddskvot för RAN för barn som var fem och ett halvt år gamla och som följdes upp vid nio års ålder [33].

Tabell 5.3 Sammanställning av resultaten i de studier som undersökte snabb automatiserad benämningsförmåga (RAN) som prediktor för att förutsäga dyslexi hos barn.

Författare År Referens	Prediktor	Utfall	Population	Samband oddskvot (95 % KI)	Sensitivitet/ Specificitet
Furnes et al 2010 [28]	RAN	Dyslexi	Skandinavien, USA, Australien	2,440 (1,519; 3,918) 2,090 (1,708; 2,557)	
Niemi et al 2011 [31]	RAN	Dyslexi	Finland	1,439 (1,37; 1,821)	
Pennington et al 2012 [19]	RAN	Dyslexi	USA, Australien, Norge, Sverige		0,427/0,935
Puolakanaho et al 2007 [33]	RAN	Dyslexi	Finland	1,52 ^Δ (p-värde: 0,009)	

^Δ I studien framgår att oddskvoten 0,657 är ett inverterat värde för prediktorn RAN. För att få jämförbara sambandsmått valde vi att invertera 0,657 vilket ger värdet 1,52



Figur 5.3 RAN som prediktor för att förutsäga dyslexi. Jämförbara data från respektive studie ingår i grafen. Det fullständiga materialet återges även i Tabell 5.3, grafen ska ses som en illustration. Principer för det illustreras i grafen beskrivs i metodavsnittet. Notera att den slutliga bedömningen av sambandet mellan RAN och dyslexi har baserats på samtliga data i Tabell 5.3.

Bedömning av evidensstyrkan

I Tabell 5.4 sammanfattas det vetenskapliga underlaget för snabb automatiserad benämningsförmåga (RAN) för att förutsäga dyslexi hos barn.

- Det finns ett begränsat vetenskapligt underlag för att brister i RAN kan förutsäga dyslexi (⊕⊕○○).

Tabell 5.4 Snabb automatiserad benämningsförmåga (RAN) för att förutsäga dyslexi hos barn.

Prediktor	Antal deltagare (studier)	Mått	Vetenskapligt underlag	Kommentar
Snabb automatiserad benämningsförmåga	4 298 (4 studier) ^a	Samband (oddskvot) Sensitivitet Specifitet	Begränsat ⊕⊕○○	Avdrag -2 pga bristande samstämmighet mellan studierna samt att olika statistiska mått användes

^a Furnes, Niemi, Pennington, Puolakanaho [19,28,31,33]

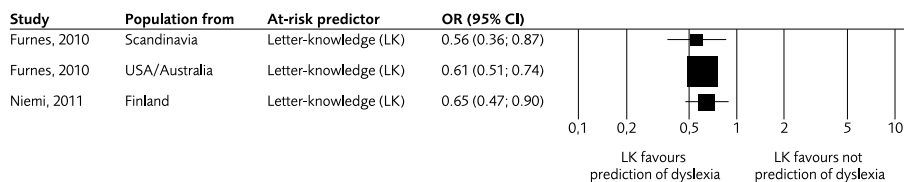
Bokstavskänedom

Beskrivning av resultatet

Tabell 5.5 sammanfattar de två studier som undersökt sambandet mellan bokstavskänedom och senare dyslexi [28,31]. Resultatet var samstämmigt mellan de två studierna som studerade sambandet med hjälp av oddskvoter (95 % KI). Puolakanaho och medarbetare redovisade sambandet med hjälp av oddskvot för bokstavskänedom för barn som var fem och ett halvt år gamla och som följdes upp vid nio års ålder [33].

Tabell 5.5 Sammanställning av resultaten i de studier som undersökte bokstavskänedom (letter-knowledge, LK) som prediktor för att förutsäga dyslexi hos barn.

Författare År Referens	Prediktor	Utfall	Population	Samband oddskvot (95 % KI)
Furnes et al 2010 [28]	LK	Dyslexi	Skandinavien, USA, Australien	0,56 (0,36; 0,87) 0,61 (0,51; 0,74)
Niemi et al 2011 [31]	LK	Dyslexi	Finland	0,649 (0,468; 900)
Puolakanaho et al 2007 [33]	LK	Dyslexi	Finland	0,464 (p-värde: 0,000)



Figur 5.4 Bokstavskänedom som prediktor för att förutsäga dyslexi. Jämförbara data från respektive studie ingår i grafen. Det fullständiga materialet återges även i Tabell 5.5, grafen ska ses som en illustration. Principer för det som illustreras i grafen beskrivs i metodavsnittet. Notera att den slutliga bedömningen av sambandet mellan bokstavskänedom och dyslexi har baserats på samtliga data i Tabell 5.5.

Bedömning av evidensstyrkan

I Tabell 5.6 sammanfattas det vetenskapliga underlaget för bokstavskänedom för att förutsäga dyslexi hos barn.

- Det finns ett måttligt starkt vetenskapligt underlag för att brister i bokstavskänedom kan förutsäga dyslexi (⊕⊕⊕○).

Tabell 5.6 Bokstavskänedom för att förutsäga dyslexi hos barn.

Prediktor	Antal deltagare (studier)	Mått	Vetenskapligt underlag	Kommentar
Bokstavskänedom	3 489 (3 studier) ^a	Samband	Måttligt starkt ⊕⊕⊕○	Avdrag -1 pga att bokstavskänedom är känsligt för överförbarhet till svenska språket

^a Furnes, Niemi, Puolakanaho [28,31,33]

Sammanlagda prediktorer

Beskrivning av resultatet

I detta avsnitt redovisas fem studier som undersökt förmågan av sammanslagda prediktorer (fonologisk medvetenhet, snabb automatiserad benämningsförmåga och bokstavskänedom) för att förutsäga dyslexi och för andra prediktorer (tester) för att förutsäga dyslexi [27,29,30,32,33].

Elbro och medarbetare redovisade att en sammanslagning av bokstavskänedom, fonologisk medvetenhet, artikulation och fonologiska representationer, mätt vid sex års ålder, hade en sensitivitet på 0,78 och en specificitet på 0,79 på att förutsäga dyslexi efter ett års formell läsundervisning [27].

Puolakanaho och medarbetare följde i en longitudinell studie 198 barn från tre och ett halvt års ålder, med uppföljningar vid fyra och ett halvt och fem och ett halvt års ålder, upp till nio års ålder då eleverna gick i andra klass [33]. Drygt hälften av barnen hade föräldrar med diagnostiserad dyslexi. När dyslexi predicerades ingick grupptillhörighet (föräldrar med eller utan dyslexi) tillsammans med övriga mått som en prediktor. Vid tre och ett halvt års ålder predicerades dyslexi med en sammanslagning av grupptillhörighet, bokstavskänedom och RAN med en sensitivitet på 0,78 och en specificitet på 0,72. Vid fyra och ett halvt års ålder,

med en sammanslagning av grupptillhörighet, bokstavskänedom och fonologisk medvetenhet, uppskattades sensitiviteten och specificiteten till respektive 0,72 och 0,74. Vid fem och ett halvt års ålder var måtten desamma som vid tre och ett halvt års ålder, och dyslexi predicerades med en sensitivitet på 0,76 och en specificitet på 0,76.

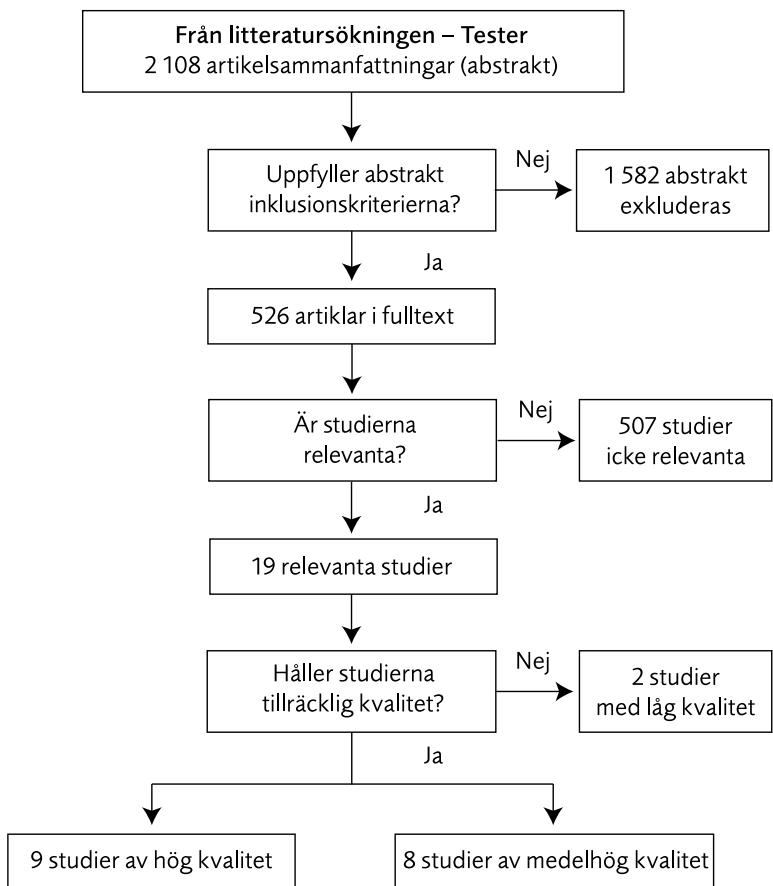
O'Connor och medarbetare följde tre kohorter med sammanlagt 445 barn [32]. Barnen testades när de var cirka fem och ett halvt år på sju tester på fonologisk medvetenhet och ett test som mätte RAN. Ungefär ett och ett halvt år senare kunde dessa tillsammans predicera dyslexi med en sensitivitet på 0,91–1,0 och en specificitet på 0,86–0,88. Skillnaderna i uppskattad sensitivitet och specificitet var försumbara mellan de tre kohorterna.

Det går inte att med exakt säkerhet tolka sensitiviteten och specificiteten i ovanstående tre studier om hur de enskilda prediktorerna förutsäger dyslexi [27,32,33]. Däremot visar resultaten samstämmigt att kombinationer av tester som mäter aspekter av fonologisk medvetenhet, bokstavskänedom och RAN kan förutsäga dyslexi med både hög träffsäkerhet och tillräcklig känslighet. Detta stödjer resultaten att brister i fonologisk medvetenhet, RAN och bokstavskänedom kan predicera dyslexi.

Två studier utvärderade andra prediktorer (tester) för att förutsäga dyslexi [29,30]. I den ena studien besvarade förskollärare och föräldrar till fem år gamla barn ett frågeformulär (the risk index questionnaire) [29]. Formuläret innehöll frågor avseende hälsa, motorik, språk, behov av pedagogiskt stöd och ärftlighet. När barnen var 11 år gamla kunde svaren förutsäga om barnen hade utvecklat dyslexi med en sensitivitet på 0,85 och en specificitet på 0,62. I den andra studien undersöktes dikotisk lyssning med hjälp av två verbala tester (en experimentell metod för att bli studera språkhantering i de båda hjärnhalvorna) [30]. Testerna genomfördes på 328 barn vid sex års ålder. I slutet av årskurs ett i skolan fick 288 av barnen genomföra ett test som avsåg att mäta korrekt läsning av ord. De som presterade minst en standardavvikelse under medelvärdet ansågs ha dyslexi. Resultatet uppvisade en specificitet på 0,91 och en sensitivitet på 0,77, dvs de dikotiska testerna var bra på att identifiera de barn som inte löpte risk att utveckla dyslexi medan känsligheten när det gällde att förutsäga dyslexi var något sämre.

6. Tester för att upptäcka och utreda dyslexi – resultat av granskningen

Litteratursökningen identifierade 2 108 artikelsammanfattningar varav 526 beställdes i fulltext. Av dessa uppfyllde 19 artiklar inklusionskriterierna. Nio studier hade hög kvalitet, åtta medelhög och två låg kvalitet. Det innebär att det vetenskapliga underlaget består av 17 studier och dessa finns sammanfattade i Tabell 17.2 i Kapitel 17.



Figur 6.1 Flödesschema över gallring och bedömning av samtliga artiklar för tester.

Beskrivning av ingående studier

Av de 17 ingående studierna genomfördes alla i USA och 6 265 barn och ungdomar ingick. Fem studier syftade till att utreda den tidiga läsutvecklingen, de flesta från förskoleåret till andra året i skolan [34–38]. Av de resterade 12 studierna hade nästan alla som primärt syfte att jämföra olika tester (hela testpaket eller enskilda deltester) med varandra eller med sig själva i test-retest-mätningar för att därigenom få en uppfattning om testernas validitet eller reliabilitet [39–50].

Samtliga studier innehöll information som kunde användas för att bedöma psykometrisk standard hos instrumenten som utvärderades. Reliabiliteten bedömdes utifrån inre konsistens (Cronbach's alfa) och/eller test-retest-korrelationer. Validiteten kunde bedömas på olika sätt. De flesta studierna undersökte hur ett test, som antas reflektera en viss förmåga, korrelerade med andra mått av denna förmåga. Tester av de olika utfallsmåtten bör följaktligen korrelera starkt eller åtminstone medelstarkt inbördes om validiteten ska kunna bedömas som tillfredsställande. Ett annat sätt att testa validiteten, som möjligen är än mer relevant, är att undersöka om instrumentet kan användas för att skilja mellan två grupper, i detta fall om testresultatet skiljer sig åt för personer med dyslexi i jämförelse med personer utan dyslexi. Endast en studie undersökte validiteten hos testinstrumenten på detta sätt [44]. Därmed fanns det endast ett arbete tillgängligt som direkt relaterade till frågan om testet kunde användas för att identifiera svårigheter som är karaktäristiska vid dyslexi.

Beskrivning av resultaten

Nedan ges en sammanfattande beskrivning av studier vars syfte var att utvärdera hur väl testerna fungerade för att upptäcka och utreda läs- och skrivfärdigheter, med fokus på dyslektiska problem, hos barn och ungdomar. För två tester, Dynamic Indicators of Basic Early Literacy Skills (DIBELS) och testbatteriet Woodcock-Johnson (WJ), fanns flera studier. Endast 10 av de 18 utvärderade testerna förekommer bland de inkluderade studierna. En sammanställning av de 18 utvärderade internationella testerna finns i Tabell 6.1 och resultaten av de ingående studierna finns sammanställda i Tabell 6.2–6.3.

Tabell 6.1 Sammanställning över de utvärderade internationella testerna.

BRIGANCE Comprehensive Inventory of Basic Skills–Revised (CIBS-R) [51]

Clinical Evaluation of Language Fundamentals (CELF) [52]

Comprehensive Test of Phonological Processing (CTOPP) [53]

Dynamic Indicators of Basic Early Literacy Skills (DIBELS) [54]

Gates McGinitie Reading Test (GMRT) [55]

Gray Oral Reading Test (GORT) [56]

Reading Assessment and Diagnostic Evaluation (GRADE) [57]

Holbrook Screening Battery (HSB) [58]

Kindergarten Reading Engagement Scale (KRES) [59]

Peabody Individual Achievement Test-Revised (PIAT-R) [60]

Stanford Diagnostic Reading Test [61]

Test of Phonological Awareness (TOPA) [62]

Test of Word Reading Efficiency (TOWRE) [63]

Texas Primary Reading Inventory (TPRI) [64]

Wechsler Individual Achievement Test (WIAT) [65]

Woodcock-Johnson Reading Mastery Test, Revised /Woodcock-Johnson III [66,67]

Wide Range Achievement Test of Spelling (WRAT) [68]

York Assessment of Reading for Comprehension (YARC) [69]

Tabell 6.2 Sammanställning av uppgifter i de studier som har undersökt deltester inom testet DIBELS för att upptäcka och utreda läsfärdighet och dyslexi.

Författare Referens	Indextest – DIBELS	Referenstest	Reliabilitet	Validitet
Burke et al [34]	Nonsense Word Fluency Oral Reading Fluency	TOWRE (Phonemic Decoding Efficacy, Sight Word Efficacy)		x
Elliot et al [40]	Letter Naming Fluency Sound Naming Fluency Initial Phoneme Phonemic Segmentation Fluency	WJ-R TOPA Teacher rating	x	x
Fien et al [36]	Nonsense Word Fluency	DIBELS (Oral Reading Fluency)	x	
Hagan-Burke et al [37]	Letter Naming Fluency Nonsense Word Fluency	TOWRE (Phonemic Decoding Efficacy, Sight Word Efficacy)		x
Hintze et al 44]	Letter Naming Fluency Initial Sound Fluency Phonemic Segmentation Fluency	CTOPP (Phonological Awareness, Phonological memory, Rapid Memory, Elision, Sound Matching)		x
McBride et al [38]	Letter Naming Fluency Initial Sound Fluency Word Usage Phoneme Segmentation Nonsense Word Oral Reading Retell	None	x	

CTOPP = Comprehensive Test of Phonological Processing [53]; DIBELS = Dynamic Indicators of Basic Early Literacy Skills [54]; TOPA = Test of Phonological Awareness [62]; TOWRE = The Test of Word Reading Efficiency [63]; WJ-R = Woodcock-Johnson Psycho Educational Battery-Revised [66]

Sex studier utvärderade screeningtestet DIBELS [34,36–38,40,44]. Tre studier på barn i förskolan undersökte reliabiliteten uttryckt som test-retest-reliabilitet [36,38,40]. En studie kom fram till att korrelationen mellan testtillfällen för deltestet Letter Naming Fluency var 0,90 [40] medan en annan studie fann en test-retest-reliabilitet på 0,86 [38]. Fien och medarbetare undersökte deltestet Nonsense Word Fluency [36]. Det kunde dock inte utvärderas då det bara undersöktes i en studie.

Fyra studier undersökte validiteten för DIBELS när testet användes på barn i årskurs ett [34,37,40,44]. I två studier ställdes DIBELS deltest Nonsense Word Fluency mot TOWRE:s deltest Phonemic Decoding Efficacy och där fann man en korrelation på 0,75 mellan deltesterna [34,37].

Hintze och medarbetare undersökte samstämmigheten mellan DIBELS och CTOPP (ett mer omfattande instrument för utredning av fonologiska funktioner) [44]. Korrelationen mellan deltestet ”fonemiskt segmenteringsflyt” i DIBELS och den fonologiska kompositen i CTOPP var 0,53. Hintze och medarbetare analyserade även DIBELS förmåga att fånga upp de barn som fallit ut som svaga på CTOPP (mer än en standardavvikelse under medelvärdet, dvs barn som ligger i riskzonen för dyslexi). Analyserna visade genomgående att DIBELS hade en acceptabel sensitivitet på 80–100 och en låg specificitet på 21–39. Elliot och medarbetare undersökte samstämmigheten mellan DIBELS, WJ-R, TOPA och lärarskattning [40]. Sammantaget visar resultatet på signifikanta korrelationer mellan liknande deltester mätt med DIBELS och dessa andra instrument, vilket stödjer validiteten hos DIBELS.

Tabell 6.3 Sammanställning av uppgifter i de studier som har undersökt testet Woodcock-Johnson (WJ) för att upptäcka och utreda läsfärdighet och dyslexi.

Författare Referens	Indextest	Referenstest	Reliabilitet	Validitet
Daub et al [39]	WJ-R	PIAT-R		x
Erford et al [41]	WJ-III	WRAT-3	x	x
Mather et al [48]	WJ-R	WRAT-R		x
Prewett et al [49]	WJ-R	PIAT-R		x
Shull-Senn et al [50]	WJ-R	None	x	

PIAT-R = Peabody Individual Achievement Test-Revised [60]; WJ-R = Woodcock-Johnson Psycho Educational Battery-Revised [66]; WJ-III = Woodcock-Johnson Test of Achievement Third Edition [66,67]; WRAT-R = Wide Range Achievement Test- Revised [68]

Fem studier utvärderade testbatteriet WJ [39,41,48–50]. Två av studierna mätte reliabiliteten [41,50]. Erford och medarbetare [41] och Shull-Senn och medarbetare [50] fokuserade på WWI/LWI som är deltester som mäter förmågan att avkoda ord och bokstäver. Erford och medarbetare fann att WWI hade en inre konsistens på 0,93 (n=165) och en test-retest korrelation på 0,96. Shull-Senn och medarbetare noterade test-retest-korrelationer mellan 0,86 och 0,96 för elever i årskurs ett till fem för LWI.

Fyra studier redovisade validitetsmått för WJ [39,41,48,49]. Samtliga studier använde en korrelationsanalys med ett annat test. Daub och medarbetare [39] respektive Prewett och medarbetare [49] ställde deltesterna LWI och PIAT-R, som mäter ordavkodning i WJ-R, mot varandra och fick fram en korrelation på 0,71 respektive 0,87. För deltesterna som mäter läsförståelse observerade Daub och medarbetare en korrelation på 0,73 och Prewett och medarbetare en korrelation på 0,82. Mather och medarbetare fann en korrelation mellan stavningsdeltesterna i WJ-R

och WRAT-R på 0,71 hos en grupp elever med inlärningssvårigheter [48]. Erford och medarbetare fann en korrelation på 0,90 mellan WJ-III deltest WWI (identifiering av ord) och WRAT-3 deltest för läsning [41].

Sex indextester utvärderades i en studie vardera: Kindergarten Reading Engagement Scale (KRES); the Brigance Comprehensive Inventory of Basic Skills-II (CIBS-II), Test of Phonological Awareness (TOPA), Comprehensive Test of Phonological Processing (CTOPP), Gray Oral Reading Test (GORT) och Peabody Individual Achievement Test (PIAT) [35,42,43,45–47]. Studierna finns förtecknade i Tabell 6.4. Underlaget är för litet för att kunna dra några slutsatser om testernas prestanda.

Tabell 6.4 Indextester som utvärderades i en studie vardera.

Författare Referens	Indextest	Referenstest	Reliabilitet	Validitet
Clarke et al [35]	KRES	None	x	
Gotch et al [42]	CIBS-II	DIBELS	x	x
Havey et al [43]	TOPA	CTOPP		x
Keenan et al [45]	GORT	PIAT/WJPC		x
Keenan et al [46]	PIAT	GORT/WJPC		x
Kilpatrick et al [47]	CTOPP	None	x	

CIBS-II = The Brigance Comprehensive Inventory of Basic Skills-II [51]; CTOPP = Comprehensive Test of Phonological Processing [53]; DIBELS = Dynamic Indicators of Basic Early Literacy Skills [54]; GORT = Gray Oral Reading Test [56]; KRES = Kindergarten Reading Engagement Scale [59]; PIAT = Peabody Individual Achievement Test [60]; TOPA = Test of Phonological Awareness [62]; WJPC = Woodcock-Johnson Passage Comprehension subtest [70]

Bedömning av evidensstyrkan

I Tabell 6.5 sammanfattas det vetenskapliga underlaget för DIBELS deltest LNF och WJ:s deltester WWI/LWI.

- Det finns ett begränsat vetenskapligt underlag för tillräcklig test-retest-reliabilitet för DIBELS deltest LNF (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för tillräcklig test-retest-reliabilitet för WJ:s deltester WWI/LWI (⊕⊕○○).

Tabell 6.5 Testernas reliabilitet för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi hos barn och ungdomar.

Test	Antal deltagare (studier)	Utfall	Vetenskapligt underlag	Kommentar
DIBELS: LNF	275 (2 studier) ^a	Reliabilitet	Begränsat ⊕⊕○○	Avdrag -2 pga att konfidensintervallen inte uppges och inte kan beräknas
WJ-III: WWI WJ-R: LWI	225 (2 studier) ^b	Reliabilitet	Begränsat ⊕⊕○○	Avdrag -2 pga att konfidensintervallen inte uppges och inte kan beräknas

DIBELS = Dynamic Indicators of Basic Early Literacy Skills [54]; LNF = Letter Naming Fluency; LWI: Letter-Word Identification; WJ-R = Woodcock-Johnson Psycho Educational Battery-Revised [66]; WJ-III = Woodcock-Johnson Test of Achievement Third Edition [66,67]; WWI = Word Identification subtest

^a Elliot, McBride [38,40]

^b Erford, Shull-Senn [41,50]

I Tabell 6.6 sammanfattas det vetenskapliga underlaget för DIBELS deltest Nonsense Word Fluency och Test of Word Reading Efficiency (TOWRE) deltest Phonemic Decoding Efficacy samt för WJ-R och PIAT-R.

- Det finns begränsat vetenskapligt underlag för validiteten för DIBELS deltest Nonsense Word Fluency och Test of Word Reading Efficiency (TOWRE) deltest Phonemic Decoding Efficacy (⊕⊕○○).
- Det finns begränsat vetenskapligt underlag för validiteten för WJ-Revised och The Peabody Individual Achievement Test-Revised (PIAT-R) med avseende på ordavkodning respektive läsförståelse (⊕⊕○○).

Tabell 6.6 Testernas validitet för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi hos barn och ungdomar.

Test (Referenstest)	Antal deltagare (studier)	Utfall	Vetenskapligt underlag	Kommentar
DIBELS: NWF (TOWRE: PDE)	414 (2 studier) ^a	Validitet	Begränsat ⊕⊕○○	Avdrag -2 pga att konfidensintervallen inte uppges och inte kan beräknas
WJ-R: LWI (PIAT-R)	153 (2 studier) ^b	Validitet	Begränsat ⊕⊕○○	Avdrag -2 pga att konfidensintervallen inte uppges och inte kan beräknas
WJ-R: Läsförståelse (PIAT-R)	153 (2 studier) ^b	Validitet	Begränsat ⊕⊕○○	Avdrag -2 pga att konfidensintervallen inte uppges och inte kan beräknas

DIBELS = Dynamic Indicators of Basic Early Literacy Skills [54]; LWI = Letter-Word Identification; NWF = Nonsense Word Fluency; PIAT-R = Peabody Individual Achievement Test-Revised; TOWRE = Test of Word Reading Efficiency; WJ-R = Woodcock-Johnson Psycho Educational Battery-Revised [66]

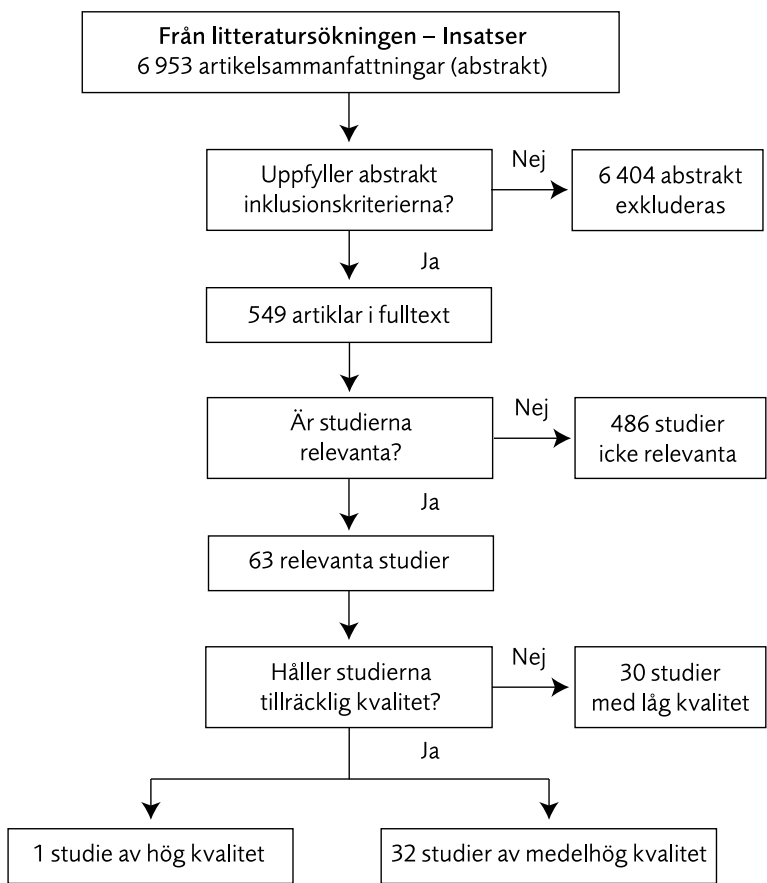
^a Hagan-Burke, Burke [34,37]

^b Daub, Prewett [39,49]

- Det finns ett otillräckligt vetenskapligt underlag för att avgöra vilka svenska tester som är tillförlitliga för att upptäcka och utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi bland barn och ungdomar, då det saknas studier (⊕○○○). De svenska testerna som finns tillgängliga för att upptäcka och utreda dyslexi och som ingick i granskningen, finns listade i Tabell 11.1 i Kapitel 11.

7. Insatser för barn och ungdomar med dyslexi – resultat av granskningen

Litteratursökningen identifierade 6 953 artikelsammanfattningar varav 549 lästes i fulltext. Av dessa uppfyllde 63 artiklar inklusionskriterierna, 33 var randomiserade kontrollerade studier och 30 var icke-randomiserade studier. Av de randomiserade studierna hade 6 låg studiekvalitet, 26 hade medelhög studiekvalitet och 1 hade hög studiekvalitet. Av de icke-randomiserade studierna hade 24 låg studiekvalitet, 6 hade medelhög studiekvalitet och inga studier hade hög studiekvalitet. Det innebär att det vetenskapliga underlaget består av 27 randomiserade kontrollerade studier och 6 icke-randomiserade studier. Dessa finns sammanfattade i Tabell 17.3 i Kapitel 17.



Figur 7.1 Flödesschema över gallring och bedömning av artiklar för insatser.

Randomiserade kontrollerade studier

Av de 27 inkluderade studierna utfördes 10 i USA, 3 i Kanada, 3 i England, 3 i Sverige, samt 2 i Nederländerna. En studie vardera genomfördes i Finland, Grekland, Spanien, Schweiz, Brasilien och Indien.

Insatserna grupperades efter vilken typ av träning som avsågs; lästräning, enbart fonologisk träning, enbart läsflytsträning eller strukturerad träning av fonem–grafem-koppling. Insatserna kunde sedan jämföras med antingen ordinarie undervisning eller annan insats med förmodad effekt. Här redovisar vi först studier där en insats jämförs med en annan insats med förmodad effekt, såsom jämförelse av två olika program. Därefter redovisas studier med insatser som jämförs med den undervisning skolan erbjuder.

Insatser som jämförs med andra insatser

Beskrivning av ingående studier

Sexton studier jämförde två insatser med förmodad effekt med varandra [71–86]. Sju av dessa studier undersökte olika metoder för att träna läsning [71–76,87]. Insatserna skilde sig åt i så hög grad att det var svårt att väga resultaten från olika studier mot varandra. En studie undersökte effekten av att läsa samma text vid upprepade tillfällen jämfört med en metod där barnen fick läsa olika texter vid upprepade tillfällen [76]. En annan studie jämförde interventioner som använde sig av explicita eller implicita avkodningsstrategier [73]. En tredje studie använde datorbaserad träning för att undersöka effekten av att fokusera på felästa eller rättlästa ord [74]. Alla tre studierna undersökte läsförståelse men hade helt olika infallsvinklar. Hit hör också Torgesen och medarbetare som ställde en intervention med ingående fonemanalys mot en induktiv analys (embedded phonics) med utgångspunkt i hela ord, men med ingående och strukturerad fonem–grafem-koppling [75].

Wise och medarbetare studerade effekten av att träna fonologisk medvetenhet och läsförståelse genom två olika datorbaserade program (ROSS) [85]. Eleverna tränade både i mindre grupper och individuellt. Eleverna

fick lika mycket träning men det skilde åt hur mycket de tränade att läsa ord i kontext. De som tränade fonologisk medvetenhet tränade extra på fonologisk medvetenhet och individuella ord, förutom att läsa ord i kontext. Eleverna som tränade läsförståelse använde all tid på att läsa ord i en kontext.

Sex studier jämförde olika metoder för att strukturerat träna fonem–grafem-koppling [77–82]. del Rosario Ortiz och medarbetare prövade en intervention för att träna bokstavsljud och fonologisk medvetenhet med eller utan övningar i taluppfattning [77]. Macaruso och medarbetare studerade effekten av att kombinera ett datorbaserat ordavkodningsprogram och ett språktränande program jämfört med enbart det språktränande programmet [81].

Tre studier kunde inte placeras i någon överordnad kategori eftersom interventionerna var unika, vilket innebär att de inte kunde jämföras med en annan insats med förmodad effekt. Hit hör t ex Calhoon och medarbetare [86] som utvärderade parläsning (dvs eleverna läser två och två för varandra) samt Mavrommati och medarbetare [83] som undersökte effekten av stavning med hjälp av bilder. En studie undersökte -träning av enbart fonologisk medvetenhet utan koppling till skriftspråket [84].

Beskrivning av resultaten

Av de 16 studierna var det 7 som undersökte olika metoder för att träna läsning [71–76,85]. I en undersökning jämfördes en fonembaserad, kontextfri och muntligt inriktad metod med en kontextbaserad skriftlig metod som använde onset-rime (onset är en konsonant i början på ordet, och rime den återstående delen som börjar med en vokal, t ex för ordet tack är "t" onset och "ack" rime) [72]. Båda insatserna var lika framgångsrika. Eleverna kunde också tillämpa kunskaperna på nytt material efter träningen. Shippen och medarbetare jämförde två läsprogram där det ena baserades på explicita och det andra på implicita avkodningsstrategier [73]. Insatserna gavs till elever med avkodningsproblem från mellanstadiet och uppåt. Båda interventionerna gav positiva effekter på ordläsning, läshastighet, läsflyt och korrekt läsning av ord. Effekten blev

större för dem som hade bättre läsförmåga när studien startade. Wexler och medarbetare jämförde en metod där barnen läste samma text vid upprepade tillfällen med en metod där barnen fick läsa olika texter [76]. Insatserna gav inga skillnader mellan grupperna. I en annan studie jämfördes insatsen PB som var fonem–grafem-baserad, med insatsen PASS som var baserad på en metod för inläring av olika kognitiva strategier, konstruerad av en av författarna [71]. Det fonem–grafem-baserade programmet ledde till att ordavkodningen förbättrades, medan PASS, enligt författarna, förbättrade både ordavkodning och ordläsning (för de elever som låg under medianen jämfört med dem som låg ovanför). Torgesen och medarbetare undersökte effekterna av två läsinterventionsprogram [75]. Det ena tränade fonemkunskap via artikulatoriska och akustiska ledtrådar, medan det andra tränade fonemkunskapen via skrivning och stavning. Båda interventionerna ledde till förbättringar i läsförmåga på både kort och lång sikt (uppföljning efter 2 år). Ett år efter interventionen behövde 40 procent av eleverna inte längre någon specialundervisning. Trots att de förbättrades när det gäller korrekt läsning och läsförståelse var läshastigheten fortfarande låg. Steenbeck-Planting och medarbetare utförde två interventioner med avsikt att förbättra korrekt läsning av ord [74]. Jämförelsen gjordes med utgångspunkt från elevernas initiala läsnivå. Metoderna fokuserade på att träna korrekt lästa ord eller felaktigt lästa ord. Författarna fann att alla barn som tränade ordläsning via dator förbättrade sin läsförmåga och sitt läsflyt på de tränade orden samt förbättrade läsförmågan även på otränade ord. Vilket fokus på träningen som var mest framgångsrikt – träning av korrekt lästa ord eller felaktigt lästa ord – berodde på elevens läsnivå och typen av ord. Information till eleven avseende vilket fokus träningen hade, påverkade resultaten positivt oavsett läsnivå.

Sex studier jämförde olika metoder för att strukturerat träna fonem–grafem-koppling [77–82]. del Rosario Ortiz och medarbetare undersökte om träning i taluppfattning har positiva effekter på fonemmedvetenhet genom att kontrastera två interventioner som båda tränade bokstavs–ljudsamband och fonemmedvetenhet och den första dessutom taluppfattning [77]. Båda interventionsgruppernas fonemmedvetenhet ökade, men bara den första gruppen uppnådde bättre resultat i läsning jämfört med kontrollgruppen.

I studien av Fälth och medarbetare jämfördes tre träningsgrupper [78]. En grupp tränade fonologisk medvetenhet och förmågan att stärka fonem–grafem-kopplingen. Den andra gruppen fokuserade på ord och meningsnivå. Den tredje gruppen använde båda träningsmetoderna. Resultaten visade att alla tre träningsgrupperna blev bättre på ordavkodning och läsförståelse men den grupp som fick det kombinerade programmet uppvisade bäst resultat.

I Kerstholt och medarbetares studie undersöktes effekten av visuellt stöd i tre olika typer av träning i fonemsegmentering [79]. Den första interventionen använde diagram och bokstäver, den andra tränade bara bokstäver, den tredje kontrollgruppen fick inget visuellt stöd utan enbart auditiv träning. Alla tre interventionsprogrammen förbättrade förmågan till fonemsegmentering, men utan inbördes skillnader. I den andra uppföljningsstudien av Kerstholt med nya elever eliminerades de skillnader i återkoppling som hade funnits i de olika interventionerna, men resultatet från första studien kvarstod [79]. Författarnas slutsats var att dessa barn, till skillnad från förskolebarn, inte drog nytta av visuellt stöd i tillägg till den auditiva träningen och att man därför inte borde träna förskolebarn och barn med svag segmenteringsförmåga med samma metoder.

Macaruso och medarbetare undersökte effekten av att, utöver det språkstödande programmet Language, använda ett datorbaserat program Lexia SOS [81]. Programmet tränar systematiskt fonologisk medvetenhet och ordavkodningsförmåga i fem olika nivåer för äldre elever. Undersökningen gjordes på mellanstadieelever som jämfördes med en grupp elever som bara undervisades i Language. Resultatet visade att undersökningsgruppen förbättrade sin läsning jämfört med kontrollgruppen. För avkodningen var skillnaderna signifikanta.

Morris och medarbetare undersökte dels om program med inslag av explicit träning i ordidentifikationsstrategier, metakognitiv inlärning och ordkunskap förbättrade läsförmågan jämfört med enbart träning av sambandet mellan fonem och grafem, dels om något av dessa kompletterande inslag gav större positiv effekt på läsförmågan än andra [80]. Undersökningen inriktades på lässvaga elever i årskurserna två och tre. Resultatet visade att traditionell träning av fonem–grafem-koppling i

kombination med antingen ordidentifikationsstrategier, metakognitiv inläring eller ordkunskap klart förbättrade läsförmågan jämfört med enbart träning av fonem–grafem-koppling.

I studien av Wright och medarbetare utvärderades två insatser [82]. Båda insatserna tränade fonologisk medvetenhet och bokstavskunskap. I den ena insatsen tränades även metalingvistiska begrepp och metakognitiva strategier. Båda elevgrupperna förbättrades, men effekten var större i gruppen som också fick träning av metalingvistiska begrepp och metakognitiva strategier.

I studien av Wise och medarbetare jämfördes läsförståelseträning med fonologisk medvetenhetsträning [85]. Båda insatserna förbättrade elevernas förmåga på ordavkodning och läsförståelse. Emellertid presterade gruppen fonologisk medvetenhet bättre på alla tester som inte var tidsbaserade såsom fonologisk medvetenhet, ordigenkänning och non-ordläsning. Läsförståelsegruppen fick bättre resultat på ordavkodning som var tidsbegränsad och på att känna igen svåra ord samt på läsförståelse. I en uppföljningsstudie av eleverna, ett och två år efter interventionerna, visade resultaten att det inte fanns några skillnader mellan grupperna avseende ordigenkänning [88]. Enligt författarna är datorbaserad träning effektiv för elever med lässvårigheter. Träningen behöver dock ges under lång tid för att vara bestående.

Endast en studie utvärderade en metod för att träna fonemisk och fonologisk medvetenhet [84]. I den studien jämfördes en intervention inriktad på fonologisk och språkstödande träning med träning av muntliga språkfärdigheter och ordkunskap. Eleverna var flerspråkiga och träningen gällde ett språk som inte var deras modersmål. Den grupp som fick fonologiskt inriktad träning uppvisade bättre resultat i läsning och stavning jämfört med den andra interventionsgruppen.

Mavrommati och medarbetare undersökte effekten av stavning [83]. Den jämförde två träningsprogram i stavning för grekiska barn med dyslexi. Den ena träningen använde bilder (piktogram) för att illustrera alternativa stavningar. Den andra träningen använde en kombination av fonem–grafem-koppling och morfologisk analys. Enligt författarna var

piktogrammetoden med bilder mer effektiv än den språkligt inriktade träningen för att förbättra stavning.

Calhoon och medarbetare jämförde ett träningsprogram i fonologi och läsförståelse som innehöll kamratstödjande insatser med ett vanligt förekommande specialundervisningsprogram i helklass utan kamratstödjande insatser [89]. Det kamratstödjande träningsprogrammet gav bättre effekt på bokstavs- och ordidentifikation, ordattack och läsförståelse med stora skillnader i standardiserad medelvärdeskilnad. Inga skillnader i läsflyt kunde påvisas mellan grupperna. Författarna diskuterade, med utgångspunkt från studien, vikten av hur man implementerar ett interventionsprogram kamrater emellan eller i helklass.

Bedömning av sammanvägd effekt och evidensstyrka

Det går inte att väga samman resultaten från studierna, grupperade i olika typer av insatser, eftersom interventionerna är alltför olika. Det medför att det vetenskapliga underlaget för att bedöma effekterna av en intervention jämfört med en annan är otillräckligt.

Insatser som jämförs med ordinarie undervisning

Beskrivning av ingående studier

Sexton studier utvärderade effekter av en intervention jämfört med ordinarie undervisning. Tre av studierna omnämns i avsnittet ovan då de inte går att väga samman med övriga studier som jämför interventioner med ordinarie undervisning [81,83,89]. Samtliga studier hade en design med pre- och posttest. Fyra studier följde upp resultatet efter sex månader till ett år för utfallsmåttet läsförståelse [78,80,90,91]. Studierna använde många olika utfallsmått, såsom att läsa ord korrekt, läsa non-ord korrekt, stavning, läsförståelse, läshastighet och fonologisk medvetenhet.

Åtta studier fokuserade på att barnen systematiskt skulle kunna tillgodogöra sig den alfabetiska principen [77,78,80,90–94]. Denna typ av träning

brukar kallas phonics på engelska. Vi har inte motsvarande term på svenska, men det innebär bl a att man tränar sambandet mellan grafem (bokstäver) och fonem (språkljud) på ett strukturerat sätt. Ofta börjar dessa interventioner på en mycket grundläggande nivå med fonemisk medvetenhet och enkel fonem–grafem-koppling för att successivt träna mer avancerade principer om skriftspråket. Hädanefter kommer vi att benämna denna typ av intervention som strukturerad fonem–grafem-koppling.

Tre av studierna om strukturerad fonem–grafem-koppling undersökte om en intervention med flera komponenter förbättrade förmågan av stavning, läsförståelse, läshastighet, fonologisk medvetenhet och korrekt läsning av ord och non-ord jämfört med ordinarie undervisning [78,80,90]. Det gick dock inte att väga samman effekterna från de olika studierna eftersom innehållet i träningen var för olika.

Tre ytterligare studier undersökte effekten av enbart fonologisk träning, dvs att man uppmärksammar den ljudmässiga sidan av språket utan att relatera det till skriftspråket [72,84,95].

Två studier fokuserade enbart på träning av läsflyt [76,96]. I dessa studier bestod träningen av så kallad upprepad läsning, vilket innebär att man läser samma ord, fraser eller sammanhängande text vid upprepade tillfällen i syfte att uppnå automatisering.

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet korrekt läsning av ord

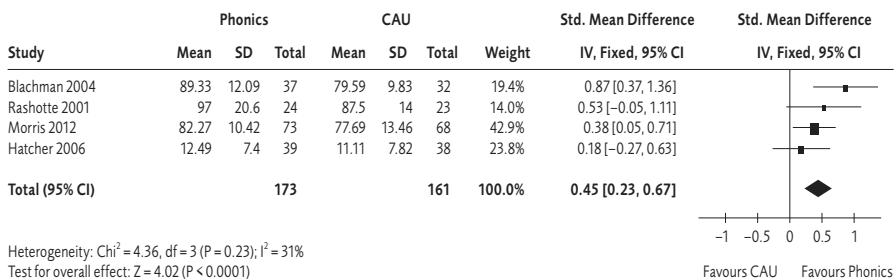
Beskrivning av resultatet

Tabell 7.1 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att läsa enstaka ord korrekt. Som framgår av Figur 7.2 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten, standardiserad medelvärdeskilnad (SMD), var 0,45 (95 % KI, 0,23 till 0,67), dvs en måttlig effekt. Notera att den sammanvägda effekten inte inkluderar studien av del Rosario Ortiz och medarbetare då den innehåller ett sammansatt mått för läsning [77].

Tabell 7.1 Sammanställning av studier som undersökte effekten av strukturerad fonem-grafem-koppling på utfallsmåttet korrekt läsning av ord (word reading accuracy – words).

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Blachman et al 2004 [91]	Korrekt läsning av ord	USA (n=37+32)	8 månader	89,33 (12,09)	79,59 (9,83)	1,31
del Rosario Ortiz et al 2002 [77]	Korrekt läsning av ord/non-ord	England (n=18+18)	10 veckor	54,05 (3,5)	51,37 (7,82)	*p<0,10 (0,44) ^a
Morris et al 2012 [80]	Korrekt läsning av ord	USA (n=73+68)	14 veckor	83,27 (10,42)	77,69 (13,46)	0,47 (0,13; 0,80) ^a
Rashotte et al 2001 [93]	Korrekt läsning av ord	Kanada (n=24+23)	8 veckor	97,00 (20,60)	87,5 (14,00)	0,73
Hatcher et al 2006 [94]	Korrekt läsning av ord	England (n=39+38)	10 veckor	12,49 (7,40)	11,11 (7,82)	*p<0,01 (0,18) ^a

^a Standardiserad medelvärdeskillnad är beräknad utifrån tabellsiffror då den inte var angiven i studien



Figur 7.2 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra läsning av enstaka ord hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,45 (95 % KI, 0,23 till 0,67), vilket indikerar att det finns en effekt av att träna fonem–grafem-koppling för att få förbättrad förmåga att läsa enstaka ord korrekt. Notera att den sammanvägda effekten inte inkluderar studien av del Rosario Ortiz och medarbetare, då denna studie innehåller ett sammansatt mått för läsning [77].

Bedömning av evidensstyrkan

I Tabell 7.2 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på korrekt läsning av ord.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, har en måttlig effekt, standardiserad medelvärdeskillnad (SMD) 0,45 (0,23; 0,67), på att förbättra korrekt läsning av ord (⊕⊕○○).

Tabell 7.2 Effekt av träning av strukturerad fonem–grafem-koppling på korrekt läsning av ord för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Vetenskapligt underlag	Kommentar
Fonem–grafem-koppling	Korrekt läsning av ord	n=370 (5 randomiserade kontrollerade studier) ^a	0,45 [0,23; 0,67]	Begränsat ⊕⊕○○	Avdrag –2 pga att insatserna inte var identiska samt smärre brister i samstämmighet

^a Blachman, Morris, Rashotte, Hatcher, del Rosario Ortiz [77,80,91,93,94]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet korrekt läsning av non-ord

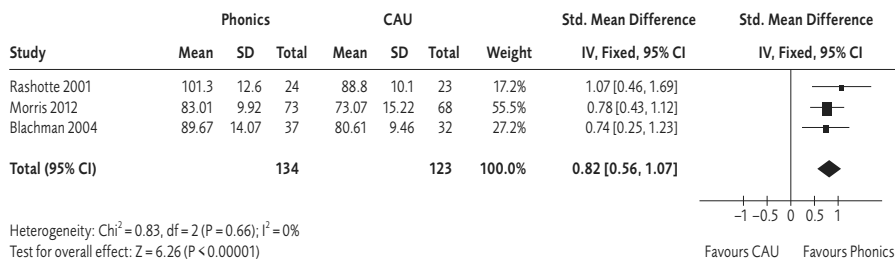
Beskrivning av resultatet

Tabell 7.3 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att läsa enstaka non-ord korrekt. Som framgår av Figur 7.3 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,82 (95 % KI, 0,56 till 1,07), dvs en stor effekt. Notera att den sammanvägda effekten inte inkluderar studien av del Rosario Ortiz och medarbetare, då denna studie innehåller ett sammansatt mått för läsning [77].

Tabell 7.3 Sammanställning av studier som undersökte effekten av strukturerad fonem–grafem–koppling på utfallsmättet korrekt läsning av non-ord (word reading accuracy – nonwords).

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Blachman et al 2004 [91]	Korrekt läsning av non-ord	USA (n=37+32)	8 månader	89,67 (14,07)	80,61 (9,46)	0,89
Morris et al 2012 [80]	Korrekt läsning av non-ord	USA (n=73+68)	14 veckor	83,01 (9,92)	73,07 (15,22)	(0,78)
Rashotte et al 2001 [93]	Korrekt läsning av non-ord	Kanada (n=24+23)	8 veckor	101,3 (12,6)	88,8 (10,10)	1,67
del Rosario Ortiz et al 2002 [77]	Korrekt läsning av ord/non-ord	England (n=18+18)	10 veckor	54,05 (3,5)	51,37 (7,82)	*p<0,10 (0,44) ^a

^a Standardiserad medelvärdeskilnad är beräknad utifrån tabellsiffror då den inte var angiven i studien



CAU = Curriculum as usual

Figur 7.3 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra läsning av non-ord hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,82 (95 % KI, 0,56 till 1,07), vilket indikerar att det finns en stor effekt av att träna strukturerad fonem–grafem-koppling och få förbättrad förmåga att läsa non-ord med korrekthet. Notera att denna sammanvägda effekt inte inkluderar studien av del Rosario Ortiz och medarbetare då den innehåller ett sammansatt mått för läsning [77].

Bedömning av evidensstyrkan

I Tabell 7.4 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på korrekt läsning av non-ord.

- Det finns måttligt starkt vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en stor effekt, SMD 0,82 (0,56; 1,07), på att förbättra korrekt läsning av non-ord (⊕⊕⊕○).

Tabell 7.4 Effekt av att träna strukturerad fonem–grafem-koppling på korrekt läsning av non-ord för barn och ungdomar med dyslexi.

Insats	Utfalls- mått	Antal deltagare (studier)	Sammanvägd effekt	Veten- skapligt underlag	Kommentar
Fonem– grafem- koppling	Korrekt läsning av non-ord	n=287 (4 randomi- serade kontrollerade studier) ^a	0,82 [0,56; 1,07],	Måttligt starkt ⊕⊕⊕○	Avdrag –1 pga att insatserna inte var identiska

^a Rashotte, Morris, Blachman, del Rosario Ortiz [77,80,91,93]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet stavning

Beskrivning av resultatet

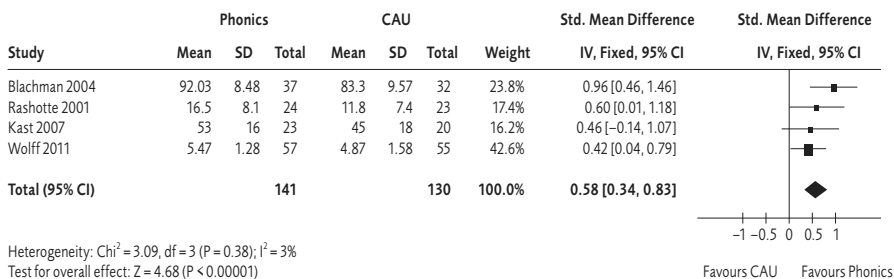
Tabell 7.5 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att stava. Som framgår av Figur 7.4 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,58 (95 % KI, 0,34 till 0,83), dvs en måttlig effekt.

Tabell 7.5 Sammanställning av studier som undersökte effekten av att träna strukturerad fonem-grafem-koppling för utfallsmättet stavning.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d)
Blachman et al 2004 [91]	Stavning	USA (n=37+32)	8 månader	92,03 (8,48)	83,3 (9,57)	1,13
Kast et al 2007 [92]	Stavning	Schweiz (n=23+20)	3 månader	53,0 (16,00) ^b	45,00 (18,00) ^b	(0,47) ^a
Rashotte et al 2001 [93]	Stavning	Kanada (n=24+23)	8 veckor	16,5 (8,1)	11,8 (7,4)	0,75
Wolff 2011 [90]	Stavning	Sverige (n=57+55)	12 veckor	5,47 (1,28)	4,87 (1,58)	0,30

^a Standardiserad medelvärdesskillnad är beräknad utifrån tabellsiffror då den inte var angiven i studien

^b Resultat speglat med antal felstavade ord



CAU = Curriculum as usual

Figur 7.4 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra förmågan att stava hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,58 (95 % KI, 0,34 till 0,83), vilket indikerar att det finns en effekt av att träna fonem–grafem-koppling för att förbättra förmågan att stava. Notera att den slutliga bedömningen av effekten baserats på samtliga data i Tabell 7.5.

Bedömning av evidensstyrkan

I Tabell 7.6 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på stavning.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, har en måttlig effekt, SMD 0,58 (0,34; 0,83), på att förbättra stavning (⊕⊕○○).

Tabell 7.6 Effekt av att träna strukturerad fonem–grafem-koppling på stavning för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Veten-skapligt underlag	Kommentar
Fonem–grafem-koppling	Stavning	n=271 (4 randomiserade kontrollerade studier) ^a	0,58 [0.34; 0.83]	Begränsat ⊕⊕○○	Avdrag –2 pga metodologiska brister samt att insatserna inte var identiska

^a Blachman, Rashotte, Kast, Wolff [90-93]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läsförståelse

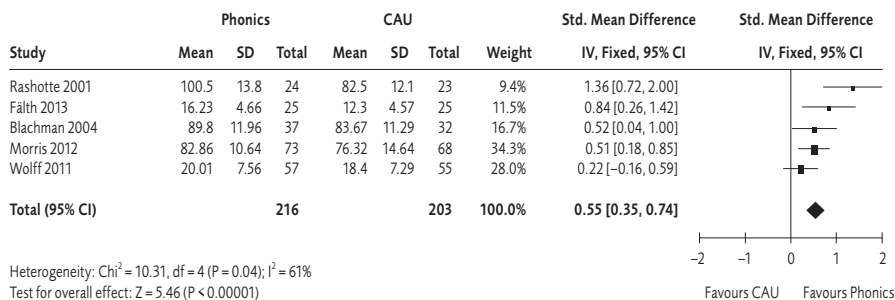
Beskrivning av resultatet

Tabell 7.7 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att förbättra läsförståelsen. Som framgår av Figur 7.5 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,55 (95 % KI, 0,35 till 0,74), dvs en måttlig effekt.

Tabell 7.7 Sammanställning av studier som undersökte träning av strukturerad fonem-grafem-koppling för utfallsmättet läsförståelse.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Blachman et al 2004 [91]	Läsförståelse	USA (n=37+32)	8 månader	89,80 (11,96)	83,67 (11,29)	0,55
Fälth et al 2013 [78]	Läsförståelse	Sverige (n=25+25)	5–8 veckor	16,23 (4,66)	12,3 (4,57)	(0,86) ^a
Rashotte et al 2001 [93]	Läsförståelse	Kanada (n=24+23)	8 veckor	100,5 (13,8)	82,5 (12,1)	1,51
Wolff 2011 [90]	Läsförståelse	Sverige (n=57+55)	12 veckor	20,01 (7,56)	18,4 (7,29)	0,41
Morris et al 2012 [80]	Läsförståelse	USA (n=73+68)	14 veckor	82,86 (10,64)	76,32 (14,64)	(0,52) ^a

^a Standardiserad medelvärdeskilnad är beräknad utifrån tabellsiffror då den inte var angiven i studien



Figur 7.5 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra läsförståelsen hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,55 (95 % KI, 0,35 till 0,74), vilket indikerar att det finns en effekt av att träna fonem–grafem-koppling för att förbättra läsförståelsen.

Bedömning av evidensstyrkan

I Tabell 7.8 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på läsförståelse.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning har en måttlig effekt, SMD 0,55 (0,35; 0,75), på att förbättra läsförståelse (⊕⊕○○).

Tabell 7.8 Effekt av träning av strukturerad fonem–grafem-koppling på läsförståelse för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Veten-skapligt underlag	Kommentar
Fonem–grafem-koppling	Läsförståelse	n=419 (5 randomiserade kontrollerade studier) ^a	0,55 [0,35; 0,74]	Begränsat ⊕⊕○○	Avdrag –2 pga att insatserna inte var identiska samt brister i samstämmighet

^a Rashotte, Fälth, Blachman, Morris, Wolff [78,80,90,91,93]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet fonologisk medvetenhet

Beskrivning av resultatet

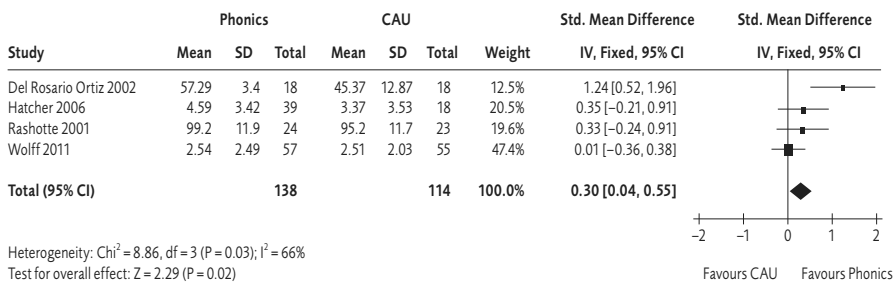
Tabell 7.9 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att förbättra den fonologiska medvetenheten. Som framgår av Figur 7.6 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,30 (95 % KI, 0,04 till 0,55), dvs en liten effekt.

Tabell 7.9 Sammanställning av studier som undersökte strukturerad fonem–grafem-koppling för utfallsmåttet fonologisk medvetenhet.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Hatcher et al 2006 [94]	Fonologisk medvetenhet	England (n=39+38)	10 veckor	4,59 (3,42)	3,37 (3,53)	(0,35) ^a
Rashotte et al 2001 [93]	Fonologisk medvetenhet	Kanada (n=24+23)	8 veckor	99,2 (11,9)	95,2 (11,7)	0,40
Wolff 2011 [90]	Fonologisk medvetenhet	Sverige (n=57+55)	12 veckor	2,54 (2,49)	2,51 (2,03)	0,43 ^a
del Rosario Ortiz et al 2002 [77]	Fonologisk medvetenhet	England (n=18+18)	10 veckor	57,29 (3,4)	45,37 (12,87)	*p<0,01 (1,27) ^a

^a Standardiserad medelvärdeskillnad är beräknad utifrån tabellsiffror då den inte var angiven i studien

^Δ Värdet 0,43 har framkommit genom att differenspoängen mellan post- och pre-test jämförts för insats och "kontrollgrupp". För att inte överskatta effekten, har studiens post-test-värde för insats och "kontrollgrupp" använts i Figur 7.6



CAU = Curriculum as usual

Figur 7.6 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra förmågan att utveckla fonologisk medvetenhet hos barn med dyslexi. Jämförbara data ingår i grafen. Den sammanvägda effekten är 0,30 (95 % KI, 0,04 till 0,55), vilket indikerar att det finns en effekt av att träna fonem–grafem-koppling för att förbättra den fonologiska medvetenheten.

Bedömning av evidensstyrkan

I Tabell 7.10 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på fonologisk medvetenhet.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, har en liten effekt, SMD 0,30 (0,04; 0,55), på att förbättra fonologisk medvetenhet (⊕⊕○○).

Tabell 7.10 Effekt av träning av strukturerad fonem–grafem-koppling på fonologisk medvetenhet för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Veten-skapligt underlag	Kommentar
Fonem–grafem-koppling	Fonologisk medvetenhet	n=252 (4 randomiserade kontrollerade studier) ^a	0,30 [0,04; 0,55]	Begränsat ⊕⊕○○	Avdrag –2 pga att insatserna inte var identiska samt brister i samstämmighet

^a del Rosario Ortiz, Rashotte, Hatcher, Wolff [77,90,93,94]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läshastighet

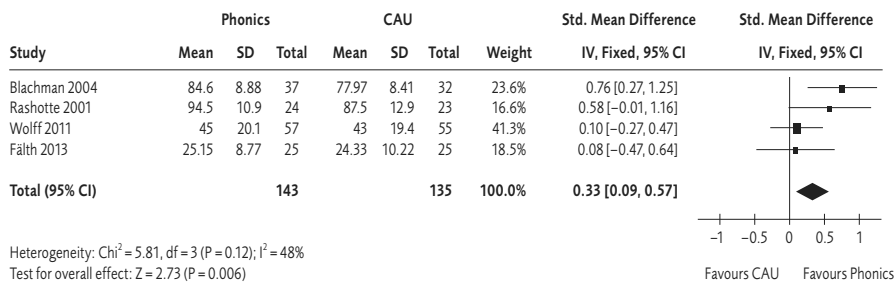
Beskrivning av resultatet

Tabell 7.11 sammanfattar resultaten av de studier som undersökte hur träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, påverkade förmågan att förbättra läshastigheten. Som framgår av Figur 7.7 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,33 (95 % KI, 0,09 till 0,57), dvs en liten effekt.

Tabell 7.11 Sammanställning av studier som undersökte effekter av träning av strukturerad fonem–grafem-koppling på utfallsmåttet läshastighet.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Blachman et al 2002 [91]	Läshastighet	USA (n=37+32)	8 månader	84,6 (8,8)	77,97 (8,41)	0,96
Rashotte et al 2001 [93]	Läshastighet	Kanada (n=24+23)	8 veckor	94,5 (10,9)	87,5 (12,9)	0,63
Wolff 2011 [90]	Läshastighet	Sverige (n=57+55)	12 veckor	45,00 (20,1)	43,00 (19,4)	0,15
Fälth et al 2013 [78]	Läshastighet	Sverige (n=25+25)	5–8 veckor	25,15 8,77	24,33 (10,22)	(0,09) ^a

^a Standardiserad medelvärdeskilnad är beräknad utifrån tabellsiffror då den inte var angiven i studien



CAU = Curriculum as usual

Figur 7.7 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra läshastigheten hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,33 (95 % KI, 0,09 till 0,57). Notera att den slutliga bedömningen av effekten av att träna fonem–grafem-koppling för att förbättra läshastigheten hos barn med dyslexi baseras på samtliga data i Tabell 7.11.

Bedömning av evidensstyrkan

I Tabell 7.12 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på läshastighet.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling, jämfört med ordinarie undervisning, har en liten effekt, SMD 0,33 (0,09; 0,57), på att förbättra läshastighet (⊕⊕○○).

Tabell 7.12 Effekt av träning av strukturerad fonem–grafem-koppling på läshastighet hos barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Vetenskapligt underlag	Kommentar
Fonem–grafem-koppling	Läshastighet	n=278 (4 randomiserade kontrollerade studier) ^a	0,33 [0,09;0,57]	Begränsat ⊕⊕○○	Avdrag –2 pga att insatserna inte var identiska samt brister i samstämmighet

^a Rashotte, Fälth, Blachman, Wolff [78,90,91,93]

Träning av strukturerad fonem–grafem-koppling för utfallsmåttet läsförståelse, efter sex månader till ett år

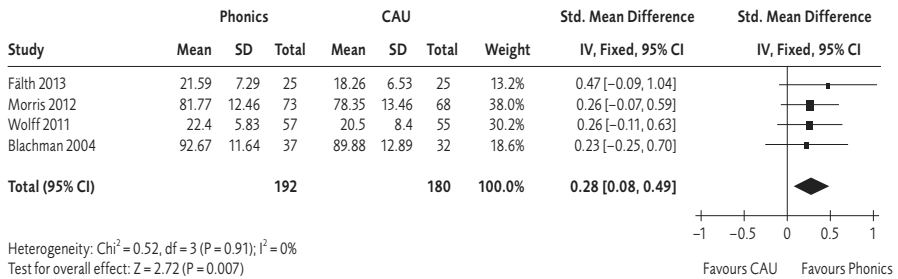
Beskrivning av resultatet

I Tabell 7.13 sammanfattas resultaten av de studier som undersökte hur träning av fonem–grafem-koppling påverkade förmågan att förbättra läsförståelsen efter sex månader till ett år. Som framgår av Figur 7.8 kunde resultaten läggas in som en metaanalys. Den sammanvägda effekten var SMD 0,28 (95 % KI, 0,08 till 0,49), dvs en liten effekt.

Tabell 7.13 Sammanställning av studier som undersökte effekten av att träna strukturerad fonem-grafem-koppling för utfallsmåttet läsförståelse efter sex månader till ett år.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) *p-värde
Blachman et al 2004 [91]	Läsförståelse	USA (n=37+32)	8 månader	92,67 (11,64)	89,88 (12,89)	0,24
Fälth et al 2013 [78]	Läsförståelse	Sverige (n=25+25)	5–8 veckor	21,59 (7,29)	18,26 (6,53)	(0,48) ^a
Wolff 2011 [90]	Läsförståelse	Sverige (n=57+55)	12 veckor	22,40 (5,83)	20,5 (8,4)	0,33
Morris et al 2012 [80]	Läsförståelse	USA (n=73+68)	14 veckor	81,77 (12,46)	78,35 (13,46)	(0,26) ^a

^a Standardiserad medelvärdesskillnad är beräknad utifrån tabellsiffror då den inte var angiven i studien



CAU = Curriculum as usual

Figur 7.8 Effekten av att träna strukturerad fonem–grafem-koppling för att förbättra läsförståelsen hos barn med dyslexi. Jämförbara data från respektive studie ingår i grafen. Den sammanvägda effekten är 0,28 (95 % KI, 0,08 till 0,49), vilket indikerar att det finns en effekt av att träna fonem–grafem-koppling för att förbättra läsförståelsen efter en uppföljning på sex månader till ett år.

Bedömning av evidensstyrkan

I Tabell 7.14 sammanfattas det vetenskapliga underlaget för effekten av träning av strukturerad fonem–grafem-koppling på läsförståelse efter sex till tolv månader.

- Det finns begränsat vetenskapligt underlag för att träning av strukturerad fonem–grafem-koppling jämfört med ordinarie undervisning efter sex till tolv månader efter avslutad insats har en liten effekt, SMD 0,28 (0,08; 0,49), på att förbättra läsförståelse (⊕⊕○○).

Tabell 7.14 Effekt av träning av strukturerad fonem–grafem-koppling på läsförståelse efter sex till tolv månader för barn med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Sammanvägd effekt	Vetenskapligt underlag	Kommentar
Fonem–grafem-koppling	Läsförståelse	n=372 (4 randomiserade kontrollerade studier) ^a	0,28 [0,08; 0,49]	Begränsat ⊕⊕○○	Avdrag –1 pga att insatserna inte var identiska och uppföljningstiden varierade mellan studierna

^a Fälth, Morris, Wolff, Blachman [78,80,90,91]

Träning av enbart fonologisk medvetenhet för utfallsmåtten korrekt läsning av ord, stavning, läsförståelse och fonologisk medvetenhet

Beskrivning av resultatet

Tabell 7.15 sammanfattar resultaten av de studier som undersökte hur träning av enbart fonologisk medvetenhet, jämfört med ordinarie undervisning, påverkade korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet.

Tabell 7.15 Sammanställning av studier som undersökte träning av enbart fonologisk medvetenhet.

Författare År, referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d) (95 % KI)
Gustafson et al 2000, [95]	Korrekt läsning av ord	Sverige (n=33+16)	5+4 månader	31,8 (10,60)	30,8 (8,90)	0,02 (-0,58; 0,62)
Nag-Arulmani et al 2003, [84]	Korrekt läsning av ord	Indien (n=31+28)	3 veckor	33,42 (5,25)	29,57 (3,16)	0,96 (0,42; 1,50)
O'Shaughnessy et al 2000, [72]	Korrekt läsning av ord	USA (n=15+15)	6 veckor	79,5 (8,6)	79,4 (7,2)	-0,44 (-1,17; 0,28) 0,01 (-0,70; 0,73)
Nag-Arulmani et al 2003, [84]	Stavning	Indien (n=31+28)	3 veckor	26,25 (2,46)	25,39 (2,03)	0,5 (-0,29; 1,29)
O'Shaughnessy et al 2000, [72]	Stavning	USA (n=15+15)	6 veckor	84,9 (8,40)	79,8 (6,70)	0,71 (-0,03; 1,44)
Nag-Arulmani et al 2003, [84]	Läsförståelse	Indien (n=31+28)	3 veckor	14,77 (3,19)	16,14 (3,81)	-0,39 (-0,91 ; 0,12)
O'Shaughnessy et al 2000, [72]	Läsförståelse	USA (n=15+15)	6 veckor	85,1 (10,00)	78,6 (8,90)	0,77 (0,03; 1,51)
Gustafson et al 2000, [95]	Fonologisk medvetenhet	Sverige (n=33+16)	5+4 månader	9,9 (2,30)	9,2 (2,80)	0,00 (-0,5971; 0,60)
Nag-Arulmani et al 2003, [84]	Fonologisk medvetenhet	Indien (n=31+28)	3 veckor	111,35 (15,28)	103,71 (18,88)	0,49 (-0,03; 1,00)

Bedömning av evidensstyrkan

I Tabell 7.16 sammanfattas det vetenskapliga underlaget för effekten av träning av enbart fonologisk medvetenhet på korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet.

- Det finns otillräckligt vetenskapligt underlag för att träning av enbart fonologisk medvetenhet, utan annan koppling till skriftspråket, jämfört med ordinarie undervisning har effekt på att förbättra korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet (⊕○○○).

Tabell 7.16 Effekt av träning av enbart fonologisk medvetenhet på korrekt läsning av ord, stavning, läsförståelse samt fonologisk medvetenhet för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Vetenskapligt underlag	Kommentar
Fonologisk medvetenhet	Korrekt läsning av ord	n=138 (3 randomiserade kontrollerade studier) ^a	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska samt stora brister i samstämmighet
Fonologisk medvetenhet	Stavning	n=99 (2 randomiserade kontrollerade studier) ^b	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska, brister i samstämmighet samt få studie-deltagare
Fonologisk medvetenhet	Läsförståelse	n=99 (2 randomiserade kontrollerade studier) ^b	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska, brister i samstämmighet samt få studie-deltagare
Fonologisk medvetenhet	Fonologisk medvetenhet	n=108 (2 randomiserade kontrollerade studier) ^a	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska, brister i samstämmighet samt få studie-deltagare

^a Nag-Arulmani, Gustafson [84,95]

^b Nag-Arulmani, O'Shaughnessy [72,84]

Träning av enbart läsflyt för utfallsmått läsning av ord på tid och läsförståelse

Beskrivning av resultatet

Tabell 7.17 sammanfattar resultaten av de studier som undersökte hur träning av enbart läsflyt, jämfört med ordinarie undervisning, påverkade läsning av ord på tid och läsförståelse.

Tabell 7.17 Sammanställning av studier som undersökte träning av läsflyt för utfallsmåten läsning av ord på tid och läsförståelse.

Författare År Referens	Utfallsmått	Population	Längd på intervention	Experiment Medelvärde (SD)	Kontroll Medelvärde (SD)	Effekt (Cohen's d)
Spencer et al 2010 [96]	Läsning av ord på tid	USA (n=121)	6 månader	2,87 (1,5)	2,75 (1,1)	0,59
Wexler et al 2010 [76]	Läsning av ord på tid	USA (n=122)	8 veckor	72,61 (15,49)	73,48 (16,38)	-0,5
Spencer et al 2010 [96]	Läsförståelse	USA (n=121)	6 månader	72,50 (8,00)	71,69 (9,7)	0,02
Wexler et al 2010 [76]	Läsförståelse	USA (n=122)	8 veckor	71,88 (14,21)	73,48 (18,04)	-0,10

Bedömning av evidensstyrkan

I Tabell 7.18 sammanfattas det vetenskapliga underlaget för effekten av träning av enbart läsflyt på läshastighet och läsförståelse

- Det finns otillräckligt vetenskapligt underlag för att träning av enbart läsflyt jämfört med ordinarie undervisning har effekt på att förbättra läsning av ord på tid och läsförståelse (⊕○○○).

Tabell 7.18 Effekt av träning av enbart läsflyt på läshastighet och läsförståelse för barn och ungdomar med dyslexi.

Insats	Utfallsmått	Antal deltagare (studier)	Vetenskapligt underlag	Kommentar
Läsflyt	Läsning av ord på tid	n=121 (2 randomiserade kontrollerade studier) ^a	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska, brister i samstämmighet samt få studiedeltagare
Läsflyt	Läsförståelse	n=122 (2 randomiserade kontrollerade studier) ^a	Otillräckligt ⊕○○○	Avdrag -3 pga att insatserna inte var identiska samt få studiedeltagare

^a Spencer, Wexler [76,96]

Icke-randomiserade studier

Sex kontrollerade studier uppfyllde inklusionskriterierna [88,97,98,99, 100,101]. Fyra studier jämförde en intervention med väntelista, dvs barnen fick samma intervention när studien hade avslutats. Två studier jämförde två interventioner sinsemellan där interventionerna bestod av träning av flera förmågor, dock inte i studien av Bhat och medarbetare som endast utvärderade träning av fonologisk medvetenhet.

Beskrivning av studierna

Bhat och medarbetare utvärderade effekten av enbart fonologisk medvetenhetsträning [97]. Två matchade grupper med vardera 20 elever från årskurs sex till åtta fick antingen intervention eller sattes på väntelista. Eleverna hade identifierats med inlärningssvårigheter och med svag fonologisk medvetenhet. Experimentgruppen förbättrades signifikant i fonologisk medvetenhet, enligt testresultat på CTOPP, jämfört med kontrollgruppen under den första interventionsperioden. Resultaten pekade på att fonologisk medvetenhet gick att träna upp hos elever med dyslexi. Ordavkodning eller non-ordläsning påverkades inte av interventionen.

Gelzheiser och medarbetare undersökte en individanpassad lästräning som var sammansatt med många olika inslag [98]. Träningen gavs av den ordinarie klassläraren på ett interaktivt sätt. Eleverna gick i skolår fyra och hade problem med läsning och gick därför i en individanpassad undervisningsgrupp. Experimentgruppen bestod av 26 elever som fick interventionen på höstterminen. Kontrollgruppen bestod av 22 elever som stod på väntelista och fick interventionen vårterminen efter. Studien har ett stort praktiskt-pedagogiskt värde. Man nämnde i studien att en mycket hög andel barn hade fått specialpedagogiskt stöd under lång tid, att en tredjedel tidigare hade gått om ett skolår, att majoriteten hade annat modersmål än engelska och visade tecken på socio-ekonomiska problem. De 26 eleverna i experimentgruppen hade vid slutet av höstterminen förbättrats signifikant, jämfört med kontrollgruppen, på både säkerhet i avkodning och läsförståelse men inte på läsflyt. Interventionen implementerades omedelbart i klassrummet, vilket gjorde det svårt att mäta effekter på kort och lång sikt. Man skulle kunna säga att forsk-

ningen i hög grad smälte samman med den pedagogiska konsekvensen i skolvardagen, något som skulle kunna liknas vid så kallad aktionsforskning. Skillnaden var att ett forskarteam ansvarade för studien, medan ordinarie lärare genomförde interventionen.

Lovett och medarbetare undersökte en intervention med träning av strukturerad fonem–grafem-koppling och läsförståelse [99]. I studien ingick 268 14-åringar i interventionsgruppen och 83 i kontrollgruppen. Träningen varade mellan 60 och 70 timmar. Vid mätning direkt efter interventionens slut hade interventionsgruppen förbättrad läsförståelse (effektstorlek (ES)=0,34), non-ordläsning (ES=0,45) och ordläsning (ES=0,15) i förhållande till kontrollgruppen.

Tijms utvärderade en datorbaserad systematiskt uppbyggd träning av avkodning för barn med dyslexi i åldrarna 9–12 år [100]. Komponenterna innefattade fonemkunskap, fonem–grafem-koppling, enstaviga ord, flerstaviga ord, morfologi, verbstruktur och lånord. Man granskade påverkan av andra möjliga individuella effekter, som kända riskfaktorer för dyslexi (fonologisk medvetenhet, verbalt korttidsminne och ordbenämning) samt kulturella och socioekonomiska faktorer. Resultaten visade att undersökningsgruppen förbättrade sin läs- och stavningsförmåga avsevärt i förhållande till jämförelsegruppen. Undersökningsgruppen låg efter insatsen t o m inom gränsen för normala färdigheter. Ett uppföljningstest visade att även läshastigheten ökade till att nå den lägre gränsen för normala färdigheter.

Olson utvärderade två datorbaserade träningsprogram [88]. Den ena gruppen elever tränade fonologi och den andra gruppen tränade läsförståelsestrategier. Studien visade att de flesta eleverna vidmakthöll eller förbättrade sina nivåer. I en uppföljningsstudie av eleverna, ett och två år efter interventionerna, visade resultaten att det inte fanns några skillnader mellan grupperna avseende ordigenkänning.

Wise och medarbetare jämförde också två datorbaserade interventioner [101]. I den ena, en kontextbaserad intervention, läste eleverna individuellt berättelser och lärde sig läsförståelsestrategier i grupp. I den andra, en fonologiskt baserad intervention, studerade eleverna fonologiska strategier i små grupper, medan den individuella träningen vid dator delades mellan fonologiska övningar och läsning av berättelser. Eleverna med fonologisk träning fick bättre resultat på tester av fonologiska förmågor och på ordläsning utan tidsbegränsning. Eleverna som tränat kontextbaserad läsning blev bättre på tidsbegränsad ordläsning. Eleverna på den lägre läsnivån drog större nytta av den fonologiska träningen än eleverna på den högre läsnivån.

Bedömning av effekter

Studierna undersökte olika interventioner och de effekter som observerades kunde därmed inte vägas samman i en metaanalys.

Effekten av strukturerad grafem–fonem-koppling var jämförbar med effekterna i randomiserade studier. De övriga interventionerna har inte undersökts i randomiserade studier.

Bedömning av evidensstyrka

Det vetenskapliga underlaget är otillräckligt för samtliga granskade interventioner som enbart har undersökts i en icke-randomiserad studie vardera.

Studien av Lovett och medarbetare som undersökte strukturerad grafem–fonem-koppling, stödjer de randomiserade studiernas resultat att interventionen har en måttlig effekt jämfört med väntelista [99]. Det vetenskapliga underlaget, baserat på icke-randomiserade studier, är dock otillräckligt eftersom det bara finns en enda studie som har medelhög kvalitet.

Systematisk översikt om musik

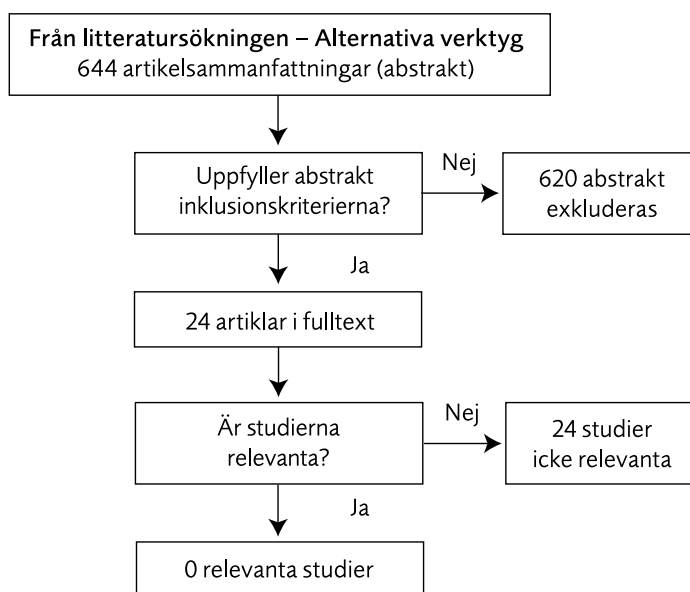
En systematisk översikt av Cochrane Collaboration hade tillräcklig kvalitet för att inkluderas [102]. Syftet med översikten var att studera effekten av musik för att förbättra läsning, läsförståelse, läsflyt, fonologisk medvetenhet och stavning. Översikten kom fram till att det saknades randomiserade kontrollerade studier, vilket betydde att inga relevanta studier kunde inkluderas i översikten.

Bedömning av evidensstyrkan

Det finns otillräckligt vetenskapligt underlag för att avgöra om musik, jämfört med en kontrollgrupp (väntelista eller ingen insats), förbättrar läsning, läsförståelse, läsflyt, fonologisk medvetenhet och stavning då det saknas relevanta studier ($\oplus\circ\circ\circ$).

8. Alternativa verktyg för barn och ungdomar med dyslexi – resultat av granskningen

Arbetet med gallring, relevans- och kvalitetsgradering resulterade i att inga studier av medelhög eller hög kvalitet identifierades.



Figur 8.1 Flödesschema över gallring och bedömning av artiklar för alternativa verktyg

Beskrivning av ingående studier

Litteratursökningen identifierade 644 artikelsammanfattningar på studier av alternativa verktyg, dvs hjälpmedel för att stödja, kompensera och utveckla läsförmåga som t ex appar i mobiltelefonen. Av dessa var inga artiklar relevanta.

Alternativa verktyg – beskrivning av resultatet

Inget resultat kunde beskrivas eftersom det saknades relevanta studier.

Bedömning av evidensstyrkan

- Det finns ett otillräckligt vetenskapligt underlag för att avgöra om alternativa verktyg förbättrar utfallsmåtten för barn och ungdomar med dyslexi, då det saknas relevanta studier (⊕○○○).

9. Etiska och sociala aspekter

Inledning

Barn och vuxna med svårigheter att läsa och skriva har historiskt sett ofta varit missförstådda. Otaliga vittnesmål finns från personer som blivit illa behandlade för att de inte lärt sig läsa i samma takt som övriga elever. De har tillskrivits egenskaper som att vara lata, slarviga och obegåvade. De har utsatts för vad man i yrkesetiska sammanhang kallar psykisk skada såsom att bli ringaktad och kanske blivit både utskällda och hånade för sin oförmåga. För många har det inneburit en livslång känsla av att ha ett lågt värde.

Vad säger skollagen?

År 2010 fastslogs i skollagen *Alla barn och elever ska ges den ledning och stimulans som de behöver i sitt lärande och sin personliga utveckling för att de utifrån sina egna förutsättningar ska kunna utvecklas så långt som möjligt enligt utbildningens mål.* Vidare står det *Om det inom ramen för undervisningen eller genom resultatet på ett nationellt prov, genom uppgifter från lärare, övrig skolpersonal, en elev eller en elevs vårdnadshavare eller på annat sätt framkommer att det kan befaras att en elev inte kommer att nå de kunskapskrav som minst ska uppnås, ska detta anmälas till rektorn. Rektorn ska se till att elevens behov av särskilt stöd skyndsamt utreds. Behovet av särskilt stöd ska även utredas om eleven uppvisar andra svårigheter i sin skolsituation. Samråd ska ske med elevhälsan om det inte är uppenbart obehövt. Om en utredning visar att en elev är i behov av särskilt stöd ska han eller hon ges sådant stöd.* Det står också *Ett åtgärdsprogram ska utarbetas för en elev som ges särskilt stöd. Av programmet ska det framgå vilka behoven är, hur de ska tillgodoses och hur åtgärderna ska följas upp och utvärderas. Elev och vårdnadshavare ska ges möjlighet att delta när ett åtgärdsprogram utarbetas.*

Resultat av litteratursökningen och urval av studier

Litteratursökningen för etiska och sociala aspekter gav 537 artikel-sammanfattningar. Av dessa var endast ett fåtal relevanta. Sökningen gjordes i databaserna PsycInfo och ERIC. Ytterligare studier hittades i sökningar som gjordes för tester och insatser. Studierna kvalitetsgranskades inte. Detta kapitel baseras dels på litteratur om de etiska och sociala perspektiven, dels på diskussioner i expertgruppen.

Utredning för barn med dyslexi

En dyslexiutredning är ofta påfrestande och omfattningen av testerna måste anpassas till det enskilda barnets reaktioner. Barnet kan utsättas för intrång i sin integritet om han eller hon genomgår många tester trots känslor av obehag eller olust. En utredning väcker förhoppning om snar hjälp som inte alltid infrias beroende på t ex resurs- eller kompetensbrist. Även när en utredning är genomförd och barnet har fått ett intyg kan insatserna dröja och ibland t o m helt utebli. En annan aspekt är att föräldrar och lärare kan ha olika uppfattning om en utredning ska påbörjas tidigt eller vänta. I sådana fall har utredaren ett etiskt ansvar att ge vägledning till ett informerat beslut. Utredarens kompetens är viktig och avgörande. Bristande kunskap och felanvända metoder utgör skaderisker. Kravet på kompetens gäller också för de som är ansvariga för att barnet sedan får adekvat pedagogisk hjälp.

Insatser för barn med dyslexi

Följande avsnitt handlar om etiska och sociala aspekter på de insatser som görs när ett barn utreds för dyslexi eller när barnet har fått en dyslexidiagnos och vikten av att de yrkesverksamma tar hänsyn till eleven och föräldrar utifrån dessa etiska och sociala aspekter [103].

Etiska principer

Många etiska principer tillämpas i olika professioners verksamheter. De uttolkas som principer, riktlinjer, regler och lagar för hur en professionell verksamhet ska bedrivas. Etiska principer och riktlinjer brukar uttala vad som är en professions specifika bidrag till samhället – det goda man vill göra – och vilka restriktioner som de professionella ålägger sig, dvs vad de anser att man inte får göra inom yrkets verksamhet (alternativt den skada man inte får göra).

Skydd mot skada är ett uppenbart och nödvändigt värde. Människor ska skyddas från kroppslig och psykisk skada och från att förlora sin hälsa och sitt välbefinnande. Professioners samhällsuppdrag är att göra gott och bidra till minskad skada i olika avseenden.

Integritet. Människor ska ha rätt att själva avgöra gränserna för vad som tillhör deras privata sfär. De har därför rätt till kontroll över vilken information om dem som andra ska ha tillgång till.

Självbestämmande. Människor ska vara fria att handla så länge deras handlingar inte strider mot andra värden. För att kunna använda rätten till självbestämmande måste individen ha tillgång till all information som kan tänkas påverka villigheten att delta i t ex forskning eller undersökning.

Lika rättigheter. Människor ska ha lika chanser att få del av det goda i samhället. Diskriminering pga kön, ras, religion, etniskt ursprung, sexuell läggning eller politisk tillhörighet berövar människor dessa chanser.

Rättvisa enligt behovsprincipen innebär att den som har större behov ska ges mer hjälp än den som har mindre behov. Rättvisepincipen, likhetsprincipen och behovsprincipen kan förefalla motstridiga men de är möjliga att förena. Åtskillnad ska göras med anledning av olika behov, inte på grundval av olikheter i de nämnda diskrimineringsgrunderna.

Forskningsetik

De sociala värdena påverkar forskningens ramar. Forskning om t ex dyslexi planeras och genomförs med ovan nämnda värden och principer samt med regler som etiska riktningsgivare. Forskning ska genomföras utan att skada forskningspersonerna. De har rätt att själva avgöra vilken information de låter forskarna ta del av. Vidare har de rätt att bestämma om de vill delta i forskningen samt rätt att avbryta sitt deltagande. De ska också behandlas lika och ha samma rätt att få del av forskningens nytta oavsett olikheter i bakgrundsfaktorer.

Sammantaget utgör dessa tillämpningar inom forskningspraktiken de viktigaste delarna av det vi kallar forskningsetik, dvs yrkesetik för forskare. Det är viktigt att tillägga att även forskningens kvalitet utgör ett etiskt värde. Forskning av låg vetenskaplig kvalitet leder inte till användbara och pålitliga slutsatser, men kan ändå ta samhällsresurser i anspråk. Det är dock inte forskningsetik som vi ska uppehålla oss vid utan andra professioners yrkesetik, dvs de professioner som möter elever och deras föräldrar i samband med utredning och åtgärder vid dyslexi.

Yrkesetik för professioner involverade vid dyslexi

Slutsatserna av den forskning om tester och insatser vid dyslexi, som redovisas i rapporten, kommer att kunna tillämpas av flera professioner såsom lärare, speciallärare, specialpedagoger, logopedier och psykologer. Även läkare och sjuksköterskor kan komma att involveras.

När man som professionell åtar sig uppdraget att hjälpa elever som ska utredas, eller som har en dyslexidiagnos, gäller samma överordnade värden som för all professionsetik. De praktikinriktade tillämpningarna i form av riktlinjer, principer och regler finns dock inte formulerade för just insatser vid dyslexi. Psykologer, logopedier, lärare, läkare och sjuksköterskor har etiska regelverk som gäller för deras verksamheter. Dessa etiska principer gäller således även för personer med dyslexi. De etiska principerna är emellertid inte alltid tydligt formulerade eller preciserade. Lärarnas etiska riktlinjer lämnar t ex stort utrymme för pedagogernas omdömen och egna överväganden [104]. Att skolan har ett ansvar för

elever som inte når målen, vare sig eleven har en diagnos eller inte, är tydligt.

Om man vill utveckla yrkesetiken för de professioner som är involverade vid dyslexi finns det två vägar att gå. Dessa utesluter inte varandra. Den ena vägen är att studera hur man hanterar yrkesetiska aspekter i andra länder och vad man betonar i sina etiska regelverk. Den andra vägen är att utgå från kända etiska risker.

Gemensamma etiska principer

Ett exempel är de gemensamma etiska principerna för professioner som arbetar med inläringssvårigheter (learning disabilities) och som anslutit sig till Council for Learning Disabilities [105]. De består av sex huvudprinciper med 3–6 preciseringar. Principerna fokuserar i första hand på elevernas och föräldrarnas rättigheter samt på kraven på kompetens och vikten av att professionerna ansluter sig till en forskningsetisk standard. Vidare understryker de vikten av samarbete för elevernas bästa. Se Exempel 9.1 för en mera detaljerad beskrivning.

Gemensamma etiska principer kan vara värdefulla när flera professioner är aktiva inom samma problemområde, så kallad områdesetik. Se Kapitel 11 där vi har konstaterat att rätten att ställa diagnos är oreglerad i Sverige. Flera professioner anlitas för utredning av dyslexi och de använder ofta olika definitioner och olika tester i utredningarna. Det är angeläget, och det tillhör yrkesetisk standard, att professionerna trots sina olikheter gemensamt sörjer för att både utredning och åtgärder gagnar eleverna och föräldrarna. Professionell oenighet ska inte belasta dem.

Att skapa gemensamma etiska riktlinjer för en specifik verksamhet förefaller inte realistiskt med tanke på att flera professioner sedan många år har sina egna yrkesetiska principer. Detta hindrar inte att man kan studera och ta intryck av de etiska principer som har formulerats i andra länder.

Exempel 9.1. Etiska principer i USA (källa: Council for Learning Disabilities).

Elevers och familjers rättigheter. De professionella ska bl a motverka diskriminering och verka för elevers lika chanser att få del av professionernas tjänster, de ska upprätthålla konfidentialitet samt upplysa barn och föräldrar om deras juridiska rättigheter och ansvar.

Den professionellas kompetens, integritet och professionalism. De professionella ska bl a utöva sin praktik inom ramarna för sin behörighet och utbildning, undvika intressekonflikter, upprätthålla sin kompetens samt till lämplig instans rapportera etiska brott i praktiskt yrkesutövande.

Vid medverkan i forskning ska den professionella bl a följa relevanta etiska regelverk, rapportera forskningsresultat korrekt med slutsatser som håller sig inom studiens begränsningar samt underlätta forskningsresultatens tillämpning i praktiken.

De professionella bidrar till undervisningens utveckling och ska bl a använda forskningsresultat för att välja interventioner och program som riktar sig till elever med inlärningssvårigheter, övervaka interventioners effektivitet samt använda bedömningsinstrument på ett valitt och reliabelt sätt.

De professionella ska samarbeta med andra professioner för elevers bästa. Den professionella ska bl a yppa konfidentiell information endast till de som enligt lagen har rätt till informationen samt vid samarbete i undervisning eftersträva att arbeta på likartat sätt.

De professionella förespråkar lagar, regler och policy för elevers bästa. De ska bl a ansluta sig till rådande lagar, tala för modifieringar av lagar och regler i elevers intresse samt underlätta möjligheten att förstå lagar, regler och policy.

Kända etiska risker

Ett annat sätt att utveckla yrkesetiken för de professioner som är involverade vid dyslexi är att utgå från identifierade etiska risker genom systematisk forskning eller genom inventering av risker som framkommit genom beprövad erfarenhet. Det senare innebär att man utgår från kända etiska risker, så kallade critical incidents. Dessa etiska risker, vilka hotar något viktigt värde, dokumenteras och utgör underlag för

utbildning, vägledning och överenskommelser. Här följer exempel på etiska risker som kan uppträda i samband med utredning av dyslexi eller undervisning av elever med dyslexidiagnos. Dessa kritiska händelser och etiska risker är i första hand exempel från experters beprövade erfarenhet.

Risk för skada vid utredning av dyslexi. En utredning är påfrestande och det kan betraktas som skada om en elev blir ålagd att genomgå alltför många tester. Skada genom en överanvändning av tester kan också uppstå om den som utreder oavbrutet erbjuder nya tester trots obehag och olust. Testerna genomförs förvisso för ett gott ändamål, men måste anpassas till elevens reaktioner. Det kan också finnas en fara med att det finns många starka kommersiella intressen i dyslexitester, och elevernas behov kan då komma att åsidosättas.

En annan aspekt på att en utredning kan vålla skada är om den inte följs av adekvata åtgärder. Att beviljas utredning väcker förhoppningar. Dessa infrias inte alltid vilket kan bero på resursbrist eller kompetensbrist. Vidare händer det att elever kan få vänta på utredning och stå på väntelista (senare insats). Även när utredningen är genomförd och intyg har skrivits kan eleven få vänta på insatser och ibland uteblir insatserna helt.

Tidpunkten för en utredning väcker frågor. Föräldrar och lärare kan ha olika uppfattning om när en utredning ska påbörjas. Det kan vara svårt att avgöra vad som är bäst för eleven. Frågan är vilket som innebär störst skaderisk, att utreda tidigt eller att vänta. I sådana fall har den professionella utredaren ett etiskt ansvar att ge vägledning till ett informerat beslut.

Utredarens *kompetens* är viktig och avgörande. Den som är utsedd att utreda och som har förskrivningsrätt ska ha adekvat kompetens. De metoder och tester som används ska vara evidensbaserade och användas enligt anvisningarna. Bristande kompetens och felanvända metoder utgör skaderisker. Det gäller i lika hög grad de som är ansvariga för att eleven får adekvat pedagogisk hjälp.

Rätten till integritet innebär att det är eleven och föräldrarna som avgör i vilken grad diagnosen ska vara privat eller offentlig. Det ställer krav på anpassning av de professionella i skolans vardag. Elever och föräldrar kan ha olika uppfattning om t ex risken för stigmatisering.

Rätten till självbestämmande ställer krav på de professionella i samband med såväl utredning som åtgärder. Ett problem som kan uppstå är om elev och förälder har olika önsknings eller uppfattningar om vad eleven ska delta i och vad som ska göras. Det kan vara svårt att avgöra om det är elevens eller föräldrarnas rätt till självbestämmande som ska gälla.

Risk för orättvisa och ojämlikhet mellan elever kan uppstå. Alla elever har rätt till lika chanser för att få hjälp för sina svårigheter. Skollagen anger ”Om en utredning visar att en elev är i behov av särskilt stöd ska han eller hon ges sådant stöd” (3 kap 8 §). Elever med dyslexidiagnos har lagstadgad rätt till hjälp om de behöver. Det finns en risk att elever med stora och outredda svårigheter får mindre hjälp eller ingen hjälp alls, vilket står i strid med skollagen (2010:800).

Tillgången till kompetens är en rättvisefråga. Bristande eller varierad utbildning kan leda till ojämlikhet i diagnostik och behandling för elever med förmodad dyslexi från olika delar av landet. Olika kommuner prioriterar olika enligt det kommunala självstyret. Där det finns otillräcklig tillgång till professionell personal kan elever få vänta på adekvata åtgärder och i värsta fall bli helt utan.

Sociala aspekter

Etiska och sociala aspekter är inte uteslutande kategorier, de är snarare sammanflätade. Handlingar och förhållanden som är oförenliga med etiken kan ofta få sociala konsekvenser. De sociala aspekterna i samband med dyslexi rör elevernas deltagande i socialt liv både i skolan och på fritiden. Det gäller kamratrelationer och tillgång till fritidsaktiviteter, där deras svårigheter inte ska vara hindrande. På längre sikt gäller det personsituation och deltagande i t ex utbildning, arbetsliv och samhällsliv. Här följer några sådana aspekter.

Låg självkänsla

Risken för att barn med dyslexi utvecklar lägre självkänsla och högre grad av ängslan än andra barn har uppmärksammats och studerats. En officiell diagnos har positiv betydelse för självkänslan och ängslan [106]. En specifik diagnos som dyslexi har befunnits ge positiva effekter, medan en generell beskrivning däremot har negativ effekt [107]. Vi kan således anta att elever med stora svårigheter att läsa och skriva, i synnerhet de som inte får en diagnos, lever med låg självkänsla och ängslan. Det kan ge konsekvenser senare i livet.

Stödinsatser

I svensk högskola gällde lagen om likabehandling under åren 2002–2008. Den gav studenter med skriftspråkliga svårigheter rätt till kompensatoriskt stöd. Den ersattes år 2009 av diskrimineringslagen vilken innebar att högskolan skulle anpassa undervisningen för att öka tillgängligheten. År 2009 beviljades i Sverige 6 000 studenter kompensatoriskt stöd varav 3 600 med skriftspråkliga svårigheter. Stödinsatserna utgjordes av bl a rättstavningsprogram och talsyntes, förlängd tid vid tentamen och litteratur inläst som talbok. I en studie av sex svenska lärosäten år 2011, fann man att studietakten för studenter med skriftspråkliga svårigheter var mellan 3 och 13 högskolepoäng lägre än de förväntade 30 [108]. Flera av studenterna visade stark motivation för studierna, vilket är en förutsättning för att vara en framgångsrik student. Den lägre studietakten kan dock få kännbara sociala konsekvenser eftersom de ackumulerade uteblivna poängen kan leda till att de inte får studiemedel. Eftersom studenterna dessutom behöver all tid till sina studier finns inte heller utrymme till att arbeta vid sidan av studierna, och de kan således tvingas avbryta studierna av ekonomiska skäl.

Svårigheter i vuxen ålder

Vuxnas problem med läsning och skrivning visar sig oftast vara kvarstående svårigheter sedan barndomen. Det går inte att undvika läsning och skrivning i vårt samhälle. Därför utvecklas mer eller mindre framgångsrika strategier för att klara studier, arbetsliv och samhällsliv. En studie av industriarbetare med skriftspråkliga svårigheter visar att de valt sitt

yrke i tron att de skulle få arbeta mest med händerna [109]. Likafullt ställs de inför sådana krav på läs- och skrivförmåga som vållar problem. Undvikande och/eller bearbetande strategier används för att möta de dagliga kraven. Att få utökad tid för att lösa arbetsuppgifterna är viktigt. Det framkommer att de skriftspråkliga kraven har ökat med datoriserad teknik, men samtidigt kan de kompensatoriska insatserna (t ex i form av appar) visa sig vara till stor hjälp i framtiden. Vidare visar det sig att upplevelsen av otillräcklighet och att betraktas som okunnig och inkompetent har följt med från skolåren in i vuxen ålder och i arbetslivet. Att låg självkänsla riskerar att följa individen bekräftas i longitudinella studier. Tidig och väl utförd utredning och diagnos samt pedagogiska insatser tycks ge förutsättningar för högre självkänsla som vuxen [110,111].

Sammanfattningsvis kan vi konstatera att de professioner som utreder och hjälper elever med dyslexi har många etiska hänsyn att beakta. Det bemötande eleven får i skolan kan vara avgörande för hans eller hennes framtida liv. I och med att de flesta tester och insatser inte är så utvärderade att det går att dra några slutsatser om deras nytta för barn med dyslexi är det svårt att välja metoder som har evidens. Tills forskningsläget har förbättrats ställs förhoppningarna till möjliggörande av pedagogisk forskning utifrån den problematik som är kännetecknande för dyslexi.

Det kvarstår också en hel del som måste hanteras individuellt beroende på elevens och föräldrarnas behov och förmåga och som kräver en etisk känslighet, ett omdöme och en kompetens. För att bistå de professionella vore det önskvärt att genomföra forskning genom en systematisk inventering av kritiska händelser i mötet mellan professionella, elever och föräldrar. Alla dessa tre grupper bör i kommande studier tillfrågas om sina erfarenheter.

10. Hälsoekonomiska aspekter

Inledning

Dyslexi är sannolikt behäftad med höga samhällskostnader till följd av minskade möjligheter för utbildning och yrkeskarriär men det saknas välgjorda kostnadsberäkningar. Dessa indirekta kostnader är antagligen de största kostnaderna utifrån ett samhällsperspektiv, men även direkta kostnader såsom utredning, diagnostisering och insatser vid dyslexi ska inkluderas i de totala samhällsrelaterade kostnaderna. För en presentation av de hälsoekonomiska begrepp och metoder som används i detta kapitel hänvisas till SBU:s metodbok [2].

Endast för insatser som har vetenskapligt underlag är det aktuellt att bedöma deras kostnadseffektivitet. De enda insatserna som i den här rapporten har visats ha vetenskapligt underlag för effekt, är de som på ett strukturerat sätt fokuserat på att pedagogiskt stärka relationen mellan bokstäver (grafem) och språkljud (fonem). Eftersom det saknas vetenskapligt underlag för de utvärderade testerna (med undantag för vissa enstaka internationella deltester) går det inte att uttala sig om deras kostnadseffektivitet. Däremot kan man uppskatta de samhällsekonomiska kostnaderna för att förutsäga, upptäcka, utreda och sätta in insatser för dyslexi hos barn och ungdomar.

Frågeställningar

- Vad kostar det utifrån ett samhällsrelaterat perspektiv att med tester upptäcka och utreda dyslexi?
- Vad är den samhällsekonomiska kostnaden för olika former av insatser vid dyslexi?
- Vad är kostnadseffektiviteten av pedagogiska insatser för att stärka relationen mellan bokstäver och språkljud vid dyslexi (strukturerad fonem–grafem-koppling)?

Resultat av litteratursökningen och urval av studier

Den hälsoekonomiska litteratursökningen genomfördes med samma söktermer som för de övriga frågeställningarna i rapporten, men med tillägg av relevanta hälsoekonomiska termer. Totalt resulterade sökningen i 285 artikelsammanfattningar. Av dessa beställdes 10 i fulltext. Ingen studie bedömdes uppfylla våra inklusionskriterier. En förteckning över exkluderade studier finns i Bilaga 2.

Analyser och egna beräkningar

Det saknas vetenskapligt underlag som besvarar de hälsoekonomiska frågeställningarna om de olika testerna för att upptäcka och utreda dyslexi samt för insatsernas kostnadseffektivitet. Det finns emellertid studier som belyser viktiga hälsoekonomiska frågor, och dessa diskuteras i det här avsnittet. Vi har gjort egna kostnadsberäkningar avseende tidsåtgången för insatsen att träna relationen mellan bokstäver och språkljud vid dyslexi (fonem–grafem-koppling) och analyserat kostnadseffektiviteten av denna insats. Detta avsnitt ska inte ses som vetenskaplig evidens för de uppställda frågeställningarna, utan syftar till att hjälpa beslutsfattare att bedöma dessa frågeställningar utifrån bästa tillgängliga kunskap.

Kostnad för test och utredning

McBride och medarbetare har i en studie beräknat kostnaden för utredning av läsproblem [38]. Studien genomfördes på 633 barn, från förskola till andra klass i skolan, från sju delstater i USA. Ett av instrumenten som studerades, Dynamic Indicators of Basic Early Literacy Skills (DIBELS), ingår i den här rapporten [54]. Det visade sig att den totala kostnaden vid användning av DIBELS var cirka 130 kronor per barn. Denna kostnad inbegrep direkta kostnader för testet (avgifter) samt en värdering av den tid som lärarna använde. Det är svårt att bedöma om beräkningen är relevant för Sverige då de olika kostnadsposterna (lön etc) inte presenterades i studien. I studien gjordes kostnadsjämförelser med andra instrument där den lägsta kostnaden uppnåddes med STAR Early Literacy [112] och den högsta kostnaden uppnåddes med Texas Primary

Reading Inventory (TPRI) [64]. Inga av dessa instrument används dock i Sverige.

Övriga tester (internationella/svenska) som inkluderats i vår rapport saknar kostnadsberäkningar i den vetenskapliga litteraturen.

Kostnaden för användning av testerna påverkas dels av priset för instrumentet, dels av tidsåtgången för att genomföra en utredning. På lång sikt får också förmågan att ställa en korrekt diagnos påverkan på kostnader. För att testernas kostnadseffektivitet ska kunna beräknas behöver kostnaderna ställas i relation till testernas sensitivitet och specificitet samt till de konsekvenser som dessa leder till. Dessa uppgifter saknar vetenskapligt underlag och det blir därmed inte möjligt att göra relevanta beräkningar av deras kostnadseffektivitet.

Kostnad för alternativa verktyg

Kostnaderna för insatser med alternativa verktyg vid dyslexi beror på pris för verktyget, underhåll, utbildning och support. Erfarenheter hos experterna i rapporten har varit att inköpskostnaden för verktygen ofta är relativt låg, medan kostnaden för utbildning och support ofta varit mycket hög. Kostnaden för support till de alternativa verktygen har ofta missbedömts vid beslutsfattande om inköp och har grovt underskattats (eller helt utelämnats). För att användningen av alternativa verktyg ska vara kostnadseffektiv krävs att verktygen ger effekt, något som inte kan fastslås i denna rapport.

Kostnad för insatser för att träna relationen mellan bokstäver och språkljud

För att bedöma kostnaden för träning av relationen mellan bokstäver och språkljud (strukturerad fonem–grafem-koppling) har vi utgått från de nio studierna som bedömts som relevanta i Kapitel 7 och gjort egna kostnadsberäkningar för dem [72,77,78,80,90–94]. Dessa studier varierade i omfattning och intensitet, vilket påverkar kostnaden för insatsen. Tidsintervallet för interventionen i studierna var mellan fyra veckor och

åtta månader samt mellan 20 sessioner om vardera 20 minuter till över 100 sessioner om vardera 50 minuter. I vissa studier gavs träning i grupp och i andra gavs den individuellt. Dessutom varierade följsamheten till träningen vilket också påverkar kostnaden. Här redovisas därför kostnadsparametrarna separat för att möjliggöra kostnadsberäkningar av olika intensiva insatser. Dessutom redovisas dessa parametrar specifikt för en studie från Sverige då den bedömdes ligga nära den aktuella och relevanta frågeställningen [90] samt för en studie från Holland [113] som även gjort beräkningar avseende kostnader och kostnadseffektivitet för insatsen.

Tabell 10.1 Karaktäristika avseende studierna för kostnadsberäkningar.

Träningens karaktäristika	Intervall samtliga inkluderade studier	Basfallet [90]	Hakaart-van Roijen [113]
Längd på interventionen	4 veckor–8 månader	12 veckor	18 månader
Antal minuter per tillfälle	20–50 minuter	40 minuter	50 minuter (plus läxa)
Antal tillfällen	20–126 tillfällen	60 tillfällen	78 tillfällen
Kunskap hos läraren	–	Särskilt utbildad lärare	Psykolog eller särskilt utbildad lärare
Storlek på grupp	1–5 deltagare	1 deltagare	1 deltagare

Kostnaden för undervisning motsvarar minst den tid som läraren genomför träning med sina deltagare, men ofta tillkommer både initial träning och förberedelse inför varje session. I basfallet antas att varje 40-minuterstillfälle tar totalt 60 minuter i anspråk av lärarens tid och att läraren redan har kompetens. Det gör att 60 timmar behövs per lärare i basfallet. Medellönen för speciallärare i Sverige år 2012 var 29 600 kronor. Efter ett påslag på 45 procent för arbetsgivaravgifter och skatter blir den totala lönekostnaden per månad 42 920 kronor. Det motsvarar 270 kronor per timme (givet 160 timmar per månad).

Totalt blir därför kostnaden 16 200 kronor (270 kr x 60 timmar) per deltagare som tränas i linje med basfallet.

Givet de nio olika studiernas intensitet och omfattning varierar kostnaden per träningsprogram från endast några tusen kronor till drygt 30 000 kronor. Det beror på träningens omfattning och om träningen sker individuellt (en deltagare per lärare) eller i grupp (upp till fem deltagare per lärare). Det saknas dock studier som beskriver vilken intensitet eller undervisningsform som är den mest effektiva respektive kostnadseffektiva.

I kostnadsberäkningen i den holländska studien bedömdes insatsen kosta cirka 32 000 kronor [113]. Vår skattning för svenska förhållanden var ungefär hälften av detta. Det beror främst på att man i den holländska studien räknade med en kostnad per timme på motsvarande 650 kronor, medan den svenska timkostnaden beräknades vara 270 kronor. Denna stora diskrepans är svårförklarad även om lönerna kan skilja sig åt. En översikt från Storbritannien beräknade kostnaden för riktad träning till barn med språksvårigheter till cirka 5 000 kronor om den genomfördes i grupp med assistent och till 12 000 kronor om den genomfördes med endast ett barn med speciallärare [114]. Dessa beräkningar ligger i linje med de svenska skattningarna.

Kostnadseffektivitet för insatser för att träna relationen mellan bokstäver och språkljud

Den enda studie som identifierats som analyserade kostnadseffektiviteten av att träna relationen mellan bokstäver och språkljud är gjord av Hakkaart-van Roijen och medarbetare, och insatsen genomfördes med hjälp av ett protokoll (The dyslexia protocol, based on a computer-aided treatment programme Gramma) [113]. Protokollet har inte utvärderats i den här rapporten, men det påminner om flera andra träningsprogram. Kostnaden för denna insats i studien beskrevs i stycket ovan. I studien användes en modell för att beräkna långsiktiga kostnader och effekter av träningsinsatsen. Insatsen jämfördes med en icke evidensbaserad träning av talpedagog eller läkare alternativt ingen träning alls, och detta benämndes som ”träning som vanligt”. I den interventionsstudie som låg till

grund för den hälsoekonomiska analysen fanns ingen kontrollgrupp, men i modellberäkningen gjordes antaganden om vad träning som vanligt skulle innebära. Interventionsstudien bestod av 193 barn och ungdomar mellan 7 och 14 år som alla hade diagnostiserats med dyslexi. Insatsen att träna med protokoll hade en signifikant positiv effekt på läsningen och stavningen jämfört med innan träningen påbörjades. I studien fick 40 procent av deltagarna normal läs- och skrivförmåga efter insatsen med protokoll. Vid jämförelsealternativet ”träning som vanligt” antogs att ingen fick normal läs- och skrivförmåga.

I studien av Hakkaart-van Roijen och medarbetare mättes effekterna av livskvaliteten i kvalitetsjusterade levnadsår (QALY) [113]. Det är ett rekommenderat utfallsmått i hälsoekonomiska beräkningar och det kombinerar livskvalitet med tiden som ett tillstånd varar. QALY-vikterna togs fram genom instrumentet EQ-5D. Instrumentet besvarades dock inte av barnen och ungdomarna själva, utan av kliniker som fick ta ställning till hypotetiska personer med allvarlig respektive mild dyslexi. För att göra om svaren på EQ-5D till QALY-vikter användes en holländsk tariff som motsvarade allmänhetens preferenser för olika hälsotillstånd. QALY-vikten för personer med allvarlig dyslexi var med denna metod 0,77 jämfört med 0,88 för övriga, på en skala mellan 0 (död) och 1 (full hälsa). Allvarlig dyslexi förväntades därmed medföra en livskvalitetsförsämring på 12,5 procent. Det författarna kallade mild dyslexi antogs inte föranleda någon livskvalitetsförlust.

Kostnaderna i analysen inkluderade både direkta medicinska kostnader och indirekta icke-medicinska kostnader som t ex extra lärarstöd. Däremot inkluderades inte kostnader för produktionsbortfall. Kostnaden för diagnostik beräknades till cirka 8 000 kronor per barn. Behandling med protokollet beräknades till cirka 32 000 kronor. Kostnaden per barn blev totalt cirka 40 000 kronor. Kostnadsberäkningen för jämförelsealternativet ”träning som vanligt” utgick från register och expertbedömningar. Det visade sig att 20 procent av barnen och ungdomarna fick hjälp av speciallärare, 15 procent fick hjälp av skolläkare och en liknande andel besökte allmänläkare för sina läs- och skrivsvårigheter. Kostnaden för ”träning som vanligt” beräknades till cirka 1 800 kronor per barn plus

en årlig kostnad på cirka 4 500 kronor för de barn som behövde extra lärarstöd.

Analysen gjordes för en tidshorisont om 6, 12 respektive 18 år. Kostnaden per lyckad behandling av allvarlig dyslexi var 327 000 kronor. Kostnaden per vunnet QALY beräknades till 528 000 kronor på 6 års sikt, 237 000 kronor på 12 års sikt och 159 000 kronor på 18 års sikt. Olika känslighetsanalyser presenterades men de hade endast mindre påverkan på resultaten.

Tabell 10.2 Viktiga parametrar och resultat i beräkningarna.

	Värde
Diagnostik	8 000 kr
Träning genom protokoll	32 000 kr
Kostnad "vård som vanligt"	6 500 kr
QALY-viktsförlust allvarlig dyslexi	0,11
QALY-viktsförlust mild dyslexi	0
Andel deltagare med normal läs- och skrivförmåga efter behandling	40 %
Andel deltagare med normal läs- och skrivförmåga utan behandling	0 %
Kostnad per QALY, 6 års tidshorisont	528 000 kr
Kostnad per QALY, 12 års tidshorisont	237 000 kr
Kostnad per QALY, 18 års tidshorisont	159 000 kr

Studien av Hakkaart-van Roijen och medarbetare hade flera brister, men den visade på några viktiga samband [113]. De stora bristerna var avsaknad av kontrollgrupp i interventionsstudien samt användandet av kliniker för att skatta barns- och ungdomars livskvalitet (via EQ-5D). Den studerade insatsen är inte tillgänglig i Sverige, men den påminner däremot om relevanta insatser. Det finns dock mycket som talar för att författarnas slutsatser även skulle kunna gälla insatser med liknande träning i Sverige,

även om de exakta beräkningarna kan skilja sig åt. De direkta kostnaderna för insatsen beräknades vara lägre i Sverige än de som beräknades i studien av Hakkaart-van Roijen och medarbetare, vilket skulle förbättra insatsens kostnadseffektivitet i Sverige [113]. När det gäller effekterna har den holländska studien räknat med att 40 procent av deltagarna fick normal läs- och skrivförmåga efter träningen, vilket kan vara en över-skattning. Det är också en grov indelning som inte fångar att många deltagare fick en viss effekt utan att för den skull få normal läs- och skrivförmåga. I den här rapporten har vi funnit att träningen har effekt på utfallsmåtten läsförmåga, stavning, läsförståelse, läshastighet och förmåga att uppmärksamma språkets ljudmässiga uppbyggnad. Vi kan dock inte uttala oss om hur många som får normal läs- och skrivförmåga. Om effekten av träningen förväntas vara lägre än 40 procent gör det att kostnadseffektiviteten försämras i jämförelse med de resultat som studien redovisar. Om exempelvis effekten skulle vara att 20 procent får normal läs- och skrivförmåga stiger kostnaden per QALY till det dubbla, vilket på 18 års sikt skulle innebära 318 000 kronor. Givet att endast 10 procent får effekt hamnar samma beräkning på 636 000 kronor.

Om resultatet från Hakkaart-van Roijen och medarbetares studie anses vara kostnadseffektivt eller inte beror på hur mycket samhällets betalningsvilja för en QALY är och på det saknas exakta definitioner [113]. Socialstyrelsens bedömningar säger dock att en kostnad per QALY på mellan 100 000 och 500 000 kronor är en måttlig kostnad. Träning av strukturerad fonem–grafem-koppling vid dyslexi kan därför anses ha en kostnadseffektivitet som antagligen befinner sig inom detta intervall och kostnadseffektiviteten förbättras ju längre tidshorisont som studeras. Det är dock viktigt att beakta att samhällets betalningsvilja för insatser även påverkas av hälsotillståndets svårighetsgrad.

Tidiga respektive sena insatser

Det skulle vara av intresse att undersöka i vilken ålder det är lämpligast att utreda och vid behov sätta in insatser för barn med dyslexi. I Finland görs t ex fler tidiga insatser och i Sverige görs fler sena insatser. De direkta kostnaderna för insatserna kan därmed vara lika stora, men de kan ge

olika effekt och därmed olika kostnadseffektivitet. Detta har inte studerats närmare i den här rapporten.

Sammanfattningsvis

Det saknas vetenskapligt underlag för hälsoekonomiska aspekter angående tester och insatser för barn och ungdomar med dyslexi.

För insatser som fokuserar på träning av strukturerad fonem–grafemkoppling finns vetenskapligt underlag för en positiv effekt, och därmed är det av värde att beräkna kostnadseffektiviteten av denna insats. I analyserna i det här avsnittet har kostnadseffektiviteten av sådan träning bedömts vara inom intervallet 100 000 till 500 000 kronor per QALY. För övriga insatser saknas vetenskapligt underlag för effekt, och det är därmed svårt att göra relevanta hälsoekonomiska bedömningar.

Det är värt att påminna om att hälsoekonomiska analyser inte syftar till att reducera kostnader, utan till att maximera samhällets totala hälsa med begränsade resurser. Det gör att insatser som har effekt kan uppfattas som kostnadseffektiva även om de medför en ökad kostnad. För personer med dyslexi kan dessutom effektiva insatser påverka kostnader och hälsa under mycket lång tid (en livstid) vilket gör att de har stor potential att bli kostnadseffektiva.

11. Inventering av tester som används i Sverige

Tillgängliga svenska tester

Skolinspektionen visade i en granskning år 2011 att skolorna i landet använde sig av många olika utredningsmaterial för att upptäcka och utreda elever med dyslexi (Skolinspektionen 2011:8). Detta kapitel kompletterar vår systematiska litteraturgranskning med en sammanställning av de svenska tester som finns tillgängliga för att upptäcka och utreda barn och ungdomar med dyslexi (Tabell 11.1).

Begreppet test används här som ett samlingsnamn för screening- och utredningstester. Sammanställningen har gjorts tillsammans med bl a Specialpedagogiska skolmyndigheten som bidragit med en förteckning över svenska tester. Vi kan inte utesluta att det kan förekomma fler tester som vi inte har identifierat.

Vi gick igenom manualerna för respektive test (några tester saknade helt manualer), för att undersöka om det fanns psykometriska data över testresultatens reliabilitet och validitet (Bilaga 4 – Mall: Utvärdering av svenska tester). Med reliabilitet menas att det finns data på testets tillförlitlighet, t ex mellan olika testare och testtillfällen. Med validitet menas att testet mäter det som det avser att mäta. I flera manualer fanns både reliabilitets- och validitetsmått rapporterade. Dessa tester är markerade med ett plustecken (+) i kolumnen för reliabilitet och validitet i Tabell 11.1. Vissa tester hade enbart normeringsdata. Med normeringsdata menas att det finns uppgifter om medelvärde och standardavvikelse i ett stickprov ur den population som testet är avsett att användas för.

En viktig komponent vid valet av ett test är att det finns vetenskapligt underlag för en god diagnostisk tillförlitlighet. Det finns också andra aspekter som bör beaktas. Det kan t ex vara om översättningen till svenska har gjorts med godkännande av upphovsmannen och att över-

sättningen är validerad genom återöversättning. Om översättningen inte är gjord ”enligt konstens alla regler” finns risk att frågorna kan tolkas annorlunda än vad som var tänkt. Att översätta språkliga tester är inte lätt. Språk skiljer sig åt på många olika sätt. Det kan t ex vara på det fonologiska systemet (ljudsystem), på ordförrådet (semantik), på grammatiska regler och strukturer av ordens böjningar och på sammanföningar i satser (morfologi och syntax). Ibland kan det vara omöjligt att direktöversätta en uppgift från ett språk till ett annat. När man översätter ett test och normerar det på svenska barn, är det möjligt att ange t ex stanine- och percentilvärden samt åldersgrupp så att barnens resultat på testet blir möjliga att värdera och följa över tid, dvs blir jämförbara med sig själva och med andra.

I den brokiga floran av tester som vi har identifierat, finns de som håller en högre standard än andra, dvs där det förekommer mått på testresultatens reliabilitet och validitet i manualerna. Inget svenskt test fick dock tillräckligt vetenskapligt underlag enligt våra kriterier. För det skulle krävas att testet hade redovisats i en publicerad forskningsartikel där det visat sig vara tillförlitligt för att upptäcka och/eller utreda läs- och skrivsvårigheter som är karaktäristiska för dyslexi bland barn och ungdomar mellan 6 och 20 år, som brukar det alfabetiska skriftspråket. Vi kan konstatera att Sverige här är i behov av nya forskningsinsatser.

Användande av de olika testerna

I SBU:s uppdrag ingick inte att kartlägga eller bedöma olika modeller för utredning. Det gäller såväl vem som gör utredningen, vad som ingår och hur utredningen genomförs. Däremot kan vi konstatera att det saknas enhetlighet inom såväl testurval som utredningsförfarande. Vi kan konstatera att vem-frågan ofta lyfts i samhällsdebatter och i praxissammanhang – inte sällan inom ramen för intressebaserade eller fackliga frågor. Det kan t ex gälla utredarens utbildning och kompetens och kanske i viss mån även erfarenhet, något som vi närmare kommenterar under rubriken ”Hur är det i Sverige idag?”

En närliggande aspekt till hur-frågan är vad utredningen består av. Testerna används i olika omfattning av olika utredare beroende på

grundutbildning, erfarenhet och kompetens samt hur bekant utredaren är med de olika testerna. Det finns ingen samlad bild av hur vanligt förekommande de olika testerna är. De tester som kräver utbildning används troligen i mindre omfattning än de generella och allmänna som kan köpas utan koppling till någon utbildning.

En viktig aspekt (även samhällsekonomiskt) är att säkra kompetensen hos lärare, speciallärare, specialpedagoger, psykologer och logopedier. De ska kunna ha evidensbaserade instrument för att göra goda pedagogiska utredningar samt kunskap att förstå när utredningen behöver involvera flera yrkeskategorier.

Grunden till en lyckad skolgång och möjligheten till en god utveckling är att kunna ta till sig den kunskap som ges. En förutsättning för det är att kunna läsa med flyt och förståelse eller att ha tillgång till alternativa verktyg för att ta till sig kunskapen. Oidentifierad dyslexi försvårar en optimal utveckling för den som är drabbad. Förskola och skola måste, så tidigt som möjligt, kunna identifiera och diagnostisera dyslexi samt tillhandahålla åtgärder för barn och ungdomar som löper risk för dyslexi eller som har dyslexi.

Hur är det i Sverige idag?

I Sverige klassades dyslexi som ett funktionshinder först år 1990. En statlig utredning konstaterade år 2010 att situationen för barn med läs- och skrivsvårigheter var diffus och att väntetiderna för dyslexiutredningar varierade i landet (SOU 2010:95:210). I enlighet med nya Skollagen ska elevhälsan involveras i de fall där skolans utredning behöver kompletteras eller göras mer djuplodande. Inom elevhälsan bör den sammantagna informationen runt en elev med t ex läs- och skrivsvårigheter finnas sammanfattad. Rätten att ställa en dyslexidiagnos är inte reglerad. Det lämnar utrymme för godtycke och ställer samtidigt krav på uppdragsgivaren att bedöma vem som ska ställa diagnosen. I praktiken är det vanligt att logopedier, psykologer, läkare, lärare, speciallärare, specialpedagoger och skolsköterskor är involverade i dyslexiutredningar. Enligt en undersökning av Lärarförbundet, görs dyslexiutredningar vid

ungefär hälften av landstingen och det sker i genomsnitt då eleverna är 13 år [115].

Professionerna utgår från olika definitioner på dyslexi vilket också försvårar en tolkning av testresultaten. Förutom kriterier i ICD-10 (International Statistical Classification of Diseases 10th version) [116] eller DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) [117] används den fonologiska förklaringsmodellen och definitionen i The International Dyslexia Association (IDA) som fastställdes år 2002. Dessutom skiftar underlagen för bedömningen av hur väl de definierade kriterierna uppfylls.

Enligt Skollagen har barn med behov av stöd för sin läsutveckling rätt till detta stöd utan någon formell diagnos. Av lagen framgår att rektorn ska se till att elevens behov av särskilt stöd skyndsamt utreds. Om utredningen visar att eleven är i behov av särskilt stöd, ska han eller hon ges det [118].

I dagsläget saknas enhetliga riktlinjer för utredning av dyslexi. Skolor saknar ibland en handlingsplan för utredning av läs- och skrivsvårigheter. De handlingsplaner som finns ser olika ut. För att kvalitetssäkra den pedagogiska utredningen utkom Specialpedagogiska skolmyndigheten år 2013 med ett stödmaterial [119].

Tabell 11.1 Lista på svenska tester.

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
Adler Färdighets- test i läsning [120]	F-vux	Individ/ grupp	Normerat	Ord-avkodning	-/-
Bedömning av språk B.A.S. [121]	F-klass Årskurs 1	Individ	Normerat	Fonologi, tal och språk	+/-

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
Bedömning av språklig medvetenhet [122]	5-åringar och uppåt	Individ	Jämförelsegrupper	Fonologi, tal och språk	-/-
Diagnostisering av elever med läs- och skrivsvårigheter [123]	Årskurs 4–6	Individ	Delvis normerat	Ordavkodning, läsförståelse, läshastighet, rättskrivning, hörförståelse, minne, tal och språk, skrivning, attityd	-/-
DJUR - lästest [124]	Årskurs 7–9, gymnasium	Grupp	Normerat	Ordavkodning, attityd, sammansatta prov	-/-
DLS bas för skolår 1 och 2 [125]	Årskurs 1–2	Grupp	Normerat	Läsförståelse, skrivning, attityd	+/+
DLS för skolår 2 och 3 [126]	Årskurs 2–3	Grupp	Normerat	Fonologi, läsförståelse, rättskrivning, ordförståelse	+/+
DLS för skolår 4–6 [127]	Årskurs 4–6	Grupp	Normerat	Läsförståelse läshastighet, rättskrivning ordförståelse	+/+
DLS för skolår 7–9 och 1 i gymnasiet [128]	Årskurs 7–9, gymnasium	Grupp	Normerat	Läsförståelse läshastighet, rättskrivning ordförståelse	+/+
DLSM – Diagnostiska läs- och skrivmaterial för mellanstadiet [129]	Årskurs 4–6	Individ	Normerat	Ordavkodning, läsförståelse, läshastighet, rättskrivning, tal och språk, skrivning, attityd, ordförståelse	+/+

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
DUVAN [130]	Gymnasium, vuxen	Grupp	Normerat	Ordavkodning, minne, attityd, ordförståelse	+/+
Materialet fonolek [131]	F-klass, årskurs 1	Grupp	Normerat	Fonologi	-/+
Lindahls högläsningsprov H4 [132]	Årskurs 1–3, 4–6, 7, 9	Individ	Normerat	Ordavkodning	-/-
Hur låter orden [133]	F-klass, årskurs 1	Grupp	Normerat	Fonologi	+/-
IL-basis, individuella prov [134]	Årskurs 1	Individ	Kvalitativt	Fonologi, ordavkodning, läsförståelse, rättstavning, hörförståelse, tal och språk, sammansatta prov	-/+
IL-basis, grupp-prov [135]	Årskurs 1	Grupp	Normerat	Fonologi, ordavkodning, rättstavning, hörförståelse, skrivning, attityd, sammansatta prov	-/+
Kvalitativa diagnoser [136]	Årskurs 4–vux	Individ	Kvalitativt	Fonologi, ordavkodning, rättstavning	-/-
Lilla DUVAN [137]	Årskurs 3, 5, 7	Grupp/ individ	Normerat	Fonologi, minne, attityd	+/+
LOGOS [138]	Årskurs 3, 4–6, 7–9, gymnasium	Individ	Normerat	Fonologi, ordavkodning, läsförståelse, läshastighet, hörförståelse, minne, sammansatta prov	+/+

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
LS klassdiagnoser [139]	Årskurs 7, 9, gymna- sium	Grupp	Normerat	Fonologi, ordavkodning, läsförståelse, rättstavning, ordförståelse	+/+
Läsa eller lyssna [140]	Årskurs 5	Individ	Normerat	Fonologi, läsförståelse, hörförståelse	-/-
LäsEttan [141]	Årskurs 1	Individ	Normerat	Fonologi, ordavkodning, RAN, självskattning	+/+
Läskedjor [142]	Årskurs 2, 3, 4–6, 7–9, gymna- sium, vuxen	Grupp	Normerat	Ordavkodning	+/+
Lässtandarder [143]	Årskurs 3, 4, 6, 7	Grupp	Kvalitativt	Läsförståelse, skrivning	-/-
LäSt [144]	Årskurs 1–3, 4, 5	Grupp	Normerat	Fonologi, ordavkodning, rättstavning	+/+
Madisons ordavkodning [145]	Årskurs 1–3, 4, 5, 7–9, gymna- sium, vuxen	Grupp	Normerat	Ordavkodning	-/-
Madisons ordförståelse [146]	Årskurs 5, 6, 7–9, gymna- sium	Grupp	Normerat	Ordförståelse	-/-
Madisons rätt- skrivningsprov alla stadier [147]	Årskurs 2, 3, 4–6, 7–9; gymna- sium, vuxen	Grupp	Normerat	Rättstavning	-/-

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
Madisons tystläsning med förståelsekontroll [148]	Årskurs 2, 3, 4–6, 7–9, gymna- sium, vux	Grupp	Normerat	Läsförståelse, läshastighet	–/–
Metod för läs- och skrivinläring [149]	Nyinlä- ring årskurs 1–3, om- inläring årskurs 4–9	Individ	Kvalitativt	Analyssamtal, självskattning	–/–
MG-kedjor [150]	Årskurs 6, 7–9, gymna- sium, vuxen	Grupp	Normerat	Ordavkodning	+/+
Materialet Morfolek [151]	Årskurs 4–6, 7, 8	Grupp	Normerat	Ordavkodning	–/–
MiniDUVAN [152]	Förskole- barn	Individ	Normerat	Fonologisk medvetenhet	+/+
Observation av språk, O.A.S [153]	F-klass	Individ	Normerat	Fonologi, hörförståelse, minne, tal och språk, ordförståelse	–/–
OLAF, individualprov [154]	Årskurs 1–3	Individ	Normerat	Ordavkodning	–/–
Ordavkodning [155]	Årskurs 4–6, 7–9	Grupp	Normerat	Ordavkodning	–/+
Ord & bild, Bild & ord [156]	Årskurs 1–2	Grupp	Normerat	Ordavkodning	–/–
Ordlek [157]	F-klass	Grupp		Fonologi	–/–

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
Ordracet, individualprov [158]	Årskurs 1–3	Individ	Normerat	RAN	–/–
OS64 och OS120 [159]	Årskurs 2, 3	Grupp	Normerat	Ordavkodning	–/–
Provia, datoriserat individualprov, 15 språkliga områden [160]	Årskurs 1–3, 4–6, 7–9	Individ	Normerat	Fonologi, ordavkodning, läsförståelse, rättstavning, hörförståelse, minne	–/–
Screening, datorprogram [161]	Gymna- sium årskurs 1	Individ	Normerat	Läsförståelse, läshastighet, rättskrivning, skrivning, ordförståelse	–/–
SL40 och SL60 [162]	Årskurs 3–5	Grupp	Normerat	Läsförståelse	–/–
Språkresan – en väg till ett andra språk [163]	F-klass från 2,5 år	Individ	Kvalitativt	Språklig nivå ur ett flerspråkigt perspektiv	–/–
Stavning L, M, H [164]	Årskurs 2, 3, 4–6, 7–9	Grupp	Normerat	Rättstavning	–/+
Stavux [165]	Vuxna	Individ	Normerat	Fonologi, rättstavning	–/–
SVAPP, ordförrådsprov för sexåringar [166]	F-klass	Grupp	Anger gränsvärde	Vokabulär (passivt ordförråd)	–/–
Testa språket med Trulle [167]	F-klass, årskurs 1	Individ	Kvalitativt	Fonologi	–/–
Testbatteriet individualprov [168]	Årskurs 3, 4–6, 7–9	Individ	Normerat	Fonologi, ordavkodning, minne, sammansatta prov	–/–

Tabellen fortsätter på nästa sida

Tabell 11.1 fortsättning

Namn Referens	Årskurs	Individ/ grupp	Normering/ kvalitativt	Testvariabel (deltest)	Reliabilitet/ validitet
UMESOL [169]	Årskurs 1–3	Individ	Resultattabell	Fonologi, ordavkodning, läsförståelse, rättstavning, självskattning	+/-
Vilken bild är rätt [170]	Årskurs 2, 3	Grupp	Normerat	Läsförståelse	-/-
Ängelholms-provet Ängelholms kommun	Årskurs 4–6	Grupp/ individ	Normerat	Fonologi, läsförståelse	-/-

Följande tester listas inte i Tabell 11.1:

Tester som är rena utvecklingsscheman: 27 glada kliv, Att lära barn läsa, God läsutveckling, God skrivutveckling, LUS, Nya Språket lyfter, Språket på väg.

Tester som inte går att få tag på idag: Bangortestet, Dyslexia screening test, Lyssna på ordet, OS400.

Tester som inte mäter de utfallsmått som bestämdes av projektets experter: ITPA, TROG, TRAS, CELF-4.

12. Diskussion

I det här kapitlet kommer först en kort sammanfattning av de viktigaste resultaten som har satts i en kontext. Därefter följer en diskussion om litteratursökningen i förhållande till den skiftande terminologin samt ett avsnitt om metodfrågor och de överväganden som gjorts under projektets gång. Till sist följer en fördjupad diskussion om resultatredovisningen för rapportens tre huvudfrågor.

Sammanfattning av resultaten

Det finns vetenskapligt underlag för att strukturerad träning av kopplingen mellan språkljud och bokstäver förbättrar läsförmåga, stavning, läsförståelse, läshastighet och förmåga att uppmärksamma språkets ljudmässiga uppbyggnad (fonologisk medvetenhet) för barn med dyslexi. För övriga insatser var det vetenskapliga underlaget otillräckligt. Det här forskningsområdet är i behov av forskning av hög metodologisk kvalitet (randomiserade kontrollerade studier). Det är viktigt att framtida studier undersöker en tillräckligt stor grupp så att slumpen inte spelar in, även om studien är randomiserad. De utfallsmått som förekom i de studier vi hittade var de som var enkla att mäta, så som t ex stavning och läsförmåga. Effekten av ökad livskvalitet, såsom självbild, självkompetens och självförtroende utvärderades inte i några av de inkluderade studierna.

Det finns tester som förefaller kunna förutsäga dyslexi redan innan barnen fått undervisning i att läsa och skriva i skolan. Brister i fonologisk medvetenhet, snabb automatiserad benämningsförmåga och bokstavskännedom har ett samband med dyslexi.

För de svenska testerna som vi utvärderade var det vetenskapliga underlaget otillräckligt. Dock fanns det enstaka internationella deltester som var utvärderade. Det behövs således mer forskning som utvärderar testernas tillförlitlighet. Traditionen att kontinuerligt utvärdera ett tests tillförlitlighet finns inte inom detta forskningsområde.

Litteratursökningen i förhållande till den skiftande terminologin

För att ta fram en systematisk översikt ingår litteratursökningen som en grundläggande del. I arbetet med att utveckla lämpliga sökstrategier för att identifiera så många relevanta studier som möjligt hittade vi några försvårande omständigheter. Det kunde t ex vara avsaknad av en heterogen terminologi eller avsaknad av viktig information i titel och artikelsammanfattning.

Begreppsbildningen är i ett internationellt perspektiv heterogen, se Faktaruta 12.1. Den heterogena terminologin kan ha flera orsaker, t ex att det inte finns någon internationellt enhetlig definition. Litteratursökningen ska vara bred för att kunna fånga så många relevanta studier som möjligt. Samtidigt måste den balanseras mot en rimlig omfattning av identifierade studier. Artikelsammanfattningarna ska granskas och ett alltför omfattande sökresultat blir svårt att hantera. En oklar och heterogen användning av begreppen kan resultera i ett mycket omfattande sökresultat eftersom alla begrepp måste vara inkluderade för att inte riskera att missa viktiga studier. För att inte utvidga sökstrategierna kan man istället komplettera databassökningen med t ex kontroll av referenslistor i centrala studier.

En svårighet kan vara att det saknas viktig information i studiernas titlar och artikelsammanfattningar. Den här informationen kan ha central betydelse för att identifiera nya studier som ännu inte har hunnit få de särskilda ämnesord som läggs på varje studie. Dessa kontrollerade ämnesord finns i databasspecifika ordlistor (tesaurus) som en hjälpfunktion i de stora internationella databaserna och läggs på varje studie av indexerare. Syftet är att underlätta utformandet av sökstrategier och skapa en större homogenitet i litteratursökningen. Avsaknad av information i titlar och artikelsammanfattningar, tillsammans med den heterogena användningen av centrala begrepp, kan även medföra heterogenitet i indexeringen eller att viktiga ämnesord helt saknas. Det innebär en risk för att studier missas, trots omfattande sökresultat.

Faktaruta 12.1 Exempel på skiftande begreppsbildning.

I USA används oftast termen **learning disabilities**. Ibland används också ett specificerande tillägg i texten såsom **reading disabilities** eller **dyslexia**. Learning disabilities är en paraplyterm för hinder i inläringen som orsakas av okända faktorer vilka inte kan knytas till någon specifik medicinsk diagnos. Typiska sådana hinder kan vara dyslexi, dyskalkyli och dysgrafi som alla betecknar svårigheter av typiskt utbildningsrelaterad karaktär. **Learning disorders** är den term som är reserverad för företeelser som har en specifik medicinsk diagnos. I realiteten tycks termerna learning disabilities, learning disorders och learning difficulties användas omväxlande.

I Storbritannien motsvaras **learning disabilities** av termerna **specific learning difficulty**, **developmental dyslexia**, **dyscalculia** och **developmental coordination disorder**. Där används termen **learning disability** istället för olika typer av mer uttalade, allmänna intellektuella nedsättningar.

Utöver detta finns i den engelskspråkiga litteraturen en rad beteckningar av mer eller mindre övergripande karaktär. **Dyslexia** används ofta i tidskrifter som specialiserat sig inom området eller har en neurologisk eller neuropsykologisk inriktning och har då en välavgränsad definition, normalt i termer av prestation i förhållande till normalfördelningskurvan. Avgränsningen kan vara olika snäv och anges antingen procentuellt, t ex 25:e percentilen, dvs de 25 procent av en population som ligger lägst i sin läs- och stavningsförmåga eller i form av standardavvikelse, t ex minus en standardavvikelse, vilket motsvarar cirka 16:e percentilen.

Specific reading disability är en beteckning som används synonymt med **dyslexia**. Andra termer av hybridkaraktär som specifikt inriktar sig på individer med dyslexi är **specific reading disorders** och **specific reading difficulties**. Mycket vanligt är beteckningar som kan rymma en mindre väl avgränsad population. Hit hör t ex **reading difficulties**, **poor readers**, **struggling readers** och **resistant readers**, alla vanligt förekommande i pedagogiskt inriktade tidskrifter. Det är ändå inte helt ovanligt att en artikel med beteckningen till exempel **struggling readers** i rubrik och/eller abstrakt i huvudtexten visar sig vara inriktad på elever med dyslektiska svårigheter.

Metodfrågor

Förutsättningar, arbetssätt och vägval inom projektet

I arbetet med att ta fram en systematisk översikt är det viktigt att klargöra vilka förutsättningar och avgränsningar som gäller. Exempel på avgränsningar i vårt projekt är att vi inte inkluderade studier om eventuella orsaker till dyslexi. Vi inkluderade bara studier som tog upp tester, som på basis av tidigare språkfärdigheter försökte förutsäga dyslexi hos barn. Vi inkluderade även ett antal tester som användes för att upptäcka (screena) och utreda dyslexi hos barn och ungdomar, samt studier som var inriktade på insatser och alternativa verktyg för barn och ungdomar med dyslexi (yngre än 20 år), vilket innebär att studier som rörde vuxna inte inkluderades.

Ett annat exempel på en avgränsning var studierna om de internationella testerna, där 18 tester valdes ut med hjälp internationella forskare (Tabell 6.1 i Kapitel 6). Projektgruppen beslöt att exkludera studier där indextestet (det studerade testet) jämfördes med ett referentstest som inte ingick bland de i förväg utvalda testerna. Detta innebär att endast studier med för projektet inkluderade tester jämfördes sinsemellan.

De svenska testerna som idag används för att identifiera och utreda barn och ungdomar med dyslexi kartlade vi genom en inventering. Den bestod i att vi tog kontakt med Specialpedagogiska skolmyndigheten och några lärosäten där specialpedagoger och speciallärare utbildas. Ett antal svenska forskare bidrog också med information om svenska tillgängliga tester. Vi studerade inte tester som avsåg att följa förändring över tid.

Att förutsäga dyslexi

För studier om tester för att förutsäga dyslexi, accepterade vi enbart longitudinella. Det innebär att studierna är framåtriktade och testerna ges till barnen innan barnen lärt sig läsa och skriva. De följs sedan upp efter det att barnen fått möjlighet att lära sig läsa och skriva under minst ett års tid. De studier, där barnen redan kunde läsa när prediktionstestet gavs, exkluderades. Likaså de studier som inte följde barnen tillräckligt länge för att de skulle ges en rimlig chans att lära sig läsa. För dessa stu-

dier valde vi att sätta ett gränsvärde för dyslexi. Om eleverna i uppföljningsstudien hade en förväntad läsförmåga som låg en standardavvikelse under medelvärdet för den aktuella åldern, ansågs dyslexi föreligga.

Svenska och internationella tester

Det finns många svenska och internationella tester som inte har utvärderats i rapporten. Det beror på att de inte uppfyllde de i förväg uppsatta inklusionskriterierna. Ett inklusionskriterium var att testet skulle utvärderas på personer i åldersspannet 6–20 år. Om en studie utvärderade testet på en population som inte täckte detta åldersspann exkluderades den. Ett exempel är det svenska testet DUVAN som utvärderades på vuxna deltagare [171,172]. Detta test, liksom andra tester såsom Läs-kedjor, LäSt och DLS, har även ingått som mått i vetenskapliga studier [125,142,144].

Interventionsstudierna

Vi inkluderade kontrollerade studier, med eller utan randomisering. Då flertalet studier inte redovisade vilken definition de utgått ifrån för att beskriva gruppen med dyslexi, var vi tvungna att sätta ett tröskelvärde, en så kallad cut-off-gräns, för att urskilja vilka studier som handlade om dyslexi. Personer med dyslexi är inte en homogen grupp utan omfattar allt från de med stora svårigheter till relativt små. Det gjorde det svårt att sätta en gräns. Experterna i projektet enades dock om att interventionsgruppen (pre-test) och kontrollgruppen skulle ha en genomsnittlig läsförmåga som avvek mer än 1,25 SD (eller 10:e percentilen) i förhållande till en jämförelsegrupp eller en norm. Med den gränsen omfattas 10 procent av befolkningen. Man kan dock förmoda att en del elever i denna grupp är uteslutna ur studierna t ex på grund av att majoritetsspråket inte är deras modersmål eller att de har en annan diagnos än dyslexi. Med läsförmåga avses här ordavkodning, non-ordläsning och/eller läshastighet av enstaka ord/non-ord alternativt av sammanhängande text samt stavning. Vi vill således, i enlighet med den antagna definitionen, inkludera elever med manifesta ordavkodningssvårigheter. Optimalt redovisas i studierna även fonologiska problem, vilket skulle förstärka antagandet att det rör sig om en population med dyslexi. Ytter-

ligare ett problem är att dyslexi i de flesta studier enbart bekräftas med hjälp av tester, fastän testernas tillförlitlighet inte är säkerställd eller att testerna är tillräckligt utvärderade. En väl definierad population är en förutsättning för att kunna uttala sig om en effekt för en viss insats. Då gränsen för dyslexi generöst sattes till den 10:e percentilen, var det svårt att få en tydligare och mer differentierad effekt än om gränsen hade satts vid en lägre percentil.

Interventionsstudier som undersökte färre än 30 barn med dyslexi exkluderades, vilket gjorde att vi exkluderade en stor mängd studier. Bland annat exkluderades studier som var baserade på endast en person, en så kallad case study. Dessa studier saknar experimentell kontroll och metoderna är oftast beskrivna på ett sådant sätt att de är svåra att upprepa, samt att de metoder som används för att kvantifiera oftast bygger på subjektiva mått såsom självrapportering. Andra exempel på exkluderade studier var de som undersökte effekten av individträning jämfört med gruppträning. Underlaget av studier som visade tillräcklig kvalitet var tunnt och därför kunde vi inte genomföra några subanalyser för att t ex ta reda på om det fanns några skillnader mellan kön, skolår, grad av dyslexi eller jämförelser av individ- respektive gruppträning.

Övriga avgränsningar

Vi valde att avgränsa litteratursökningen för teststudier till år 1990, för interventionsstudier till år 1980 och för studier om alternativa verktyg till år 2000. Forskningen om dyslexi har förändrats mycket de senaste 30 åren och det är därför inte relevant att basera resultaten på äldre studier eftersom de då inte blir överförbara till dagens situation.

Forskning från hela världen inkluderades med språk som bygger på den alfabetiska principen. Rapporten baseras dock i huvudsak på studier från USA och Europa. Det speglar sannolikt den forskning inom utbildningsområdet som sker internationellt samt att mycket av den högkvalitativa forskningen sker i USA och i Europa.

I kvalitetsgranskningen har vi beaktat många faktorer som kan påverka och snedvrider studiens resultat, så kallad bias. De risker för bias, som

bedömts för varje kvalitetsgranskad studie av hög eller medelhög kvalitet, är följande: selektionsbias (brister i randomiseringsmetod), behandlingsbias (otillräcklig blindning), samt de mer självklara riskerna: bedömningsbias, bortfallsbias, rapporteringsbias samt intressekonflikter. Risken för bias ökar om forskarna har en tendens att överrapportera positiva fynd, samtidigt som de underlåter att rapportera utebliven effekt. Att intressekonflikter redovisas i studierna förekommer alltför sällan. Det är en särskilt viktig fråga när det gäller teststudier, då det kan handla om att forskaren vill framhäva sitt eget test.

Projektet har arbetat enligt den strikta metodik som tillämpas inom SBU. Arbetssättet liknar det tillvägagångssätt som används inom Cochrane Collaboration. Mycket strikta kriterier gjorde att många studier inte uppfyllde våra inklusionskriterier och att de inte höll för relevans- och kvalitetskriterierna.

Det var en utmaning att bedöma test- och insatsstudiernas kvalitet på ett systematisk och transparent sätt. Alla SBU-rapporter använder GRADE för att evidensgradera resultaten. GRADE har utarbetats internationellt och används bl a av Världshälsoorganisationen (WHO) och av Cochrane Collaboration. GRADE används, förutom av SBU och Socialstyrelsen, även av andra aktörer som gör kunskapsöversikter i Sverige. Fördelarna med att GRADE används av allt fler är att möjligheterna ökar för att kunna jämföra kunskapssammanställningarna.

GRADE tillämpas för att bedöma tillförlitligheten hos de sammanvägda resultaten. Utgångspunkten är att randomiserade studier har hög evidensstyrka medan observationsstudier har lägre. Från det utgångsläget kan evidensstyrkan sänkas eller höjas. Nedgradering sker om det finns brister i överförbarhet mellan olika studier och evidensgraden kan höjas för observationsstudier om stora effekter ses och det finns få förväxlingsfaktorer, så kallade confounders.

Fördjupad resultatdiskussion

Att förutsäga dyslexi

Det är eftersträvansvärt att kunna förutsäga vilka barn som har sådana svårigheter att de löper extra stor risk att få problem som är karakteristiska för dyslexi. De kan då få en anpassad träning redan från början. Tester som kan förutsäga dyslexi måste bygga på förmågor som har betydelse för läsning, och ska kunna användas redan innan barnen är läskunniga. I själva begreppet förutsäga ligger att testets tillförlitlighet bäst mäts genom att undersöka samstämmigheten mellan testresultatet och förekomsten av dyslexi vid ett senare tillfälle.

Det finns en omfattande forskning som har studerat hur olika kognitiva och språkliga faktorer hos barn korrelerar med senare läsutveckling. Majoriteten av undersökningarna har inte syftat till att förutsäga dyslexi specifikt utan snarare studerat sambandet mellan tidig förmåga och senare läsförmåga generellt hos en population. Det är möjligt att samma eller liknande förmågor som predicerar dyslexi (dvs den andel av en population som har stora svårigheter att lära sig läsa) också kan förutsäga läsutveckling hos en oselektad population. Det går dock inte att använda resultat från undersökningar som studerat tester för att förutsäga generell läsutveckling som vetenskapligt belägg för att de också predicerar dyslexi. Av den anledningen begränsades urvalet av studier till de som primärt syftade till att förutsäga dyslexi.

Vi ställde upp ett antal kriterier för inklusion av studierna men få uppfyllde kriterierna. Många var tvärsnittsstudier som jämförde olika prediktorer med varandra utan att kontrollera om barnen vid ett senare tillfälle visade sig ha dyslexi. Vi sökte longitudinella studier. Tyvärr var studierna behäftade med osäkerhet pga bristande redovisning. Ett problem var på vilka grunder barnet bedömdes ha dyslexi. I flera studier var det tveksamt om kartläggningen av barnets problem var tillräckligt noggrann för att de bristande förmågorna skulle klassificeras som dyslexi. I studierna identifierades oftast läs- och skrivproblem med ett litet antal tester. På samma sätt var det i praktiken svårt att fastställa om barnet verkligen inte hade fått någon träning i läsning. De flesta studierna var genomförda i anglosaxiska länder och åldern för när barnen får börja

bekanta sig med bokstäver och läsning varierade. En begränsning i generaliserbarheten av våra resultat var att det ofta förekom översampling av vissa grupper. Ett exempel var två studier med samkönade en- och tvååggstvillingar [19,28] och två andra studier som inkluderade barn där föräldrarna hade dyslexi [29,33]. I en studie slog man samman flera prediktorer (fonologisk medvetenhet, automatiserad benämningsförmåga, bokstavskännedom osv) till ett kompositmått och undersökte hur väl de sammanslagna prediktorerna kunde förutsäga dyslexi [27]. Effekten av varje enskild prediktors förmåga att förutsäga dyslexi kan då bli både under- och överskattad. Därför bör graden av samband för enskilda prediktorer tolkas med försiktighet.

Tester

Tester ska i så hög utsträckning som möjligt vara tillförlitliga (reliabla), trovärdiga och sanna (valida), standardiserade och fria från bias. Testresultat kan ge viktig information vid ställande av dyslexidiagnos, men en dyslexiutredning är mer komplex än så och bör inte grunda sig på enstaka testresultat. De flesta tester består av ett antal deltester. Om enbart vissa deltester ingår i utvärderingsarbetet är det vanskligt att dra slutsatser om hela testbatteriet.

Det finns ett begränsat antal studier som undersöker psykometrisk prestanda hos de tester som ofta används i läs- och dyslexiforskning och i kliniska sammanhang för att diagnostisera dyslexi [173]. Denna slutsats gäller både de svenska och internationella testerna och testbatterierna vi har granskat. Inga studier som undersökte de svenska testernas prestanda inkluderades i rapporten då det saknades relevanta studier. En bedömning av evidensstyrkan är således inte möjlig att göra för de svenska testerna då informationen är otillräcklig.

I regel ingår ett relativt litet antal personer i de psykometriska utvärderingarna, varför försiktighet måste tillämpas vid tolkning av resultaten. Särskilt råder det brist på studier rörande testernas förmåga att identifiera dyslektiska svårigheter (snarare än att testa hela variationen i läsförmåga). Man kan tänka sig att det beror på att det saknas konsensus för såväl diagnostik (och definition) av dyslexi, som att det saknas en gold

standard för testinstrument. Mot denna bakgrund är vår generella slutsats att man i själva verket vet relativt lite om psykometriska prestanda för de internationella testerna. Det gäller även tester som är väletablerade i USA, såsom DIBELS och möjligen med något undantag, Woodcock-Johnsons testbatteri. I endast en studie undersökte man huruvida ett screeningtest (DIBELS) kunde användas för att fånga in en grupp barn som presterade svagt enligt ett mer omfattande testbatteri av fonologisk förmåga (CTOPP) [44]. Resultatet visade att screeningtestet hade hög sensitivitet, men låg specificitet. Detta är en vanlig egenskap hos screeninginstrument och betyder att man med instrumentet fångar in de barn som faktiskt har problem med fonologisk förmåga, men att man också fångar in många barn som inte har några uppenbara svårigheter (falska positiva). I regel får man väga specificitet och sensitivitet mot varandra, och vid screening anses sensitiviteten vara den viktigaste parametern av de två, se Bilaga 5 Bakgrund och överväganden angående statistik.

Man kan också konstatera att forskning inom detta område ofta bedrivs utan att man redogör för psykometriska egenskaper hos de tester som beskriver den grupp som ingår i studien. Istället refererar man i regel till testmanualerna där det ibland redogörs för psykometriska egenskaper från standardiseringsarbetet. Det finns vissa problem med detta tillvägagångssätt. För det första har arbetet med manualerna i regel inte fackgranskats, till skillnad från studier som publiceras i internationella tidskrifter. För det andra är diagnostisk tillförlitlighet inte en stabil egenskap hos ett test i alla sammanhang, utan det kan variera stort för olika populationer. Man kan därför säga att tillförlitligheten hos testet avser de resultat som erhållits vid ett tillfälle med testet och inte själva testet i sig. Därför måste prestandan i testet kontinuerligt utvärderas [174]. I teststudierna som granskades saknades ofta information om resultatet från referenstestet hade tolkats oberoende av resultatet från det undersökta testet. Det saknades också i många studier information om de som utför testerna hade adekvat och tillräcklig kompetens. Utbildning krävs för den som ska administrera testet och tolka resultatet. Det finns också många starka kommersiella intressen involverade i testerna. Screening- och utredningstester är kostsamma, vilket innebär att det finns risk för att nytugåvor med uppdateringar inte köps in utan att de äldre versionerna används med risk för att ge felaktiga resultat.

Trots att ett antal problem har lyfts fram med de olika teststudierna har vår omfattande granskning resulterat i att vi har kunnat utvärdera ett antal deltester. Även om resultaten ska tolkas med försiktighet pga det begränsade underlaget så framträder en generell bild av att det går att utveckla reliabla och valida utredningsmaterial för att undersöka läs- och skrivfärdighet. Framför allt kan vi uttala oss om Woodcock-Johnsons deltest Letter-Word Identification (eller Word Identification, som den heter i en tidigare upplaga av testet). När det gäller detta deltest så är konklusionen att det finns begränsat vetenskapligt underlag för tillfredsställande reliabilitet och validitet. Även om Letter-Word Identification endast är ett deltest i ett större testbatteri så mäter testet potentiellt fler aspekter av barnets färdighet i användandet av den alfabetiska principen vid läsning, såsom säker bokstavskänedom och avkodningsförmåga av vanliga och ovanliga ord som presenteras i isolering. Att orden presenteras utan betydelsebärande sammanhang är en viktig aspekt eftersom avkodning och läsning är särskilt svåra för personer med dyslexi, potentiellt pga underliggande fonologiska svårigheter [1].

Insatser

Det finns en omfattande litteratur som beskriver ett antal träningsmetoder för barn och ungdomar med dyslexi. Ändå identifierade vi endast en metod, strukturerad träning av fonem–grafem-koppling. Metoden har ett vetenskapligt underlag för att den förbättrar barnens förmåga att läsa ord och non-ord korrekt samt förbättrar stavning, läsförståelse, läshastighet och fonologisk medvetenhet. Man skulle möjligen kunna komma att ifrågasätta denna slutsats i rapporten, med hänvisning till att effekten av träningen i strukturerad fonem–grafem-koppling i vissa fall har utvärderats med tester vars vetenskapliga underlag bedömts som otillräckligt i vår rapport. Att det inte anses finnas vetenskapligt underlag innebär i det här sammanhanget att testet inte har varit utvärderat i någon studie som har publicerats i en vetenskaplig artikel. Det behöver alltså inte betyda att testerna inte är reliabla och valida. Dessutom jämförs alltid testresultaten mellan grupper som har, eller inte har, erhållit den aktuella interventionen.

Endast 33 studier om insatser vid dyslexi bedömdes som tillräckligt tillförlitliga för att ingå i det vetenskapliga underlaget. Tillförlitligheten i data från en randomiserad kontrollerad studie påverkas framför allt av hur randomiseringen har fungerat, om barn och utförare är blinda för vilken insats som ges (aktiv eller kontroll) och bortfall.

Blindning förekommer normalt inte för studier inom detta område eftersom det kan vara svårt att åstadkomma. En möjlighet, som sällan utnyttjats i studierna, är att den som utvärderar effekten av träningen inte känner till vilken grupp barnen tillhör. Å andra sidan berättar barn gärna vilken insats de får, så även blindning av utvärderare kan bli problematisk. Bortfallet och bortfallsanalysen var bristfälligt beskrivna i vissa studier.

Ett annat problem är att de flesta studierna var små, eller mycket små. Förvånansvärt många utvärderingar av träningsmetoder omfattade en handfull barn och dessa kom över huvud taget inte ifråga för vår granskning. Risken med små studier, med enkel randomisering, är att slumpen ändå kommer att påverka fördelningen mellan grupperna. Vi tillämpade en generös regel och inkluderade studier med minst 30 deltagare totalt för kontroll- och interventionsgrupp. Det bidrar dock till ytterligare osäkerhet i underlaget.

En osäkerhet som är svår att bedöma är hur väl träningsprogrammen följdes. Många studier redovisade inte om den så kallade programtroheten (fidelity) kontrollerades, eller i vilken utsträckning barnen faktiskt deltog i träningen. Skillnader i effekter mellan studier skulle i vissa fall kunna bero på vem det är som förmedlar träningen. Är det den entusiastiske forskaren som utvecklat träningen ifråga eller är det den ordinarie personalen på skolan?

Tolkningen av resultaten försvårades av de olika valen av kontrollåtgärd. Studierna jämförde antingen två träningsmetoder med varandra eller en experimentell metod mot ordinarie undervisning (curriculum as usual, training as usual). Nackdelen med ordinarie undervisning är att den inte beskrivs och att innehållet i den därmed är okänt. Sannolikt varierar

den ordinarie undervisningen mellan studier och risken är att man i en metaanalys i viss mån jämför äpplen med päron.

Ytterligare en orsak till det magra utfallet var att studierna mätte effekter av träning för olika läsmått. Endast två studier undersökte t ex effekterna av att träna läsflyt för korrekt läsning av ord och endast två studier undersökte för läsförståelse. Om det hade funnits en större enighet i vilka mått som använts hade vi kunnat tillvarata mer av utvärderingarna.

De studier som jämförde två förmodat effektiva metoder med varandra beskrev oftast interventionerna i acceptabel grad. Däremot var kontrollmetoderna så olika i studierna att det inte gick att väga samman resultaten i en metaanalys. Det är fullt möjligt att det finns effektiva metoder bland dem som utvärderas på detta sätt, men vi har inte kunnat bedöma dessa. Det kan vara så att två träningsprogram som ställs mot varandra till stor del innehåller samma komponenter och därför inte resulterar i några större skillnader mellan grupperna. Det kan tolkas som att det är principerna för träningsinsatsen snarare än enskilda program som är effektfulla.

Vi kan notera att effektstorleken tenderar att vara större i de nordamerikanska studierna jämfört med andra länder. En möjlig förklaring kan vara att deltagarna i dessa länder företrädesvis kommer från skolor och områden med låg socioekonomisk status. Naturligtvis är dyslexi lika vanligt förekommande i alla nivåer av socioekonomisk status, men då vi inte känner till deltagarnas fonologiska förmåga skulle det kunna innebära att den låga ordavkodningen är en reflektion av miljön snarare än dyslexi. Dessutom är det troligt att deltagarna, med mer transparenta språk än engelska, redan innan interventionen fick undervisning baserad på en mer eller mindre strukturerad fonem–grafem-koppling. Man skulle alltså kunna anta att de därför utgjorde mer av en ”hård kärna” och alltså inte var lika mottagliga för interventionen.

Alternativa verktyg

Alternativa verktyg är ett begrepp som har använts de senaste åren för att beskriva olika hjälpmedel för att stödja och utveckla läsförmågan för personer med läs- och skrivsvårigheter. Främsta avsikten med dessa verktyg är att ”gå runt” läs- och skrivproblemen och göra det möjligt att ta till sig och förmedla text utan att läsa och skriva på det traditionella sättet. Många studier har gjorts på datorområdet och olika program för att genomföra interventioner som ska förbättra läs- och skriveförmågan. Det är dock få studier som belyst just alternativa verktyg där huvudsyftet inte har varit att träna den aktuella förmågan utan att verkligen kompensera för läs- och skrivsvårigheterna. Det är fortfarande oklart vad som egentligen är kompenserande program, dvs vilka program som ”går runt” läs- och skrivproblemen och vilka program som tränar förmågan. Det är t ex vanligt med studier som utforskat hur en talsyntes kan göra att läsförmågan utvecklas. Oftast har man då försökt mäta förbättringar med olika typer av lästester. Resultaten har varit skiftande, från goda resultat till ingen förbättring alls. Ett problem är att studierna försökt mäta t ex avkodning, vilket är en förmåga som inte direkt tränas av en talsyntes.

Vår avsikt har varit att försöka hitta studier som använt alternativa verktyg för att studera om deltagarna blivit bättre på att förmedla och ta till sig text efter att ha använt verktygen. Vi har inte funnit någon studie som uppfyllt de uppställda metodologiska kraven. Studierna har oftast haft för små undersökningsgrupper, inga kontrollgrupper, svårtolkade resultat, för lite information om undersökningsgruppen (som t ex vilken typ av läs- och skrivsvårigheter eleverna hade och hur grava problemen var). När det gäller smarta telefoner, surf- och läsplattor och tillhörande applikationer (så kallade appar) har vi funnit ytterst få studier. Inga har varit metodologiskt acceptabla. Detta förklaras till stora delar av att tekniken är förhållandevis ny och få forskare har hunnit utvärdera hjälpmedlen. Området är i behov av mer omfattande forskning med större antal deltagare i undersökningarna. Dessutom behöver forskningen problematisera skillnader mellan tränande program och utrustning och alternativa verktyg, dvs på vilket sätt kan alternativa verktyg kompensera för en persons läs- och skrivsvårigheter? Genom att ”gå runt svårigheterna”, träna upp förmågan eller både och [175]?

Jämförelser med resultat från andra översikter

Intervention for Dyslexia

Singleton publicerade en icke-systematisk översikt med titeln *Intervention for Dyslexia* år 2009 [176]. Denna översikt var inte systematisk. Den redovisade inte hur studierna hade inkluderats eller sökts fram, inte heller hur de hade granskats. Syftet med översikten var att sammanfatta publicerad forskning om effekten av specialundervisning för barn från 5 till 18 år med dyslexi. Den tog även upp screeningtester samt alternativa verktyg.

Översikten beskrev tydligt hur studierna definierade begreppet dyslexi. Att dyslexi främst påverkade förmågan att få flyt i läsningen, att stavningen påverkades och att dyslexi ofta ledde till svårigheter med bl a fonologisk medvetenhet och verbalt minne. Här beskrevs att dyslexi förekom på alla begåvningsnivåer och att det var bäst att se dyslexi som ett kontinuum, utan några tydliga gränser.

Singleton kommenterade att det fanns få randomiserade studier på området. Studierna var i huvudsak kontrollerade men utan randomisering, något som Singleton kallade silver standard. Singleton valde att redovisa forskningsstudierna för interventionerna i tre olika kapitel (England, USA och andra länder) pga skiftande skolsystem samt behov av specialundervisning. Eftersom det inte är en systematisk översikt finns heller inte någon redovisning över hur litteraturen söktes fram. Singleton inkluderade också mycket grå litteratur, som t ex kapitel ur böcker och avhandlingar.

Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties

Sir Jim Rose fick i uppdrag av regeringen i Storbritannien att skriva en rapport om tidig riskbedömning och insatser för barn med dyslexi [177]. Rapporten publicerades år 2009 och fick ett stort genomslag i medierna. Dåvarande utbildningsminister Ed Balls lovade att implementera rekommendationerna i rapporten.

Rose-rapporten är inte en systematisk översikt. Den redovisar inte hur studierna hade inkluderats eller sökts fram och inte heller hur de hade granskats. Man byggde resultatet på rapporter från konsultationer, besök i skolor samt kunskap från 659 föräldrar till barn med dyslexi, 75 barn och vuxna med dyslexi samt 129 lärare och forskare.

Rapporten definierade tydligt dyslexi, och definitionen accepterades av Storbritanniens nationella dyslexiorganisation. Rapporten innehåller en mängd rekommendationer. Här betonas bl a vikten av lärarnas roll och kompetensutveckling av lärare i arbetet med att upptäcka elever med dyslexi. Riskbedömning av alla barn med hjälp av så kallade screeningtester anses inte lämpligt, då dessa tester ännu inte är tillräckligt utvärderade. Däremot hänvisar rapporten till ett antal publicerade studier som förespråkar riskbedömning med hjälp av prediktorerna bokstavs-kännetecken samt fonologisk medvetenhet [176,177]. Rapporten framhåller vikten av tidig systematisk träning av fonem–grafem-koppling och noterar vidare att det saknas evidens för effektiva insatser för äldre barn.

13. Konsekvenser av rapportens bedömningar

SBU:s roll är att systematisera aktuell och kvalitetssäkrad kunskap. Det är sedan andra aktörer som kommer att använda denna kunskap som vägledning. Den här rapporten innehåller inte några förslag till förändringar i praxis.

Enhetlig utredning

Idag saknas tydliga riktlinjer för hur barn och ungdomar ska utredas vid dyslexi. Kunskapssammanställningen visar på att det förekommer över 50 olika kartläggnings- och utredningstester runt om i landet. Vissa tester har en teoretisk förankring men är inte vetenskapligt utvärderade. Det finns därför behov av att utvärdera testerna och prioritera mellan dem. På så sätt kan man säkerställa att utredningar görs med relevanta tester med god tillförlitlighet samt att de mäter det som de avser att mäta. Det skulle leda till ett mer enhetligt utredningsförfarande och skulle vara kostnadsbesparande. Utredningarna skulle ske på samma sätt oberoende av skola, kommun eller landsting och det skulle vara lättare att upprätthålla en hög och jämn kompetensnivå hos utredarna över hela landet.

Tester för att förutsäga dyslexi

Det finns vetenskapligt underlag för att de testmetoder som mäter brister i fonologisk medvetenhet, snabb automatiserad benämningsförmåga och bokstavskänedom kan förutsäga dyslexi hos barn som inte lärt sig läsa och skriva. Här finns möjligheter att tidigt uppmärksamma de barn som skulle ha extra glädje av språkstimulerande insatser.

Insatser

Det finns vetenskapligt underlag för att strukturerad träning av fonem–grafem-koppling har positiv effekt på att förbättra ordavkodning avseende korrekthet och snabbhet, stavning, läsförståelse samt fonologisk medvetenhet jämfört med ordinarie undervisning. Det är därför viktigt med insatser i skolundervisningen, där strukturerad träning av fonem–grafem-koppling ingår, i synnerhet för barn med dyslexi. Det kan påpekas att även om det inte finns vetenskapligt underlag för att träning av enbart fonologisk medvetenhet har effekt, så ingår fonemisk medvetenhetsträning vanligen som en del i den strukturerade fonem-/grafem träningen.

Resurser och utbildning

Rapporten har identifierat en rad kunskapsluckor, se Kapitel 14. Forskningsområdet *tester och insatser* har omfattande behov av forskning av hög metodologisk kvalitet. Det är viktigt att höja standarden och statusen på den forskning som produceras inom utbildningsområdet. För det krävs en forskningsvänlig skolmiljö där skolan får ha sin vardag och ta vara på den tid som ges till elevernas kunskapsinhämtning, samtidigt som forskarna kan ges möjlighet att genomföra forskning med ett sunt vetenskapligt tänkande och med väl utprovade metoder. En sådan satsning får inte störa skolan i dess vardagliga miljö utan bör ges extra pedagogiska resurser av staten för att alla elever ska ges möjlighet att ingå i den vanliga undervisningen. För att höja kunskaperna inom området dyslexi bör skolan ges extra resurser för att regelbundet kunna vara en del i framtidens forskning.

När nya tester utvecklas för att förutsäga, upptäcka och utreda dyslexi är det viktigt att utbilda dem som ska använda testerna. Det är nödvändigt med fortbildning och kompetensutveckling baserat på vetenskapligt evidensbaserade metoder och tester. Det är också viktigt att dessa kunskaper förmedlas i lärarutbildningarna.

Framtida utmaningar för forskningen

Många studier uppfyller inte vedertagna vetenskapliga kriterier. De flesta studier som vi granskade var inte randomiserade. I de studier som var randomiserade var tillvägagångssättet inte utförligt beskrivet och randomiseringsförfarandet påverkar i stor utsträckning hur tillförlitlig en studie är. En förklaring kan vara att det inom forskningsfältet traditionellt har betraktats som tillräckligt att uppge om en randomisering har skett och inte hur den har skett.

Forskarna måste kritiskt fundera över antalet personer som inkluderas i studierna, utföra så kallade power-beräkningar, så att en tillräcklig styrka uppnås för att kunna genomföra statistiskt tillförlitliga analyser. Det är lämpligt att man i studierna fastställer vilka de primära utfallsmåtten är så att det blir tydligt vad försökspersonerna ska bli bättre på. I studien av Everatt lyftes andra metodologiska problem fram som försvårar tolkningen av forskningsresultaten inom dyslexiområdet [180].

Om interventionsstudier inte är väl kontrollerade kan det leda till att den interna valideringen blir lidande. Det betyder att man inte kan vara säker på att man testar det man avser att testa. På motsvarande sätt är det viktigt att beskriva programtrohet (fidelity) till de insatser/program som används. Att vedertagna kontrollgrupper ofta saknas kan förklaras av etiska skäl, nämligen att det inte anses etiskt försvarbart att undanhålla en potentiellt effektiv insats för barnen.

För teststudier saknas en gold standard, dvs ett test som anses vara det bästa tillgängliga. Det ställer till problem eftersom man ska utvärdera vilket test som är bäst lämpat för användning.

Eftersom det internationellt inte finns någon gemensam definition av dyslexi är sammanvägning av data för grupper med dyslexi världen över en risk.

14. Kunskapsluckor och angelägna forskningsområden

Vår systematiska genomgång av vetenskapliga studier om tester och insatser för barn och ungdomar med dyslexi visar att det finns behov av mer forskning med hög vetenskaplig kvalitet.

Tester för att förutsäga dyslexi

Det finns få framåtriktade studier, dvs longitudinella, av tillräcklig kvalitet för att identifiera underliggande förmågor som kan förutsäga att ett barn har dyslexi redan innan det fått lästräning. Många studier är inte tillräckligt tydliga med hur de definierar gruppen barn med dyslexi. En anledning till att så få studier identifierades är att många författare inte undersökte sambandet mellan prediktor (den förmåga som mättes) och manifest dyslexi. Det är också viktigt att kommande studier följer ett stort antal barn över en lång tid, gärna under flera år, samt att sambandsanalysen mellan test(er) och senare dyslexi redovisas på ett tydligt sätt.

Tester för att upptäcka och utreda dyslexi

Det går inte att avgöra vilka svenska tester som är tillräckligt bra för att upptäcka och utreda dyslexi bland barn och ungdomar. Det behövs således mer forskning om de enskilda testernas tillförlitlighet. En orsak till denna kunskapslucka kan vara att det inte finns en tradition inom fältet att utvärdera och publicera forskning om testernas prestanda. Dessutom saknas ett bästa möjliga test (referenstest eller så kallad gold standard) att jämföra testerna med. Man kan då inte uttala sig om vilket test som är bättre än ett annat, utan kan bara jämföra de olika testerna med varandra för att undersöka om de mäter liknande förmågor. Vanliga referenstester (tester som man jämför det studerade testet mot) i Nordamerika är Comprehensive Test of Phonological Processing (CTOPP) och Test of

Word Reading Efficiency (TOWRE). Det saknas dock studier där dessa tester själva utvärderas.

Insatser

Enligt Skollagen ska de metoder som används i skolan vila på vetenskaplig grund, dvs det ska finnas stöd i forskningen för de metoder som används. Det behövs därför mer forskning där effekten av olika interventioner som används i den svenska skolan utvärderas. Det är ofta individuella insatser som utvärderas eller små grupper om två till fem barn, men det skulle behövas fler studier där effekten av klassrumsinsatser (stora grupper) utvärderas. Effekten av de insatser som ges mäts oftast direkt efter avslutad insats (så kallad post-test-mätning) men det skulle även vara värdefullt att nyproducerad forskning mäter effekten av insatser efter längre tid, t ex ett till två år efter avslutade interventioner (så kallad follow-up-mätning). Det saknas även forskning som studerar insatser för barn och ungdomar med dyslexi för utfallsmåten självkompetens, självförtroende, livskvalitet, kunskapsutveckling och självständighet.

Alternativa verktyg

Det finns inga studier av tillräcklig kvalitet som undersöker värdet av alternativa verktyg för barn och ungdomar med dyslexi. Begreppet alternativa verktyg och kompenserande program skulle även behöva en tydligare definition.

Det saknas studier om vilka effekter kompenserande insatser har för att utveckla den skriftspråkliga förmågan eller snarare för att ta till sig och förmedla text. Den teknik som nu finns tillgänglig för att kompensera för läs- och skrivsvårigheter kräver utbildningsinsatser för både elev och lärare. Bara ett fåtal studier finns som studerar om smarta telefoner, surf- och läsplattor och tillhörande applikationer (så kallade appar) kan utgöra ett stöd för personer med dyslexi. Inga relevanta studier identifierades som uppfyllde inklusionskriterierna. Den snabba utvecklingen av nya tekniska hjälpmedel gör att det är svårt att utvärdera forskning inom detta område.

Systematiska översikter saknas

Det saknas nästan helt systematiska översikter inom dyslexiområdet. Vi identifierade en systematisk översikt av tillräcklig kvalitet [102]. Översikten studerade musikterapi för barn och ungdomar med dyslexi, men det saknades relevanta studier för att kunna avgöra om musik förbättrar läsning, läsförståelse, läsflyt, stavning eller fonologisk medvetenhet.

Etik

Det behövs mer forskning kring etiska aspekter för barn och ungdomar med dyslexi. De få studier som finns berör enbart forskningsetiska frågeställningar. Det behövs också forskning som utvärderar händelser i mötet mellan professionella, elever och föräldrar. Dessa tre grupper borde tillfrågas om sina erfarenheter.

Hälsoekonomi

Vi hittade inga studier om de olika insatsernas kostnadseffektivitet för att diagnostisera och erbjuda insatser för barn och ungdomar med dyslexi. Det skulle behövas fler hälsoekonomiska studier som underlag till beslutsfattare som ska prioritera mellan olika screening- och utredningstester samt insatser.

15. Ordförklaringar och förkortningar

Abstrakt	Artikelsammanfattningar
Accuracy	Exakthet
Alfabetiska skriftspråket	Alfabetiska skriftspråk använder tecken för olika språkljud; vissa bara konsonanter (hebreiska), vissa konsonanter samt vokaler (latinska/grekiska) samt vissa stavelser (japanska hiragana/katakana)
Blindning	Maskering, åtgärder för att hemlighålla vissa centrala omständigheter i en undersökning tills den är avslutad och resultaten ska bearbetas
Bias	Förutfattad mening. I forskningssammanhang ett metodproblem som ökar risken för systematiska tolkningsfel
Cochrane Collaboration	Internationellt nätverk av läkare, forskare och patienter; som utför systematiska översikter och metaanalyser av publicerad forskning
Confounder	Förväxlingsfaktor
Cronbachs-alfa	Sammanfattande korrelationsmått mellan samtliga items inom ett test
Cut-offgräns	Ett gränsvärde för avskiljning av en grupp i en fördelning

Diagnos/ kriteriediagnos	Bestämmande eller konstaterande att ett visst tillstånd föreligger enligt någon klassificering som gjorts inom medicinen, psykologin eller vårdvetenskapen. Kriteriediagnos kallas det då man behöver uppfylla ett visst antal kriterier, t ex minst sex av tio kriterier, för att en diagnos ska kunna ställas
Differentialdiagnostik	Typ av medicinsk uteslutningsmetod som sjukvårdspersonal använder sig av för att ställa en diagnos
Evidens	Något som bedöms tyda på att ett visst förhållande gäller, vetenskapliga belägg
Evidensstyrka	Beskriver tillförlitligheten i rapportens sammanvägda resultat
Evidensgradering	SBU använder det internationellt utarbetade systemet GRADE för evidensgradering i syfte att bedöma det vetenskapliga underlagets styrka
Fidelity	Programtrohet innebär den grad av modellen som faktiskt utförs och genomförs som den var ämnad
Follow up-test	Mätning som utförs en tid efter att en insats är genomförd
Fonem–grafem-koppling	Kopplingen mellan språkljud (fonem) och bokstäver (grafem)
Fonem	Språkljud (språkets minsta betydelseskiljande delar)
Fonemisk medvetenhet	Förmågan att identifiera och manipulera språkljud

Fonologisk medvetenhet	Förmåga att uppmärksamma språkets ljudmässiga uppbyggnad. Man förstår t ex principen för rim och att ordet tåg är kortare än ordet nyckelpiga, trots att tåget är långt
GRADE	Grading of Recommendations Assessment, Development and Evaluation. SBU använder det internationellt utarbetade systemet GRADE för evidensgradering i syfte att bedöma det vetenskapliga underlagets styrka
Grafem	Skriftspråkets minsta betydelskiljande enhet. Det kan vara bokstäver eller siffror
Gold standard	Bästa möjliga test/metod som de övriga kan jämföras med
Icke-systematisk översikt	En översikt som oftast bygger på studier som författaren själv valt och känner till. Det ingår ofta ingen förteckning över hur studier identifierats med hjälp av systematisk litteratursökning
Implementering	Att införa nya tester och metoder
Indextest	Det studerade/undersökta testet
Inre konsistens	Ett reliabilitetsmått på hur väl de enskilda uppgifterna i ett test mäter samma sak (se Cronbach's alfa)
Intervention	Ingrepp eller åtgärd för att åstadkomma en specifik förändring
KI	Konfidensintervall
Kohort	Grupp av personer som har vissa definierade egenskaper gemensamt

Kohortstudie	I en kohortstudie följs en eller flera grupper av individer över en tidsperiod då såväl exponering som utfall mäts. Kohortstudier kan följa individerna under en period framåt i tiden (prospektiv studie) eller kartlägga sådant som har hänt tidigare (retrospektiv studie)
Kognitiv förmåga	Förmåga att uppfatta, bearbeta, lagra och återge information – t ex minne och inlärningsförmåga
Kompositmått	Sammansatt mått
Konfidensintervall	Ett talintervall som med viss angiven sannolikhet innefattar det sanna värdet av t ex ett medeltal eller en oddskvot. Konfidensintervallet innehåller alla tänkbara värden som inte kan förkastas på grundval av föreliggande data. Vanligen anges övre och nedre gräns för ett konfidensintervall som har 95 procents sannolikhet
Korrelationsanalys	Metoder för att mäta styrkan i sambandet mellan två variabler
Longitudinell studie	Studie som följer en individ under en (inte obetydlig) tid
Läsflyt	Den sammanhållande länken, kittet, mellan avkodning och förståelse. När man har gott läsflyt läser man med passande läshastighet och betoning så att det blir förståelse i det lästa
Metaanalys	En statistisk metod som utnyttjar resultaten från flera studier för att ge ett sammanvägt resultat
Non-ord	Bokstavslängor som är uppbyggda som vanliga ord men utan betydelse

Normering	Statistisk behandling av testresultat så att man kan ange någon form av standardpoäng. För att normeringen ska vara tillförlitlig måste ett tillräckligt stort antal individer delta i utprovningen
Oddsquot	Kvoten mellan två odds. Exempel: odds exponerade/icke exponerade bland personer med viss sjukdom dividerat med motsvarande odds bland friska. Oddsquoten ger en uppfattning om hur starkt sambandet är mellan exponeringen och sjukdomen. Förkortningen OR används ofta för den engelska termen odds ratio
Ortografi	Stavningssystemet i ett språk
Ortografisk förmåga	Förmåga att stava rätt. Ortografi är läran om ordens rätta stavning
Percentilvärde	Percentilgränserna delar upp en serie rangordnade mätetal i 100 lika stora delar
Phonics	Träning av sambandet mellan grafem och fonem på ett strukturerat sätt
Post-test	Mätning direkt efter insatsens slut
Prediktor	Ett testresultat eller annat förhållande som anses kunna utsäga något om framtida skeenden, t ex att individen löper ökad risk att få dyslexi
Psykometri	Statistisk disciplin som avser att kvalitetsgranska test
RAN	Rapid Automated Naming vilket betyder snabb automatiserad benämningsförmåga

Randomiserad kontrollerad undersökning	Jämför två eller flera grupper till vilka deltagarna har fördelats slumpmässigt. Upplägget görs på så sätt att grupperna blir lika, förutom den faktor som studeras
RCT	Randomised controled trial (randomiserad kontrollerad undersökning)
Referenstest	Test som man jämför det studerade testet mot
Reliabilitet (precision)	Beskriver hur pass tillförlitligt och noggrant en mätning genomförs. En hög reliabilitet innebär en liten andel slumpmässiga mätfel
Riskbedömning	Bedömning av riskerna för att någon kan komma att drabbas av dyslexi
Screening	Engelskspråkig term för gallring, när man med hjälp av test eller annan metod vill skilja ut en grupp för närmare undersökning eller forskning
Självkompetens (self-efficacy)	Engelskans self-efficacy innebär att individen har tilltro till den egna förmågan att kunna klara en specifik uppgift med ett visst resultat. Ju mer en person tror sig klara av en uppgift desto större är chanserna att hon eller han gör det. En person med hög self-efficacy väljer ofta mer utmanande uppgifter och har högre förmåga till uthållighet än personer med lägre grad av self-efficacy. Det är svårt att översätta begreppet self-efficacy till svenska men uttrycket självkompetens har använts
Sensitivitet	Ett mått på i vilken utsträckning testet korrekt förmår identifiera elever med dyslexi, dvs sannolikheten för att elever med dyslexi får ett positivt testresultat

Signifikant	Ett statistiskt mått på om en observation beror på slump eller inte
SMD	Standardiserad medelvärdeskillnad, används som en sammanfattande statistik i metaanalyser då studierna bedömer samma effektmått, men ofta använder sig av olika skalor. I de fallen är det nödvändigt att standardisera resultaten av studierna till en enhetlig skala innan de kan kombineras. En vanlig metod för att beräkna SMD är Cohen's d
Socioekonomiska faktorer	Faktorer såsom inkomst, utbildning, yrke och arbete/sysselsättning
Specificitet	Ett mått på i vilken utsträckning testet korrekt förmår identifiera elever som inte har dyslexi, dvs sannolikheten för att elever utan dyslexi får ett negativt testresultat
Språklig medvetenhet	Förmåga att reflektera över det egna språket, att kunna se på fler aspekter av språket än de rent innehållsmässiga. Sådana aspekter kan t ex vara att kunna rimma, att kunna uppmärksamma ord i meningar, att kunna markera stavelser och att kunna urskilja enstaka språkljud
Standardiserat test	Test som utprovats på en större grupp, t ex en viss årskurs eller en viss åldersgrupp. Testpoängen omvandlas till standardpoäng i en standardskala, t ex stanineskalan
Stanineskala	Normerad skala, niogradig, 1-3; under genomsnittet 4-6; normalvärde 7-9; över genomsnittet

Systematisk översikt	En översikt som avser en tydligt formulerad fråga och som använder systematiska och explicita metoder för att identifiera, välja ut och kritiskt bedöma relevanta studier och som samlar in och analyserar uppgifter från dessa. Statistiska metoder (metaanalys) används ibland för att analysera och sammanfatta resultaten av de inkluderade studierna
Test-retest-reliabilitet	Resultatet ska vara detsamma vid upprepade mätningar
Tvärsnittsstudie	I en tvärsnittsstudie mäts exponering och utfall vid ett enda tillfälle
Validitet (accuracy)	Mått på hur väl man mäter det man vill mäta, dvs hur väl syfte och resultat stämmer överens

16. Personer som medverkat till rapporten

Projektgrupp

Stefan Samuelsson (ordföranden)

Professor i specialpedagogik, Linköpings universitet, Linköping

Barbro Bruce

Legitimerad logoped, med dr i logopedi, universitetslektor i utbildningsvetenskap/specialpedagogik, Malmö högskola, Malmö

Gunnel Colnerud (avsnitt om etiska och sociala aspekter)

Legitimerad psykolog och professor i pedagogik, Linköpings universitet, Linköping

Christina Hellman

Fil dr, universitetslektor i lingvistik, Stockholms universitet, Stockholm

Mona Näreskog

Lärare, legitimerad läkare och specialist i barn- och ungdomspsykiatri, Stockholm

Idor Svensson

Docent i psykologi, legitimerad psykolog, specialist i klinisk och pedagogisk psykologi, Linnéuniversitetet, Växjö

Ulrika Wolff

Docent i pedagogik, Göteborgs universitet, Göteborg

Jakob Åsberg

Lärare, fil dr psykologi, Göteborgs universitet, Göteborg

Extern referensexpert

Petter Gustavsson

Professor i differentiell psykologi, Karolinska Institutet, Stockholm

Extern referensperson

Gunilla Salo

Specialpedagogiska skolmyndigheten, Stockholm

Kansli

Agneta Brolund (informationsspecialist)

SBU, Stockholm

Thomas Davidson (hälsoekonom)

SBU, Stockholm

Kickan Håkanson (projektadministratör)

SBU, Stockholm

Agneta Pettersson (biträdande projektledare)

SBU, Stockholm

Karin Stenström (projektledare)

SBU, Stockholm

Externa granskare

SBU anlitar externa granskare till sina rapporter. De har kommit med värdefulla kommentarer, som i hög grad bidragit till att förbättra rapporten. I slutversionen av rapporten har SBU dock inte kunnat tillgodose alla ändrings- eller tilläggförslag från de externa granskarna, bl a därför att de inte alltid varit samstämmiga. De externa granskarna står därför inte nödvändigtvis bakom samtliga slutsatser eller andra texter i rapporten.

Carsten Elbro

Professor, Institut for Nordiske Studier og Sprogvidenskab,
Københavns universitet, København, Danmark

Bente Eriksen Hagtvet

Professor, Institutt for spesialpedagogikk, Oslo, Norge

Pekka Niemi

Professor, psykologiska institutionen, Åbo universitet, Finland

Eva Wigforss

Fil doktor, leg logoped, Lunds universitet

Bindningar och jäv

Sakkunniga och granskare har i enlighet med SBU:s krav inlämnat deklARATION rörande bindningar och jäv. Dessa dokument finns tillgängliga på SBU:s kansli. SBU har bedömt att de förhållanden som redovisas där är förenliga med krav på saklighet och opartiskhet. I arbetet med att relevans- och kvalitetsgranska studier bedömde de sakkunniga experterna inte artiklar där de själva var medförfattare eller på annat sätt jäviga.

17. Studier som ligger till grund för resultat och slutsatser

Table 17.1 At-risk studies of high or moderate quality used for results and conclusions in the present report.

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Catts 1991 [26] USA	<p>Design Longitudinal study</p> <p>Aim Examine if speech-language impairment children are at risk for development of reading problems upon entering school</p> <p>Setting Kindergarten in 1 of 2 Midwest public school districts</p> <p>Population 41 speech-language impairment 30 comparison group (without speech-language impairment)</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p> <p>Follow-up Grade 1</p>	<p>Pre-test Phonological awareness (deletion and blending tasks) Phonetic recording (RAN)</p> <p>Post-test/predictors Subtest from the WJ-R Word Attack Word Identification The poor readers performance criterion of at least 1 standard deviation below the mean on at least 2 of 3 receptive language measures and/or 3 of 3 expressive language measures</p> <p>Participants n=41 speech-language impairment 6 girls, 32 boys Age: 6 years 4 months</p> <p>n=30 12 girls, 18 boys Age: 6 years, 1 month</p>	<p>Percentage of correct classification of reading group membership for each of the language measures (Table III)</p> <p>RECL: 78.1 EXPL: 61.0 DELETION: 75.6 BLENDING: 70.7 RANAN: 75.6 RANOBJ: 51.2 RANCOL:70.7 SENMEN: 75.6</p> <p>Correct classification Good at phonological awareness - good at reading: 87.5% Bad at phonological awareness - later poor readers: 89% (among the speech-language impairment subjects)</p>	<p>Moderate</p> <p>Other comments Only speech-language impairment</p>

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Elbro et al 1998 [27] Denmark	<p>Design Longitudinal study</p> <p>Aim Examine the possibility to predict dyslexia from kindergarten</p> <p>Setting Schools in Copenhagen</p> <p>Population n=96. Half of the participants in the study were children of dyslexic parents (n=51); the other half were children of normally reading parents (n=45)</p> <p>Inclusion criteria None</p> <p>Exclusion criteria Families were excluded from the study in cases where found a mismatch between the self-reported reading history and the objective findings. If no reliable information about the reading status of one of the parents</p> <p>Follow-up Grade 2</p>	<p>Pre-test Letter naming, initial-phoneme deletion, phoneme identification, pronunciation accuracy, distinctness of phonological representations</p> <p>Post-test/predictors Several reading and language tests (non-word reading, identification of pseudo homophones) Cut-off point at reading level: 10–12% of the poorest readers</p> <p>Participants 1 dyslexic parent n=49 26 girls, 23 boys Normal reading parents n=42 20 girls, 22 boys</p> <p>Drop-out rate n=5 (between pre- and post-test) n=2 (repeated kindergarten) n=2 (moved) n=1 (not tested due to technical error)</p>	<p>Logistic regression analysis with backwards stepwise selection of predictors; Letter naming $X^2=7.3$ ($p<0.01$) Phoneme identification $X^2=5.5$ ($p<0.05$) Distinctness of phonological representations $X^2=10.4$ ($p<0.01$)</p> <p>All 5 measures tested at kindergarten (pre-tests) are predictors of reading difficulties with cut-off value $p=0.25$ Sensitivity: 0.78 Specificity: 0.79 Correct classification rate: 0.79</p>	<p>Moderate</p> <p>Other comments There are also sensitivity and specificity values for the cut-off value of $p=0.5$ available in the article</p>

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Furnes et al 2010 [28] Norway	<p>Design Longitudinal study</p> <p>Aim To compare the prediction of reading and spelling difficulties at the end of Grade 1 and 2 across transparent and opaque orthographies</p> <p>Setting Australia, Norway and USA</p> <p>Population n=2 006 754 same-sex twin pairs from Australia and the USA and 249 same-sex twin pairs from Scandinavia, a total of 1 003 pairs or 2 006 children</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p> <p>Follow-up Grade 1 and 2 (for Norwegian sample)</p>	<p>Pre-test LK PA RAN</p> <p>Post-test/predictors Subtest from TOWRE measured phonological decoding and word recognition. Subtest from WRAT measured spelling</p> <p>Cut-off criterion at the 15th percentile (1 standard deviation) on composite word reading score, spelling skills and spelling difficulties</p> <p>Participants In preschool Australia/USA n=1 508 738 girls, 769 boys</p> <p>Mean age: 4.8 years In preschool Scandinavia n=498 249 girls, 249 boys Mean age: 5.1 years</p> <p>Drop-out rate USA/Australia sample, n=133 (Pres-Grade 2) Scandinavia sample, n=218 (Pres-2)</p>	<p>Results of logistic regression analyses predicting difficulties in word reading in Grade 1 and 2 from letter knowledge (LK), phonological awareness (PA) and RAN in kindergarten OR (95% CI)</p> <p>Australia/USA sample Grade 1 LK: 0.61 (0.51; 0.74) p<0.001 PA: 0.30 (0.22; 0.42) p<0.001 RAN: 2.09 (1.71; 2.56) p<0.001</p> <p>Grade 2 LK: 0.83 (0.70; 0.99) p<0.05 PA: 0.36 (0.27; 0.48) p<0.001 RAN: 2.13 (1.75; 2.60) p<0.001</p> <p>Scandinavia sample Grade 1 LK: 0.56 (0.36; 0.87) p<0.05 PA: 0.36 (0.16; 0.82) p<0.05 RAN: 2.44 (1.52; 3.92) p<0.001</p> <p>Grade 2 LK: 0.51 (0.33; 0.79) p<0.001 PA: 0.80 (0.44; 1.76) RAN: 2.38 (1.53; 3.72) p<0.001</p> <p>Results of logistic regression analyses predicting difficulties in spelling in Grade 1 and 2 from letter knowledge, phonological awareness and RAN in kindergarten OR (95% CI)</p> <p><i>Results continue on the next page</i></p>	Low

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
<i>Continued</i>			<i>Australia/USA sample</i>	
Furnes et al 2010 [28] Norway			Grade 1 LK: 0.66 (0.56; 0.79) p<0.001 PA: 0.36 (0.27; 0.47) p<0.001 RAN: 1.42 (1.91; 1.69) p<0.001	
			Grade 2 LK: 0.80 (0.68; 0.96) p<0.05 PA: 0.28 (0.20; 0.37) p<0.001 RAN: 1.52 (1.26; 1.83) p<0.001	
			<i>Scandinavia sample</i>	
			Grade 1 LK: 0.43 (0.27; 0.69) p<0.001 PA: 0.13 (0.04; 0.39) p<0.001 RAN: 1.94 (1.18; 3.18) p<0.01	
			Grade 2 LK: 0.65 (0.42; 0.99) p<0.05 PA: 0.37 (0.17; 0.82) p<0.05 RAN: 1.55 (1.01; 2.39) p<0.05	

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Helland et al 2011 [29] Norway	<p>Design Longitudinal study</p> <p>Aim To predict dyslexia in children ahead of formal literacy training. Data gather at age 5 as predictors of dyslexia at age 11</p> <p>Setting 9 Norwegian preschools</p> <p>Population 109, 5-year old children followed 5 years (from age 5 to age 8)</p> <p>Inclusion criteria All 4 counties in western Norway (Rogaland, Hordaland, Sogn og Fjordane and Møre og Romsdal) and both urban and rural districts should be represented. Norwegian as first language. Dyslexia group defined as cut-off at the low 10% score on the literacy tests</p> <p>Exclusion criteria Impaired sight or hearing, intellectual disability according to DSM-IV or any other impairment included in DSM-IV (ADHD, autism spectrum disorders and neurological impairments)</p> <p>Follow-up 4 years (from age 5 to 8 years old)</p>	<p>Pre-test Risk index questionnaire (RI-5) to caretakers of the children and to the respective preschool teachers was used as a screening instrument; questions about health, laterality, motor skills, and language, special need education and heredity. The RI-5 test sorted the children into 3 categories: at-risk, matched controls and remainder group. Test of reading and spelling</p> <p>Post-test/predictors New questionnaire (RI-11) filled out by parents. Individual testing of defined risk group and match controls. Tests of reading and spelling</p> <p>Participants At baseline, age 5 n=25 at-risk group n=24 matched controls n=53 remainder group</p> <p>Follow-up at age 8: n=25 at-risk group n=24 control</p> <p>Follow-up at age 11: n=22 at-risk group n=20 control</p> <p>Drop-out rate n=7</p>	<p>Sensitivity RI-5:0.85</p> <p>Specificity RI-5:0.62</p>	<p>Moderate</p> <p>Risk of bias affected by Case-control design</p> <p>Other comments The questionnaire approach could help to identify children with dyslexia at the age of 5</p> <p>There are inter-rater correlations available in the article (Table 3)</p>

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Lamm et al 1999 [30] Israel	<p>Design Longitudinal study</p> <p>Aim Examine kindergarten dichotic listening as a predictor of first grade reading achievement</p> <p>Setting North of Israel</p> <p>Population 328 kindergartens, all monolingual native speakers of Hebrew</p> <p>Inclusion criteria None</p> <p>Exclusion criteria Hearing deficits</p> <p>Follow-up K-1 (end of year 1)</p>	<p>Pre-test 2 verbal dichotic tests [181]</p> <p>Post-test/predictors 2 verbal dichotic tests. Cluster of 3 test: designed to evaluate word reading, reading comprehension, and listening comprehension</p> <p>Participants Baseline: n=328 147 girls, 143 boys</p> <p>Follow-up n=290 children tested dichotically in both years, 288 were also administrated tests of word reading, reading comprehension and listening towards the end of their first year in school</p> <p>Mean age: 5.8–6 years</p> <p>Drop-out rate n=38 (changed their area of residence or remained in kindergarten)</p>	<p>Subjects' distribution according to kindergarten dichotic listening performance by first-grade word reading (n=288). Data calculated from table 4: Sensitivity: 0.77 Specificity : 0.93 Correct classification rate: 0.91</p>	Low

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Niemi et al 2011 [31] Finland	<p>Design Longitudinal study</p> <p>Aim To delineate predictors of resistance to treatment that are evidenced as little or no reading progress during Grade 1</p> <p>Setting Municipalities in Finland</p> <p>Population 1 987 children complete data for both fall and spring of Grade 1, kindergarten readers (n=702)</p> <p>Inclusion criteria All were non-readers at school entrances</p> <p>Exclusion criteria None</p> <p>Follow-up Grade 1</p>	<p>Pre-test Phonological awareness Language knowledge RAN Number sequence skills</p> <p>Post-test/predictors Reading performance and match Cut-off points of -1 standard deviation and +1 standard deviation in reading and math</p> <p>Participants n=1 285 571 girls, 714 boys</p> <p>Drop-out rate n=211</p>	<p>Summary of multinomial logistic regression analysis for variables predicting membership in groups of slow reading and match acquisition. OR (95% CI); n=1 074</p> <p>Slow learners in reading Phonological awareness: 0.822 (0.643; 1.050) Language knowledge: 0.625 (0.479; 0.815) RAN: 1.439 (1.137; 1.821)</p>	<p>Moderate</p> <p>Risk of bias affected by: Case-control design. Slow learners in reading are compared with fast readers</p> <p>Other comments Data on slow learners in reading and math is available in article</p>

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
O'Connor et al 1999 [32] US	<p>Design Longitudinal study</p> <p>Aim To design a small set of phonological, letter, and memory tasks that would reliably identify children likely to develop reading difficulties</p> <p>Setting Cohort 1: 4 schools in a district located in a midsize agricultural community in the Pacific Northwest</p> <p>Population 197 children from kindergarten in cohort 1</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p> <p>Follow-up K-1</p>	<p>Pre-test Syllable deletion, Segmentation phonemes, Rapid letter naming, Sound repetition, PPVT-R, Phonological measures, Syllable blending and segmentation, Blend phonemes, First-sound isolation, Rhyme production, Dynamic segmentation</p> <p>Post-test/predictors Subtests: Word Identification and Word Attack from WRMT</p> <p>To classify children as reading difficulties they had to score more than 1.4 standard deviation below the mean score (from the Grade 1 norm table) on the combined Word Identification and Word Attack subtest of WRMT</p> <p>Participants n=129</p> <p>Drop-out rate n=28 (ongoing kindergarten intervention) n=23 (children moved out from the area) n=17 (children scored lower than 70 on PPVT-R)</p>	<p>Classification Rates for predictors of November and April for cohort 1 (n=129) Data calculated from table 4:</p> <p>November Predictors: Segmentation phonemes, Rapid letter naming, Syllable deletion Sensitivity: 1.0, specificity: 0.88 Correct classification rate: 0.88</p> <p>April Predictors: Segmentation phonemes, Rapid letter naming, Sound repetition Sensitivity: 1.0, specificity: 0.87 Correct classification rate: 0.88</p> <p>Table 3, Cohort 1: November Best predictors: Syllable deletion (0.996), Segmentation phonemes (0.485), Rapid letter naming (0.429)</p> <p>April Best predictors: Rapid letter naming (0.785), Sound repetition (0.757), Segmentation phonemes (0.755)</p>	Low

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Pennington et al 2012 [19] USA	<p>Design Prospective cohort, International Longitudinal Twin Study sample</p> <p>Aim To determine the clinical utility of models (PA only, L only, PS/NS only, Any Single, 2–3 Deficits, Any Deficit) for predicting and diagnosis of dyslexia</p> <p>Setting Preschool age twins from 3 countries (Australia, Norway, and USA) followed into early school age</p> <p>Population There were in total 809 twin pairs and here they randomly selected 1 twin from each pair for the study sample of 809 individuals: n=82 individuals as dyslexic n=727 non-dyslexic</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p> <p>Follow-up Grade 1</p>	<p>Pre-test 3 cognitive measures: PA, L, PS or NS</p> <p>PA: CTOPP (Elision, Blending, Words and Sound Matching)</p> <p>L: Grammatic Closure: Morphology; WPPSI, Vocabulary; 100 Pictures Naming Test; and Test for Reception of Grammar. PS/NS: TOWRE (Phonemic Decoding and Sight Word Efficiency)</p> <p>Post-test Dyslexia defined as $\leq 10^{\text{th}}$ percentile on a reading composite measure (i.e., the average of both subtests from the TOWRE A and TOWRE B) that measured the construct of single word and non-word reading accuracy and fluency</p> <p>Participants n=809 children 398 girls, 411 boys</p> <p>Drop-out rate Not reported</p>	<p>Predictive value (% or ratio) of different deficit profiles</p> <p>Sensitivity (%) PA only: 15.9 PA: 41.5 L only: 2.4 L: 19.5 PS/NS only: 22.0 PS/NS: 42.7 Any Single: 40.2 2–3 Deficits: 28.0 Any Deficit: 69.3</p> <p>Specificity (%) PA only: 96.7 PA: 93.5 L only: 93.7 L: 90.9 PS/NS only: 96.1 PS/NS: 93.5 Any Single: 86.5 2–3 Deficits: 96.0 Any Deficit: 82.7</p> <p>Positive Predictive Power (true positives) (%) PA only: 35.1 PA: 42.7 L only: 4.2 L: 19.5 PS/NS only: 39.1 PS/NS: 42.7 Any Single: 25.2 2–3 Deficits: 44.2 Any Deficit: 30.8</p> <p><i>Results continues on the next page</i></p>	<p>Low</p> <p>Comments This study has also cross-sectional data available</p>

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Continued			Negative Predictive Power (%) PA only: 91.0 PA: 93.5 L only: 89.5 L: 90.9 PS/NS only: 91.6 PS/NS: 93.5 Any Single: 92.8 2-3 Deficits: 92.2 Any Deficit: 95.8	
Pennington et al 2012 [19] USA				

The table continues on the next page

Table 17.1 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Pre-test Post-test Participants Drop-out rate	Classification accuracy	Risk of bias Comments
Puolakanaho et al 2007 [33] Finland	<p>Design Longitudinal study</p> <p>Aim To examine early predictors for estimating individual risk of reading disability</p> <p>Setting Schools in Finland</p> <p>Population n=198 (n=106 familial risk of dyslexia, n=92 no familial risk). The present data were drawn from the follow-up project of Jyväskylä Longitudinal study of Dyslexia [182–184]</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p> <p>Follow-up Grade 2</p>	<p>Pre-test PA S-TM RAN EV PS LN</p> <p>Post-test/predictors Measures of reading and spelling (5 tasks providing 8 measures)</p> <p>To classified with reading difficulties the child's skills were at or below the 10th percentile either a. in at least 3 out of the 4 accuracy measure or at least 3 out of the 4 fluency measures, or b. in 2 accuracy measures and in 2 fluency measures</p> <p>Participants n=46 (RD) n=152 (nRD)</p> <p>Pre-test at the age phases: 3.5, 4.5 and 5.5 years. Post-test at the age: 8.9 years</p> <p>Drop-out rate None</p>	<p>Classification accuracy at the different probability cut-off levels using the age-specific logistic models. Analyses of RD (n=46) and nRD (n=152) children.</p> <p>The weighted procedure was used and the lowest cut-off level applied (p=0.14 for age 3.5 and 4.5; p=0.17 for age 5.5) or the sensitivity level around 90%</p> <p>Age 3.5 Sensitivity: 89.6% Specificity: 70.5% Correct classification rate: 72.8%</p> <p>Age 4.5 Sensitivity: 64.0% Specificity: 78.1% Correct classification rate: 76.4%</p> <p>Age 5.5 Sensitivity: 79.8% Specificity: 83.7% Correct classification rate: 83.2%</p>	<p>Moderate</p> <p>Risk of bias affected by Statistical bias (family risk is included in the results)</p>

CI = confidence interval; CTOPP = Comprehensive Test of Phonological Processing; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders-fourth version; EV = Expressive vocabulary; ILTS = International Longitudinal Twin Study; L = language speed; LK = Letter knowledge; LN = Letter naming; nRD = Not RD; NS = naming speed, NSS = number sequence skills; OR = odds ratio; PA = Phonological awareness; PPVT = The Peabody Picture Vocabulary Test; PPVT-R = Peabody Picture Vocabulary Test-Revised; PS = processing speed, PS = Pseudo word repetition; RAN = Rapid automatized naming; RD: reading

disabilities; RI-11 = Risk Index age 11; RI-5 = Risk Index age 5; S-TM = Short-term memory; TOWRE = Test of Word Reading Efficiency; WJ = Woodcock Johnson; WJ-R = Woodcock Reading Mastery Tests-Revised; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; WRAT = Wide Range Achievement Test; WRMT = Woodcock Reading Mastery Test;

Table 17.2 Test studies of high or moderate quality used for results and conclusions in the present report.

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Burke et al 2007 [34] USA	<p>Design Cross-sectional study</p> <p>Aim Examine the concurrent criterion-related validity of first grade measures from DIBELS</p> <p>Setting 3 elementary schools, 10 different kindergarten classrooms from mid-sized city in Northwestern Massachusetts</p> <p>Population 213 kindergarten students</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>DIBELS Subtests: Dibels oral reading fluency (DORF) and Nonsense word fluency (NWF)</p> <p>Participants n=213 children 86 girls, 109 boys</p> <p>Demographic information was unavailable for 18 participants. Age range: First grade</p> <p>Drop-out rate Not reported</p>	<p>TOWRE Subtests: Phoneme decoding efficiency (PDE) and Sight word efficacy (SWE)</p>	<p>Intercorrelations between subtests of DIBELS and subtests of TOWRE</p> <p>Validity Non-word reading DIBELS NWF-TOWRE PDE: 0.75 (p<0.05)</p> <p>Oral reading fluency DORF-TOWRE SWE: 0.92 (p<0.05)</p>	<p>Moderate (validity)</p> <p>Risk of bias affected by Little information on scoring practices</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Clarke et al 2004 [35] USA	<p>Design Prospective cohort study (K–1 Grade)</p> <p>Aim Examined psychometric properties reliability of the KRES, a brief teach-report measure of classroom reading engagement</p> <p>Setting Low-income, African-American community</p> <p>Population 27 K-students followed from K through first grade</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test KRES</p> <p>Participants n=27 9 girls, 18 boys Age: 6.2 years</p> <p>Drop-out rate Not reported</p>	Reference test N/A	<p>KRES demonstrate strong internal consistency: Cronbach α: 0.96</p> <p>Test-retest reliability adequate: 0.66 (3 weeks, n=14)</p>	<p>Moderate (reliability) Low (validity)</p> <p>Risk of bias affected by Limited information about scoring practices</p> <p>Other comments Very small sample makes generalisation difficult</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Daub et al 1996 [39] USA	<p>Design Cross-sectional study (3 weeks between the tests)</p> <p>Aim To test the correlation between different reading subtests of WJ-R, PIAT-R and DAB-2</p> <p>Setting Elementary students in a suburban school in Georgia</p> <p>Population 35 students in third and fourth-grade with learning disabilities in the area of reading</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test WJ-R</p> <p>Participants n=35 11 girls, 24 boys Age: 9–10 years</p> <p>Drop-out rate Not reported</p>	Reference test PIAT-R	<p>Pearson correlation between the reading subtest of WJ-R and PIAT-R</p> <p>Validity <i>Reading Recognition</i> WJ-R-PIAT-R: 0.71</p> <p><i>Reading Comprehension</i> WJ-R-PIAT-R: 0.73</p>	<p>Low (validity)</p> <p>Other comments Small sample size makes generalisation difficult</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Elliot et al 2001 [40] USA	<p>Design Cross-sectional study (2 weeks between the tests)</p> <p>Aim Examine the psychometric properties of a set of preliteracy measures modified from the DIBELS; DIBELS-M, which includes: Letter naming fluency Sound naming fluency Initial phoneme fluency Phonemic segmentation ability</p> <p>Setting 4 classrooms in 3 elementary schools in a moderate-size Midwestern city in US</p> <p>Population 75 kindergarten children</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test DIBELS</p> <p>Participants n=75 34 girls, 41 boys</p> <p>Drop-out rate Not reported</p>	<p>Reference test WJ-R Teacher rating questionnaire TOPA</p>	<p>Reliability estimates for the DIBELS-M measures</p> <p>Interrater reliability (n=50) Letter naming fluency: 0.94 Sound naming fluency: 0.82 Initial phoneme fluency: 0.89 Phonemic segmentation ability: 0.87</p> <p>Test-retest (n=75) Letter naming fluency: 0.90 Sound naming fluency: 0.83 Initial phoneme fluency: 0.74 Phonemic segmentation ability: 0.85</p> <p>Equivalent Forms (n=75) Letter naming fluency: 0.80 Sound naming fluency: 0.82 Initial phoneme fluency: 0.64 Phonemic segmentation ability: 0.84</p> <p>Concurrent validity of the DIBELS-M level estimate (average) scores</p> <p>Letter word identification Letter naming fluency - WJ-R: 0.71 Sound naming fluency - WJ-R: 0.62</p> <p>Prereading Letter naming fluency - Teaching rating: 0.63 Sound naming fluency - Teaching rating: 0.62 Initial phoneme fluency - Teaching rating: 0.46 Phonemic segmentation ability - Teaching rating: 0.53</p> <p>Phonological awareness Initial phoneme fluency - TOPA: 0.62 Phonemic segmentation ability - TOPA: 0.52</p>	<p>Low (reliability) Low (validity)</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Erford et al 2005 [41] USA	<p>Design Cross-sectional</p> <p>Aim Study the characteristics of Slosson Phonics and Structural analysis test (SP-SAT)</p> <p>Setting 2 schools located in the central Maryland area</p> <p>Population Children in grade 1–4 Grade 1 (n=38) Grade 2 (n=41) Grade 3 (n=44) Grade 4 (n=42)</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test WJ-III subtests WWI, WWA and WPC</p> <p>Participants n=165 85 girls, 80 boys Age: 6.5–12 years</p> <p>Drop-out rate Not reported</p>	Reference test WRAT-3	<p>Reliability - inter item consistency coefficient</p> <p>Word decoding WRAT-3: 0.90 WWI: 0.93</p> <p>Non-word decoding WWA: 0.92</p> <p>Reading comprehension WPC: 0.87</p> <p>Reliability - test-retest coefficient WRAT-3: 0.90 WWI: 0.96 WWA: 0.67 WPC: 0.90</p> <p>Pearson correlation coefficients between WJ-III and WRAT-3 for the total sample</p> <p>First administration WWI-WRAT-3: 0.90</p> <p>Retest WWI-WRAT-3: 0.90</p>	<p>Moderate (validity) Low (reliability)</p> <p>Risk of bias affected by Little information about scoring procedure and scorer knowledge</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Fien et al 2008 [36] USA	<p>Design Prospective cohort study (K–2 Grade)</p> <p>Aim Examine the validity of DIBELS NWF as an index of beginning reading proficiency for student in kindergarten through second grade. In addition, the study includes test-retest data on DIBELS NWF and ORF subtests</p> <p>Setting 34 Reading First schools in Oregon. Half of the schools in large urban areas, and the remaining schools equally divided between cities with population between 50 000 and 100 000 (8 schools) and rural areas (9 schools)</p> <p>Population Kindergarten through second grade students, divided in English learners and English speakers groups</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test DIBELS NWF</p> <p>Participants English learners n=3 310 1 718 girls, 1 592 boys</p> <p>English speakers n=11 654 6 434 girls, 5 220 boys</p> <p>Drop-out rate According to table 1: English learners group: no drop-out (n=+536) from NWF test K-fall 2nd</p> <p>English speakers group: 15.7% drop-out NWF test K-fall 2nd</p>	Reference test DIBELS ORF	Test-retest (3 weeks) reliability coefficients, for a random 20% of the population and across grades, range from: 0.84–0.90 on DIBELS NWF 0.94–0.99 on DIBELS ORF	Low (reliability)

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Gotch et al 2011 [42] USA	<p>Design Cross-sectional study</p> <p>Aim Investigate evidence of construct validity of the composite scores on the CIBS-II-Readiness by examining the internal structure of the instrument and correlations with external measures</p> <p>Setting 4 main regions in the USA</p> <p>Population A stratified random sample of kindergarteners (n=383)</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test CIBS-II</p> <p>Participants n=383 192 girls, 191 boys Mean age: 5.92 years</p> <p>Drop-out rate Not reported</p>	Reference test DIBELS	<p>Reliability Internal consistency in the reading and phonemic awareness factors: 0.85–0.96 (Table 1)</p> <p>Test-retest coefficient: 0.73–0.99</p> <p>Correlations of CIBS-II-Readiness composites with the DIBELS measure CIBS-II Reading–DIBELS Initial sound fluency: 0.58 CIBS-II Reading–DIBELS Letter naming fluency: 0.62 CIBS-II Reading–DIBELS Nonsense word fluency: 0.59 CIBS-II PA–DIBELS Initial sound fluency: 0.59 CIBS-II PA–DIBELS Letter naming fluency: 0.47</p>	<p>Moderate (validity and reliability)</p> <p>Risk of bias affected by Limited information on procedure</p> <p>Other comments Pattern of correlations seen in this study provides evidence that the Readiness measures assess skills similar to those assessed by DIBELS</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Hagan-Burke et al 2006 [37] USA	<p>Design Prospective cohort study (K–2 Grade)</p> <p>Aim To examine the convergent validity of the DIBELS with a standard measure of phonological decoding ability and sight word reading fluency, the TOWRE</p> <p>Setting Public primary school from middle to lower middle class families and located in a semi-rural area in northeast Georgia</p> <p>Population 202 kindergarten students followed to 2nd grade</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test DIBELS Subtests: Letter naming fluency (LNF) and Nonsense word fluency (NWF)</p> <p>Participants n=202 86 girls, 116 boys</p> <p>Drop-out rate None</p>	<p>Reference test TOWRE Subtests: Phoneme decoding efficiency (PDE) and Sight words (SW)</p>	<p>Intercorrelations between DIBELS subtest (fall of first grade) and TOWRE subtests DIBELS NWF-TOWRE PDE: 0.748 (p<0.01)</p>	<p>Low (validity)</p> <p>Risk of bias affected by Limited information about scoring procedure and scorer training</p> <p>Other comments DIBELS is a viable set of measures for targeting and monitoring the critical component skills that are required to learn to read</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Havey et al 2002 [43] USA	<p>Design Cross-sectional study</p> <p>Aim To determine the degree to which 2 measures of phonological awareness/ability, TOPA; CTOPP correlate with each other and with measure of reading</p> <p>Setting Public school in a small city in the Midwestern USA</p> <p>Population 81 second-semester kindergarten children</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test TOPA</p> <p>Participants n=81 34 girls, 47 boys Mean age: 6 years</p> <p>Drop-out rate Not reported</p>	Reference test CTOPP	<p>Correlations among phonological and reading measures</p> <p>Phonological awareness TOPA-CTOPP: 0.63 (p<0.001)</p> <p>Phonological memory TOPA-CTOPP: 0.47 (p<0.001)</p> <p>Rapid naming TOPA-CTOPP: 0.19</p> <p>Letter-word TOPA-CTOPP: 0.49 p<0.001)</p>	<p>Low (validity)</p> <p>Other comments CTOPP were the best predictor of performance on a measure of letter knowledge and word reading</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Hintze et al 2003 [44] USA	<p>Design Cross-sectional study</p> <p>Aim Examine the concurrent validity and the diagnostic accuracy of the DIBELS compared to the CTOPP</p> <p>Setting 3 elementary schools, 10 different kindergarten classrooms from midsized city in Northwestern Massachusetts</p> <p>Population 86 kindergarten students randomly selected for participation</p> <p>Inclusion criteria None</p> <p>Exclusion criteria None</p>	<p>DIBELS 3 subtests: Letter naming fluency (LNF) Initial sound fluency (ISF) Phonemic segmentation fluency (PSF)</p> <p>Participants n=86 children 44 girls and 42 boys Age range 5 years, 3 months to 6 years, 7 months</p> <p>Drop-out rate Not reported</p>	<p>CTOPP 3 subtests: Phonological awareness (PA) Phonological memory (PM) Rapid memory (RM) Elision (ELI) Sound matching (SM)</p>	<p>Intercorrelations for scores on the DIBELS and CTOPP</p> <p>Phonological awareness DIBELS PSF-CTOPP ELI: 0.47 DIBELS PSF-CTOPP SM: 0.25 DIBELS PSF-CTOPP PACom: 0.53</p> <p>The study evaluated different cut-off values. The data show high sensitivity and insufficient specificity regardless of cut-off level</p> <p>Performance of the DIBELS subtests (ISF, LNF and PSF) over cut-scores 24, using the CTOPP composite score as the criteria</p> <p>ISF (CTOPP Phonological Awareness Composite) Sensitivity 100 Specificity 39</p> <p>ISF (CTOPP Phonological Memory Composite) Sensitivity 91 Specificity 36</p> <p>ISF (CTOPP Rapid Naming Composite) Sensitivity 90 Specificity 37</p> <p>LNF (CTOPP Phonological Awareness Composite) Sensitivity 87 Specificity 34</p> <p>LNF (CTOPP Phonological Memory Composite) Sensitivity 100 Specificity 35</p>	<p>Low (validity)</p> <p>Other comments The order between the 2 measures was counterbalanced. Approximately 20 minutes sessions for each measure. Maximum 3 school days between the administrations of the 2 measures</p>

Results continues on the next page

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
<i>Continued</i>				LNF (CTOPP Rapid Naming Composite) Sensitivity 100 Specificity 36	
Hintze et al 2003 [44] USA				PSF (CTOPP Phonological Awareness Composite) Sensitivity 93 Specificity 23	
				PSF (CTOPP Phonological Memory Composite) Sensitivity 100 Specificity 23	
				PSF (CTOPP Rapid Naming Composite) Sensitivity 80 Specificity 21	

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Keenan et al 2008 [45] USA	<p>Design Cross-sectional</p> <p>Aim Examine the validity of the most popular reading comprehension measures used in research and clinical practice in USA: GORT, WJPC and PIAT</p> <p>Setting 27 different school districts in Colorado</p> <p>Population 510 children (n=180 identical twins, n=290 fraternal, n=40 siblings)</p> <p>Inclusion criteria English as first language, no uncorrected sensory deficits and Full-Scale IQ greater than 85 measured by WISC</p> <p>Exclusion criteria Not reported</p>	<p>Test GORT</p> <p>Participants n=510 Age: 8–18 years</p> <p>Drop-out rate Not reported</p>	<p>Reference test PIAT WJPC</p>	<p>Intercorrelations among the reading comprehension tests GORT-PIAT: 0.51 GORT-WJPC: 0.54 WJPC-PIAT: 0.70</p>	<p>Moderate (validity)</p> <p>Other comments Twins and siblings in the study. Modest inter correlations among the tests suggested that they were measuring partly different skills. If the clinician used WJPC and PIAT to test the child's comprehension skill, a child with only poor decoding skills would appear as if he or she also had poor comprehension skills</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Keenan et al 2012 [46] USA	<p>Design Cross-sectional</p> <p>Aim Examine the implications of test differences for defining and diagnosing comprehension deficit using reading comprehension tests</p> <p>Setting Children were recruited for a behavioral genetic study of comprehension skills as part of the Colorado Learning Disabilities Research Center [185]</p> <p>Population 995 children (n=888 twins, n=107 siblings) 8 not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test PIAT</p> <p>Participants n=995 Age: 8–18</p> <p>Drop-out rate Not reported</p>	Reference test GORT and WJPC	<p>Correlations between pair of reading comprehension tests</p> <p>Bivariate correlations (current sample n=995) PIAT-WJPC: 0.68 PIAT-GORT: 0.54 WJPC-GORT: 0.53</p>	<p>Moderate (validity)</p> <p>Other comments The study use a cut-off of the lowest 100 score (lowest 10% of the sample) so that they compared the same sample size across all tests</p>
Kilpatrick 2012 [47] USA	<p>Design Cross sectional study</p> <p>Aim Examines the relationship between 3 phonological awareness subtests taken from the CTOPP (elision, segmentation and blending) within an unselected population of first grade and second grade students</p> <p>Population First grade (n=67) and second grade (n=49) students</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test CTOPP</p> <p>Participants Grade 1 (n=67) and Grade 2 (n=49)</p> <p>Drop-out rate Not reported</p>	Reference test None	<p>Intercorrelations for CTOPP subtest, presented separately for Grade 1 (n=67) and Grade 2 (n=49)</p> <p>Grade 1: Elision – Segmentation: 0.2 (ns) Elision – Blending words: 0.29 (p<0.05) Segmentation – Blending: 0.35 (p<0.05)</p> <p>Grade 2: Elision – Segmentation: 0.26 (p<0.05) Elision – Blending words: 0.47 (p<0.001) Segmentation – Blending: 0.55 (p<0.001)</p>	<p>Low (reliability)</p> <p>Other comments Correlations tend to be higher in second grade compared with first grade</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Mather et al 1991 [48] USA	<p>Design Cross-sectional</p> <p>Aim Investigate the construct and concurrent validity of WJ-R</p> <p>Setting Private school in the metropolitan Washington DC area</p> <p>Population 47 students with severe learning disabilities from grade 6–8</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test WJ-R</p> <p>Participants n=47 14 girls, 33 boys Age:10–14 years</p> <p>Drop-out rate Not reported</p>	Reference test WRAT-R	Correlation among the WJ-R Spelling and WRAT-R Spelling for the students with learning disabilities WJ-R–WRAT-R: 0.71 (p<0.001)	Low (validity)
McBride et al 2010 [38] USA	<p>Design Prospective cohort (K–2)</p> <p>Aim Examine technical adequacy and information/cost return for 4 early reading measures; DIBELS, SEL, GRADE and TPRI</p> <p>Setting 7 states in USA</p> <p>Population Including 200 students per grade</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test DIBELS</p> <p>Participants n=200 children per grade</p> <p>Drop-out rate Not reported</p>	Reference test None	<p>DIBELS Retest reliability</p> <p>Kindergarten Initial sound fluency: 0.66 Letter naming fluency: 0.86 Word usage: 0.75</p> <p>Grade 1 Phoneme segmentation: 0.57 Nonsense word: 0.85 Word usage: 0.69</p> <p>Grade 2 Word usage: 0.69 Oral reading: 0–89 Retell: 0.75</p>	<p>Moderate (reliability)</p> <p>Risk of bias affected by Limited information regarding administration and scoring</p>

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Prewett et al 1991 [49] USA	<p>Design Cross-sectional</p> <p>Aim Examine the relationship among the reading subtest of WJ-R, PIAT-R, K-TEA and WRAT-R</p> <p>Setting Suburban school district in central Ohio, 1989–1990</p> <p>Population 118 elementary school students (n=1 Kindergarten, n=25 first grade, n=16 second grade, n=33 third grade, n=18 fourth grade, n=25 fifth grade)</p> <p>Inclusion criteria Students included were those students referred to the school psychologist for psychoeducational evaluation because of their lack of academic progress</p> <p>Exclusion criteria Not reported</p>	<p>Test WJ-R</p> <p>Participants n=118 49 girls, 69 boys Age: 6–12</p> <p>Drop-out rate Not reported</p>	Reference test PIAT-R	<p>Pearson correlation between the reading subtests of the WJ-R, PIAT-R and WRAT (n=118)</p> <p>Word decoding WJ-R–PIAT-R: 0.87</p> <p>Reading comprehension WJ-R–PIAT-R: 0.82</p>	Low (validity)

The table continues on the next page

Table 17.2 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Index test Participants Drop-out rate	Reference standard	Accuracy/correlation	Risk of bias Comments
Shull-Senn et al 1995 [50] USA	<p>Design Longitudinal study (first/third/fifth grade)</p> <p>Aim Examine a the stability reliability for the WJ-R</p> <p>Setting Public schools in Midwest USA</p> <p>Population 60 elementary age students</p> <p>Inclusion criteria Not reported</p> <p>Exclusion criteria Not reported</p>	<p>Test WJ-R</p> <p>Participants n=20 students at each grade level 10 girls , 10 boys</p> <p>Drop-out rate None</p>	Reference test None	<p>Reliability, test-retest of WJ-R subtests for first to fifth grade Letter word identification: 0.86–0.96 (p<0.01) Passage comprehension: 0.79–0.95 (p<0.01) Broad reading cluster: 0.90–0.97 (p<0.01) Dictation: 0.35–0.94 (p<0.01) Writing skills: 0.83–0.87 (p<0.01) Broad written language cluster: 0.65–0.93 (p<0.01)</p>	Moderate (reliability)

BWLC = Broad Written Language Cluster; CIBS-II = The Brigance Comprehensive Inventory of Basic Skills-II; CTOPP = Comprehensive Test of Phonological Processing; DAB-2 = Diagnostic Achievement Battery-2; DIBELS = Dynamic Indicators of Basic Early Literacy Skills; DORF = DIBELS Oral Reading Fluency; GORT = Gray Oral Reading Test; GRADE = Group Reading Assessment and Diagnostic Evaluation; K = Kindergarten; KRES = Kindergarten Reading Engagement Scale; N/A = not applicable; NPP = Negative Predictive Power; ns = not significant; NW = Nonsense Word; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PACom = Phonological Awareness Composite; WISC = Wechsler Intelligence Scale for Children; PC = Passage comprehension; PIAT = Peabody Individual Achievement Test; PIAT-R = Peabody Individual Achievement Test-Revised; PPP = Positive Predictive Power; SEL = Star Early Literacy; TOPA = Test of Phonological Awareness; TOWRE = Test of Word Reading Efficiency; WA = Word Attack; WIAT = Wechsler Individual Achievement Test; WID = Word Identification; WISC = Wechsler Intelligence Scale for Children; WJ-III = Woodcock-Johnson Test of Achievement Third Edition; WJPC = Woodcock-Johnson Passage Comprehension subtest; WJ-R = Woodcock-Johnson Psycho Educational Battery-Revised; WPC = WJ-III Passage Comprehension subtest; WRAT-3 = Wide-Range Achievement Test-Third Edition Reading subtest; WRAT-R = Wide Range Achievement Test-Revised; WRMT-R = Woodcock Reading Mastery Test-Revised; WWA = WJ-III Word Attack subtest; WWI = WJ-III Word Identification subtest

Table 17.3. Intervention studies of high or moderate quality used for results and conclusions in the present report.

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Bachman et al 2004 [91] USA	<p>Design Cluster RCT with 2 cohorts</p> <p>Aim Evaluate an intervention for students with poor word-level reading skills, and monitor long term gains on reading skills</p> <p>Setting 11 schools in 4 school districts in upstate New York. Socio-economic status from poor urban to middle class suburban</p> <p>Population n=723 children in 2nd and 3rd grade (mean age 7.88 years), deemed by teachers to be among 20% poorest readers were invited n=295 accepted. Inclusion criteria: standard score <90 on the word attack test WRMT-R, and Verbal IQ >80 n=100 eligible (65% male); 11 boys were eliminated with respect to the neuroimaging component of the study</p> <p>Follow-up 1 year</p>	<p>Intervention Each lesson built on a 5 step plan. Included explicit and systematic instruction to help children develop an understanding of the phonologic and orthographic connections in words and many opportunities to read texts that were phonetically controlled and texts that were not phonetically controlled</p> <p>Intensity 50 minutes of one-to-one tutoring, 5 days per week from end of September to beginning of June (8 months); on average 126 tutoring sessions or 105 hours of tutoring</p> <p>Participants n=37</p> <p>Drop-out rate 5 children (various reasons) between randomisation and training</p> <p>Fidelity 96% of the observed lessons included all 5 of the required steps, with 100% inter-rater agreement</p>	<p>Comparison group: Remedial reading instruction as usual</p> <p>Intensity N/A</p> <p>Participants n=32 students</p> <p>Drop-out rate: n=9</p> <p>Fidelity N/A</p>	<p>Word reading (accuracy) <i>Words</i> Post-test: WRMT Word ID: Cohen's d=1.31 p=0.0001, GORT III accuracy: Cohen's d=0.72 p=0.0041 Follow-up: WRMT Word ID: Cohen's d=1.05 p=0.0001, GORT III accuracy: Cohen's d=0.30 p=0.215</p> <p><i>Non-words</i> Post-test: WRMT Word attack: Cohen's d=0.89 p=0.0006 Follow-up: WRMT Word attack: Cohen's d=0.56 p=0.0243</p> <p><i>Text reading fluency</i> Post-test: GORT III rate: Cohen's d=0.96 p=0.0002 Follow-up: GORT rate: Cohen's d=0.81 p=0.0014</p> <p><i>Spelling</i> Post-test: WRAT spelling: Cohen's d =1.13 p=0.0001 Follow-up: WRAT spelling: Cohen's d=0.81 p=0.0015</p> <p><i>Reading comprehension</i> Post-test: GORT Comprehension, Cohen's d=0.55 p=0.027 Follow-up: GORT Comprehension, Cohen's d=0.24 p=0.32</p>	<p>Moderate</p> <p>Risk of bias affected by The randomisation was highly constrained by a parallel research project on neuroimaging. 2 children were reassigned from the control group to the intervention group</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Bhat et al 2003 [97] USA	<p>Design Quasi-experimental (within-group repeated-measure design). The instructional procedure in both groups was exactly the same, except that instruction for I2 began after instruction for I1 had been completed (mid-test)</p> <p>Aim Investigate if school students with learning disabilities could improve their phonological awareness skills after instruction, and if these skills could impact word recognition skills</p> <p>Setting Middle school students with learning disabilities</p> <p>Population (=participants) n=40 (20 girls, 20 boys) Average age: 12.90 years old</p> <p>Inclusion criteria Students (n=50) with learning disabilities were screened using the Lindamood Auditory Conceptualisation Test. Teachers identified students who had poor attendance records or who, for any other reason, might be able to complete the instructional process. These students were eliminated from the pool of potential participants, leaving 40 students</p> <p>Follow-up None</p>	<p>Intervention I1: Training group (intervention between pre-test and mid-test)</p> <p>Participants I1: n=20 (8 girls, 12 boys) Mean age 12.90 years</p> <p>Intensity Instructions in phonological awareness were provided on a one-on-one basis, during 4 weeks. 3 days per week, and they were taught 2 lessons per day. Total number of lessons for both groups was 18</p> <p>Drop-out rate None</p> <p>Fidelity Not reported</p>	<p>Comparison group I2: Waiting list; no intervention between pre-test and mid-test; intervention between mid-test and post-test</p> <p>Participants I2: n=20 (12 girls, 8 boys) Mean age: 12.90 years</p> <p>Intensity See I1</p> <p>Drop-out rate None</p> <p>Fidelity Not reported</p>	<p>Phonological awareness (CTOPP) Test scores for group I1 <i>Pre-test (mean)</i> 67.35 <i>Mid-test (mean)</i> 94.65 <i>Post-test (mean)</i> 94.70 Test scores for group I2 <i>Pre-test (mean)</i> 63.75 <i>Mid-test (mean)</i> 67.85 <i>Post-test (mean)</i> 96.65</p> <p>Word decoding (WRMT-R) Test scores for group 1 <i>Pre-test (mean)</i> 41.35 <i>Mid-test (mean)</i> 43.20 <i>Post-test (mean)</i> 44.95 Test scores for group I2 <i>Pre-test (mean)</i> 40.45 <i>Mid-test (mean)</i> 41.4 (7.21) <i>Post-test (mean)</i> 42.7 (6.57)</p>	<p>Moderate</p> <p>Comments Training groups did not differ on pre-test measures</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Calhoon et al 2005 [89] USA	<p>Design RCT, randomised at teacher level</p> <p>Aims Examine the effect of a peer-mediated instructional approach for students with reading disabilities</p> <p>Setting Special education, self-contained language arts classrooms for students with reading disability in 2 middle schools in 1 school district</p> <p>Population: Teachers: n=4, self-selected Students: n=38 (32 sixth, 5 seventh, and 1 eighth grade; mean age approximately 12 years (66% boys). Reading at least 3 grade levels below their current grade placement</p> <p>Follow-up Post-test only</p>	<p>Intervention Linguistics skills training + Peer assisted learning strategies:</p> <ol style="list-style-type: none"> Mediated verbal rehearsal which models the procedural steps for verbal rehearsal Step-by-step feedback Frequent verbal and written interaction between tutor and tutee Reciprocity (i.e. each student serves in the role of tutor and tutee during each session) <p>Intensity 3 days a week, in total 51 hours</p> <p>Participants n=2 teachers n=18 students</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Comparison group: Saxon Phonics reading instruction but no peer-mediated reading activities. Structure similar to the intervention</p> <p>Intensity 3 days a week, in total 51 hours</p> <p>Participants n=2 teachers n=20 students</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Word reading (accuracy) <i>Words</i> WJ-III Letter-word identification post score: Effect size (Cohen's d) growth: 1.10, p<0.05</p> <p><i>Non-words</i> WJ-III Word attack Effect size (Cohen's d) growth: 0.99, p<0.05</p> <p>Text reading fluency Post-test: WJ-III Reading fluency Effect size (Cohen's d) growth: -0.19/ns</p> <p>Reading comprehension Post-test: WJ-II Passage comprehension Effect size (Cohen's d) growth: 0.94, p<0.05</p>	<p>Moderate</p> <p>Risk of bias affected by Fidelity not sufficiently monitored. Possible teacher effects due to teacher randomisation. Statistical analyses appear not to have been adjusted for group level randomisation</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Cogo- Moreira et al 2012 [102] Brazil	<p>Design Systematic review</p> <p>Aim To study the effectiveness of music education on reading skills in children and adolescents with dyslexia</p> <p>Setting Public and private schools</p> <p>Population Children and adolescents with dyslexia</p> <p>Inclusion criteria RCT (including quasi-randomised or cluster-randomised trials). At least 1 of our primary outcomes (oral reading skills, reading comprehension, reading fluency, phonological awareness, and spelling, measured through validated instruments*). Secondary outcomes were self-esteem and improved academic performance</p> <p>* Valid instruments for example: GORT, WRMT, CTOPP, TORF, NRST and CELF</p> <p>Follow-up Short term (up to 6 months) Medium term (between 6 and 12 months) and long term (more than 12 months)</p>	<p>Intervention Any approach of music education** compared with a waiting list or no training control group.</p> <p>** Individual or group music lessons or musical training with a music advisor or teacher at music school or at the school where the children are receiving their formal instruction, either as part of the general curriculum or as addition tuition. Children may be exposed to song and tonal or atonal and rhythmic patterns in a natural musical setting or may be exposed to a specific musical methodological approach. They may be encouraged to practice music in small or large groups</p> <p>Intensity N/A</p> <p>Participants N/A</p> <p>Drop-out rate N/A</p> <p>Fidelity N/A</p>	<p>Comparison group N/A</p> <p>Intensity N/A</p> <p>Participants N/A</p> <p>Drop-out rate N/A</p> <p>Fidelity N/A</p>	No studies were included in this review, since no studies fulfilled the inclusion criteria	Low

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Das et al 2008 [71] (study 1) Canada	<p>Design RCT</p> <p>Aim: Examine the effectiveness of 2 reading intervention programs for poor readers</p> <p>Setting A reservation school in Western Canada</p> <p>Population: Mean age: approximately 9.2 years</p> <p>Inclusion criteria >1 standard deviation or more below the mean on at least 1 of WJ-III Word identification or WJ-III Word attack (non-words)</p> <p>Follow-up: N/A</p>	<p>Intervention Pass reading enhancement program: Small group training (3–4 students). 4 successive processing modules and 4 simultaneous processing modules, each involving a global and curriculum-related bridging component</p> <p>Intensity Twice weekly for a total of 20 sessions. Approximately 13.5 hours of instruction was delivered in total from December–May</p> <p>Participants n=21 13 girls, 8 boys</p> <p>Drop-out rate Not reported</p> <p>Fidelity High training integrity (95%, combined for intervention and comparison groups)</p> <p>Average student attendance: 11.5 hours</p>	<p>Comparison group: Phonics based intervention in small groups (3–4 students)</p> <p>Intensity Twice weekly for a total of 20 sessions. Approximately 13.5 hours of instruction was delivered in total from December–May</p> <p>Participants n=21 10 girls, 11 boys</p> <p>Drop-out rate Not reported</p> <p>Fidelity High training integrity (95%, combined for intervention and comparison groups)</p> <p>Average student attendance: 11.5 hours</p>	<p>Word reading (accuracy) Words Post-test: WJ-III Word identification: Cohen's d=0.2, p=ns</p> <p>Non-words Post-test: WJ-III Word attack Cohen's d=0.006, p=ns</p> <p>Phonemic awareness Post-test: CTOPP Elision Cohen's d=0.07, p=ns CTOPP Segmenting Cohen's d=0.59*, p=ns. When corrected for pre-test differences</p> <p>* Not corrected for pre-test differences</p>	<p>Moderate</p> <p>Risk of bias affected by Fidelity problems. Only the first part of study (study 1) could be evaluated due to inclusion criteria</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
del Rosario Ortiz et al 2002 [77] Spain	<p>Design RCT</p> <p>Aims: Examine the effect of adding speech perception training to interventions for reading skills in children with reading disability</p> <p>Setting Third, fourth, and fifth grades of state primary schools</p> <p>Population: n=53 with specific reading difficulties according to their teachers. Mean age: 9.8 years</p> <p>Inclusion criteria IQ >90 Pseudo-word reading performance <15th percentile on the PROLEC reading test was inclusion criteria</p> <p>Follow-up: Post-test only</p>	<p>Intervention I1: Speech perception training + letter sound and phonological awareness (SP/LPA). Intensive systematic instruction in letter sounds and phonological awareness + speech perception training I2: Letter sound and phonological awareness (LPA): Emphasised the acquisition of individual letter-sound correspondences and phonological awareness skills. Each session consisted of 5 tasks. Delivered by post graduate students in groups of 4 children</p> <p>Intensity 20 sessions, each 20 minutes during 4 weeks</p> <p>Participants n=17 (SP/LPA) n=18 (LPA)</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not specified: "met every week to review the training integrity"</p>	<p>Comparison group: C: Curriculum as usual</p> <p>Intensity N/A</p> <p>Participants n=18</p> <p>Drop-out rate Not reported</p> <p>Fidelity N/A</p>	<p>Word reading (accuracy) <i>Words/Non-words</i> I1: 55.05 (2.46) I2: 54.04 (3.5) C: 51.37 (7.82) Word reading (PROLEC; word and pseudo word reading) Effect size reading: 0.252; I1 >C</p> <p>Phonemic awareness I1: 56.41 (3.46) I2: 57.29 (3.4) C: 45.37 (12.87) Effect size phonemic awareness: 0.258; I1=I2>C</p>	<p>Moderate</p> <p>Risk of bias affected by Fidelity not sufficiently monitored. Unclear use of certain statistical terms, including type of effect size</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Fälth et al 2013 [78] Sweden	<p>Design RCT</p> <p>Aim Examine 3 computerised interventions strategies for children with reading disabilities: "bottom"-up training focusing on decoding/phonology, "top down"-training focusing on word and sentence processing and a combination of the both</p> <p>Setting 2nd grade in 41 Swedish schools, general educational settings. No school drop-outs after randomisation</p> <p>Population: n=100 (30 girls, 70 boys). Mean age: not reported, but approximately 8 years</p> <p>Inclusion criteria: Special instruction in reading in grade 2, at least 0.75 standard deviation below the mean of the typical readers in word reading</p> <p>Exclusion criteria Recent immigrants or gross neurological disturbances or sensory deficits</p> <p>Follow-up: Post-test, follow-up (1 year)</p>	<p>Intervention I1: Comprehension training program [186] I2: Phonological training program [187] I3: I1+I2</p> <p>Intensity 25 training sessions one-to-one Most sessions lasted 15–25 min</p> <p>Participants I1: n =25 children I2: n=25 children I3: n=25 children</p> <p>Drop-out rate Low</p> <p>Fidelity Unclear. The authors monitored training fidelity by personal meetings, e-mail correspondence and phone calls</p>	<p>Comparison group: C: Special instructions as usual</p> <p>Intensity N/A</p> <p>Participants n=25 children</p> <p>Drop-out rate Low</p> <p>Fidelity N/A</p>	<p>Word reading (speed) - words (Sight word reading) Mean (standard deviation) Post-test I1: 23.69 (8.81) I2: 25.15 (8.77) I3: 23.84 (7.97) C: 24.33 (10.22)</p> <p>Follow-up 2 I1: 32.90 (13.06) I2: 34.24 (11.40) I3: 37.89 (8.64) C: 28.92 (12.24)</p> <p>Word reading (speed) - Non-words (Non-word reading) Post-test I1: 12.52 (4.72) I2: 12.95 (3.80) I3: 14.96 (5.84) C: 13.04 (5.60)</p> <p>Follow-up 2 I1: 15.95 (6.33) I2: 16.50 (6.06) I3: 20.65 (6.77) C: 15.26 (6.59)</p> <p>Reading comprehension Post-test I1: 12.24 (4.86) I2: 16.33 (4.66) I3: 14.43 (5.76) C: 12.30 (4.57)</p> <p>Follow-up 2 I1: 17.67 (4.99) I2: 21.59 (7.29) I3: 20.48 (6.89) C: 18.26 (6.53)</p>	<p>Moderate</p> <p>Comments The randomisation is not clearly described but was constrained with the aim of "monitoring group differences". Nonetheless, there were substantial pre-test differences on some tests</p>

Results continues on the next page

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>				Phonemic awareness (segment subtraction) Post-test I1: 4.80 (2.36) I2: 7.00 (3.31) I3: 6.04 (3.00) C: 6.42 (2.93) Follow-up 2 I1: 6.08 (2.53) I2: 8.09 (2.68) I3: 7.96 (2.88) C: 6.87 (2.46)	
Fälth et al 2013 [78] Sweden					
Gelzheiser et al 2011 [98] USA	Design Quasi-experimental (respon- sive to intervention) Aim Investigate the efficacy of the Interactive Strategies Approach- Extended (ISA-X), which was provided by public school teachers in a daily, one-to-one format to grade 4 struggling readers with individualised education programs Setting Grade 4 students from Albany, New York area Population (=participants) n=50 Inclusion criteria IQ above 80 Follow-up None	Intervention I1: Interactive Strategies Approach- Extended (ISA-X) Participants I1: n=26 8 girls, 18 boys Intensity An initial workshop for 5 days, a 3-day refresher workshop, and semimonthly 2-hours after-school meetings Drop-out rate None Fidelity Not reported	Comparison group I2: Waiting list (general reading instructions in classroom) Participants I2: n=22 5 girls, 17 boys Intensity Reading instructions in the general classroom, called "business-as- usual" Drop-out rate n=2 Fidelity Attendance was highly variable, and 25% of the sample was absent from instruction once every 2 weeks, or 10% of the time	Effect sizes for pre- to post-semester changes during the intervention and nonintervention Word reading (accuracy) <i>Words</i> I1: 1.28 I2: – Reading fluency I1: 0.19 I2: 0.20 Reading comprehension I1: 1.68 I2: –	Moderate Comments Half the students received the intervention in the fall semester while the others served as a waiting-list control group who received the intervention in the spring

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Gustafson et al 2000 [95] Sweden	<p>Design RCT, randomised at school level</p> <p>Aims Examine the effect of phonological awareness training for children with dyslexia</p> <p>Setting 14 schools with middle class children, from 1 medium sized city in Sweden</p> <p>Population: n=48 children in 4th grade Age: 10–11 years</p> <p>Follow-up: Post-test only</p>	<p>Intervention I: Phonological awareness training in small groups (2–4 children) consisting of 7 types of exercises (without written letters) delivered by teachers</p> <p>Intensity Approx 760 hours of training February–May and September–December + about 40 sessions 20 minutes each of "regular special education"</p> <p>Participants n=33 children</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Comparison group: C: Special education as usual</p> <p>Intensity About 40 sessions, 20 minutes each</p> <p>Participants n=16</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>The Intervention group made more progress in phonological awareness ($p < 0.01$ in an overall statistical analysis), but no significant effect on reading index</p> <p>Word reading (accuracy) Words Raw mean scores on a Reading index Post-test I: 31.8 (10.6) C: 30.8 (8.9)</p> <p>Phonemic awareness Raw mean scores on UMESOL Segment subtraction Post-test: I: 9.9 (2.3) C: 9.2 (2.8)</p>	<p>Moderate</p> <p>Risk of bias affected by Fidelity not sufficiently monitored. Statistical analyses appear not to have been adjusted for school level randomisation</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Hatcher et al 2006 [94] United Kingdom	<p>Design RCT</p> <p>Aim Examine the effectiveness of an intervention for reading-delayed children</p> <p>Setting 16 primary schools; 2 were excluded due to small numbers of poor readers</p> <p>Population n=635 children in first grade screened from 16 schools using group spelling test and teacher nomination. The 6 children with lowest reading skills from each of the 14 schools were included. Mean age: approximately 5.6 years</p> <p>Exclusion n=9 (3 children due to severe behavior problems, 4 due to very low cognitive ability, and 2 due to general poor school attendance)</p> <p>Follow-up 69/76 children were retested approximately 11 months later. However, as the children in general continued to have the same assistants as those who delivered the intervention, this part is not considered to be of sufficient quality for consideration here</p>	<p>Intervention I: Individual training in small groups (n=3) with letter-sound and phoneme identification activities</p> <p>Intensity Daily 20 minutes lessons for 10 weeks, delivered by trained teaching assistants, in total 16.5 hours</p> <p>Participants n=39 children</p> <p>Drop-out rate Following randomisation 1 school withdrew, i.e. n=13 participated. No drop-out from intervention</p> <p>Fidelity High</p>	<p>Comparison group C: Waiting list for 10 weeks</p> <p>Intensity N/A for the first 10 weeks</p> <p>Participants n=38</p> <p>Drop-out rate n=1</p> <p>Fidelity Teaching assistants supported by investigators</p>	<p>Word reading (accuracy) at 10 weeks Words Post-test: Mean (standard deviation) I: 12.49 (7.4) C: 11.11 (7.82) EWR: Cohen's d=0.79, p<0.001</p> <p>Phonemic awareness Post-test: Mean (standard deviation) I: 4.59 (3.42) C: 3.37 (3.53) Phonological Awareness TEST: Phoneme completion Cohen's d=0.46, p=0.044</p>	<p>Moderate</p> <p>Risk of bias affected by Statistical analyses corrected for individual RCT at each school, and corrected for pre-test differences between study groups</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Kairaluoma et al 2009 [188] Finland	<p>Design RCT</p> <p>Aim To test the effects on reading of dietary supplementation of unsaturated fatty acids in children with dyslexia</p> <p>Setting Grades 3–6 in 15 schools in central Finland</p> <p>Population n=61 26 girls, 35 boys Children with observed reading difficulties. Screened by teachers with a reading test, Lukilasse [189]. Mean age: 10.5 years</p> <p>Inclusion criteria Score ≥ 4 standard points below age. IQ >80</p> <p>Exclusion criteria Neurological or psychiatric disorder</p> <p>Follow-up Post-test only</p>	<p>Intervention Ethyl-EPA (500 mg/day)/carnosine (400 mg/day) per orally</p> <p>Intensity Daily for 90 days</p> <p>Participants n=30 children 12 girls, 18 boys</p> <p>Pre-test ANCOVAS was performed</p> <p>Drop-out rate None</p> <p>Fidelity High</p>	<p>Comparison group Placebo pills, triglycerides and cellulose orally</p> <p>Intensity Daily for 90 days</p> <p>Participants n=31 children 14 girls, 17 boys</p> <p>Pre-test ANCOVAS were performed</p> <p>Drop-out rate None</p> <p>Fidelity High</p>	<p>Post tested with ANCOVAS. No significant results for: Word reading (accuracy) word and word reading (accuracy) non-words, word reading (speed) words, word reading (speed) non-words, text reading fluency, spelling and phonemic awareness</p>	<p>Moderate</p> <p>Risk of bias affected by Method of randomisation is not specified</p> <p>Other comments Follow-up N/A</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Kast et al 2007 [92] Switzerland	<p>Design RCT, matched groups, recruited by advertisement in newspapers</p> <p>Aim To test whether computer based multisensory learning help to mitigate writing errors in children with development dyslexia</p> <p>Setting Home</p> <p>Population n=43, 15 girls, 28 boys Age: 9–11 years (mean age: 10.3 years)</p> <p>Inclusion criteria Developmental dyslexia IQ >85. Native German speaking</p> <p>Exclusion criteria ADHD Planning problem IQ <85 Germany as secondary language</p> <p>Follow-up Post-test</p>	<p>Intervention I: Dybuster, a computer-based writing training. The basic principle is to learn associations between the visual and auditory stimuli</p> <p>Intensity 15–20 minutes 4 days a week for 3 months then 3 months without training</p> <p>Participants n=20 children</p> <p>Drop-out rate n=2</p> <p>Fidelity N/A</p>	<p>Comparison group C: Waiting group</p> <p>Intensity 3 months</p> <p>Participants n=23 children</p> <p>Drop-out rate Not reported</p> <p>Fidelity N/A</p>	<p>Spelling Post-test: Dybuster, 100 words (50 learned, 50 not-learned); Mean (standard deviation). No of errors. I: 45 (18) C: 53 (16) Cohen's d=0.47</p>	<p>Moderate</p> <p>Risk of bias affected by Method of randomisation is not specified</p> <p>Other comments The authors have produced the program</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Kerstholt et al 1994 [79] (Study 1) The Netherlands	<p>Design RCT</p> <p>Aim Effects of different types of phonemic segmentation on phonemic segmentation, reading and spelling. Has visual support an additional value in phonemic segmentation?</p> <p>Setting 7 schools for children with learning disabilities</p> <p>Population n=132 After screening in schools for children with learning disabilities (letter knowledge, phonemic segmentation skill, reading, spelling, non-verbal intelligence) n=50, 15 girls, 35 boys Age: 7–10 years (mean: 8.4 years)</p> <p>Inclusion criteria: Poor phonemic segmentation</p> <p>Follow-up Post-test</p>	<p>Intervention Computerised training I1: Structure training: phonemic segmentation + structure support, correct grapheme displayed. Alphabet and diagram I2: Memory training: phonemic segmentation + memory support. Auditory signal as error warning. Alphabet only</p> <p>Intensity 15–20 minutes, 2 times per week during 5 weeks</p> <p>Participants I1: n=16 children I2: n=16 children</p> <p>Drop-out rate: n=2</p> <p>Fidelity Adequate</p>	<p>Comparison group C: Phonemic segmentation only; computerised</p> <p>Intensity 15–20 minutes, 2 times per week during 5 weeks</p> <p>Participants n=16 children</p> <p>Drop-out rate: None</p> <p>Fidelity N/A</p>	<p>Word reading (accuracy) <i>Words</i> AVI word cards' system; Reading words, mixed measure, speed and accuracy, max 32; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 20.69 (12.33) I2: 23.56 (7.79) C: 26.25 (9.27)</p> <p>Word reading (speed) <i>Words</i> Post-test: See Word accuracy</p> <p>Spelling The CITO-test, max 36; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 17.50 (10.75) I2: 20.63 (9.13) C: 24.50 (7.36)</p> <p>Phonemic awareness Phonemic segmentation, researcher designed, max 36; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 23.19 (9.72) I2: 25.50 (9.83) C: 27.75 (7.36)</p>	<p>Moderate</p> <p>Risk of bias affected by Method of randomisation not clear</p> <p>Other comments Do not take initial scores into account. No separate measures of reading speed and accuracy</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Kerstholt et al 1994 [79] (Study 2) The Netherlands	<p>Design RCT, stratification based on phonemic segmentation, reading, spelling, non-verbal IQ</p> <p>Aim Effects of oral feedback and different types of phonemic segmentation effect on phonemic segmentation, reading and spelling</p> <p>Setting 4 schools for children with learning disabilities</p> <p>Population n=136 Age: 7–10.4 years (mean age: 8 years 6 months)</p> <p>Inclusion criteria Poor phonemic segmentation (30% error rate, at least 14 months behind), after 24 months of reading instruction. A year and a half behind in reading and spelling</p> <p>Follow-up Post-test only</p>	<p>Intervention Training on phonemic segmentation, reading and spelling and oral feedback</p> <p>I1. Structure training Visual (no of segments) + corrective oral feedback. Correct grapheme was displayed regardless of child's response</p> <p>I2. Memory training Visual + corrective oral feedback</p> <p>Intensity 15–20 minutes, 2 times per week during 5 weeks, total 10 sessions</p> <p>Participants I1: n=17 children I2: n=16 children</p> <p>Drop-out rate n=2</p> <p>Fidelity Adequate</p>	<p>Comparison group Training as usual + oral feed back</p> <p>Intensity 15–20 minutes, 2 times per week during 5 weeks, total 10 sessions</p> <p>Participants n=16 children</p> <p>Drop-out rate None</p> <p>Fidelity N/A</p>	<p>Word reading (accuracy) Words AVI word cards' system; Reading words, mixed measure, speed and accuracy, max 32; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 21.35 (9.39) I2: 20.69 (10.56) C: 23.88 (9.05)</p> <p>Word reading (speed) Words Post-test: See Word accuracy</p> <p>Spelling The CITO-test, max 36; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 20.94 (10.84) I2: 22.06 (7.66) C: 22.31 (9.33)</p> <p>Phonemic awareness Phonemic segmentation, researcher designed, max 36; Mean (standard deviation) Post-test: ns (according to ANOVA) I1: 23.82 (11.81) I2: 24.50 (10.03) C: 25.31 (10.05)</p>	<p>Moderate</p> <p>Risk of bias affected by Method of randomisation not clear</p> <p>Follow-up N/A</p> <p>Other comments Do not take initial scores into account. No separate measures of reading speed and accuracy</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Lovett et al 2000 [190] Canada	<p>Design RCT</p> <p>Aim Examine the heuristic value of the double-deficit hypothesis from 2 perspectives; a) to understand the core processing deficits underlying severe form of developmental reading disability b) to assess the amenability of these deficits to treatment efforts</p> <p>Setting Special laboratory classroom at a pediatric teaching hospital or in affiliated satellite schools</p> <p>Population n=166, 53 girls, 113 boys Age 7–13 years n=140 after a pre-training diagnostic testing</p> <p>Inclusion criteria Reading disability. Substantial underachievement (<20th percentile) replicated on 4 out of 5 standardised reading achievement measures. A phonological deficit and or a visual naming-speed deficit. English as first language. Average intelligence</p> <p>Follow-up None</p>	<p>Intervention I1: Phonological analysis and blending/direct instruction (PHAB/DI) I2: Word identification strategy training (WIST)</p> <p>Intensity 35 hours: 1 hour/day, 4 days a week</p> <p>Participants I1: n=52 Double deficit: 27 Phonological deficit: 12 Visual naming speed: 13 I2: n=53 Double deficit: 32 Phonological deficit: 8 Visual naming speed: 13</p> <p>Drop-out rate Not reported</p> <p>Fidelity Probably adequate</p>	<p>Comparison group C: Training in academic survival skills and self-help strategies</p> <p>Intensity 35 hours: 1 hour/day, 4 days a week</p> <p>Participants n=35 Double deficit: 17 Phonological deficit: 11 Visual naming speed: 7</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Word reading (accuracy) Words (WRMT-R Word Identification) Post-test Double deficit group I1: 65.9 (12.2) I2: 58.3 (10.4) C: 65.5 (16.6)</p> <p>Post-test Phonological deficit group I1: 68.8 (11.5) I2: 69.6 (15.3) C: 74.5 (7.9)</p> <p>Post-test Visual naming speed group I1: 75.9 (10.3) I2: 73.8 (12.1) C: 76.1 (6.6)</p> <p>Non-words (WRMT-R Word Attack) Post-test Double deficit group I1: 71.7 (11.11) I2: 62.0 (13.4) C: 68.0 (14.7)</p> <p>Post-test Phonological deficit group I1: 78.3 (5.5) I2: 76.0 (14.2) C: 66.2 (10.0)</p> <p>Post-test Visual naming speed group I1: 84.1 (9.0) I2: 77.9 (10.3) C: 78.0 (9.2)</p>	Moderate

Results continues on the next page

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>				WRMT-R; Multiple regressions with initial scores. Standardised Beta, t-values and p-values are reported	
Lovett et al 2000 [190] Canada				Post-test: I vs C: -0.201; -4.01; 0.001 WIST vs PHAB/DI: -0.177; -3.35; 0.001 GFW; Standardised Beta, t-values and p-values are reported I vs C: -0.192; -3.35; 0.001 WIST vs PHAB/DI: -0.194; -3.32; 0.001	
Lovett et al 2012 [99] Canada	<p>Design Quasi-experimental</p> <p>Aim Investigate the effect of research-based reading intervention designed for struggling readers</p> <p>Setting High school students from 19 large multicultural and linguistically diverse urban school district in Canada, the Toronto Catholic District School Board</p> <p>Population (=participants) n=351, 164 girls, 187 boys Average age: 14.7 years old</p> <p>Inclusion criteria Overall, their average reading performance was between 1.5 and 2 standard deviations below age expectations at program entry</p> <p>Follow-up 1 year (follow-up data were not available on the control participants, who crossed over into the intervention condition following post-testing at 70 hours)</p>	<p>Intervention I1: PHAST PACES program integrate word identification and text comprehension strategy instruction and phonological awareness</p> <p>Participants I1: n=268</p> <p>Intensity 60–70 hours</p> <p>Drop-out rate Not reported</p> <p>Fidelity High</p>	<p>Comparison group I2: Control</p> <p>Participants I2: n=83</p> <p>Intensity Not reported</p> <p>Drop-out rate See intervention</p> <p>Fidelity High</p>	<p>Effect sizes (Cohen's d) were calculated for fixed effect</p> <p>Word reading Words Post-test: WRMT-R Word Identification Cohen's d=0.15</p> <p>Word reading Non-words Post-test: WRMT-R Word Attack Cohen's d=0.45</p> <p>Reading comprehension Post-test: WRMT-R Passage comprehension Cohen's d=0.34</p>	Moderate

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Macaruso et al 2009 [81] USA	<p>Design RCT, randomised at class room level</p> <p>Aims Evaluate CAI* program as a supplement to Language!**</p> <p>Setting 3 class rooms in a public school in St George, Utah</p> <p>Population n=47. 6th and 7th grade students involved in a remedial reading program</p> <p>Inclusion criteria Low score on Developmental Reading Assessment and teachers referrals</p> <p>Exclusion criteria 5 students omitted due to: 2 students - IQ <70 2 students - brain injury 1 student - autism</p> <p>Follow-up Post-test only</p> <p>* CAI, Computer-Assisted Instruction, designed to build phonological awareness and phonics skills</p> <p>** Language! is a comprehensive programme for teaching language-based literacy</p>	<p>Intervention Language! combined with Read Right*** + CAI (Lexia S.O.S)</p> <p>Intensity 20–30 minutes 2 or 3 times per week from late September 2005 and ended May 2006. Students received on average 50 sessions</p> <p>Participants n=27 students 12 girls, 15 boys</p> <p>Drop-out rate None</p> <p>Fidelity Not reported</p> <p>*** Read Right for fluency development</p>	<p>Comparison group Language! combined with Read Right</p> <p>Participants n=15 students 8 girls, 7 boys</p> <p>Drop-out rate None</p> <p>Fidelity Not reported</p>	<p>All tests are subtests from Woodcock-Johnson III Tests of Achievement</p> <p>Word reading (accuracy) <i>Non-words</i> Post-test: F (1, 39)=5.04; p=0.03 favoring the training group</p> <p>Text reading fluency Post-test: N/A (mixed measure)</p> <p>Spelling Post-test: I: 74.6 (13.6) C: 81.1 (12.3)</p> <p>Reading comprehension Post-test: I: 79.6 (15.3) C: 84.5 (8.8)</p>	<p>Moderate</p> <p>Comments Small sample</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Mavrommati et al 2002 [83] Greece	<p>Design RCT</p> <p>Aim To compare the pictographic teaching method (PICTO) with the traditional teaching method (TRAD)</p> <p>Setting 4th to 6th grades</p> <p>Population n=72, 20 girls, 52 boys age: 9–11 years</p> <p>Inclusion criteria Poor speller, ≥5 positive indicators on Bangor Dyslexia Test, IQ ≥95, Greek as first language, free from psychiatric or emotional problems</p> <p>Follow-up Post-test: 1 month after cessation of teaching (PICTO group: 2+4 weeks; TRAD group: 12+4 weeks)</p>	<p>Intervention I1: PICTO I2: TRAD</p> <p>Intensity PICTO - 10 sessions, 40 minutes each during 2 weeks TRAD - 60 sessions, 40 minutes each during 12 weeks</p> <p>Participants I1: n=2x12 students (trained on different sets of words and method, Set A and Set B) I2: n=2x12 students (trained on different sets of words and method, Set A and Set B)</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Comparison group C: No extra teaching of any kind</p> <p>Intensity Untreated group was given only 1 spelling test of all the 50 words. This was done 1 month after the end of the 14th week, which was the last week of the whole teaching programme</p> <p>Participants n=24</p> <p>Pre-test Mean (standard deviation) C: 8.32 (4.32)</p> <p>Drop-out rate Not reported</p>	<p>Spelling Mean (standard deviation): post-test 2 sets A and B</p> <p>Set A I1: 22.65 (2.42) I2: 3.00 (3.13) C: 1.32 (1.49)</p> <p>Set B I1: 20.83 (3.05) I2: 4.67 (2.44) C: 1.41 (1.76)</p>	<p>Moderate</p> <p>Risk of bias affected by Method of randomisation not specified. Drop-out rate is not reported</p> <p>Other comments The spelling test included the 25 trained words</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
McPhillips et al 2000 [191] United Kingdom	<p>Design RCT, individually matched, double blind, placebo-controlled</p> <p>Aim Examine whether the Asymmetrical tonic neck reflex (ATNR) could be inhibited by stereotypical movements of the primary reflex system and whether reading skills would improve as ATNR persistence decreased</p> <p>Setting Home</p> <p>Population Children from ordinary primary schools, recruited through local dyslexia groups. n=98. Age: 8–11 years. n=66 after screening, but 6 students were not possible to match, n=60</p> <p>Inclusion criteria Reading difficulties, at least 12 months behind, (Salford sentence Reading test, WORD), IQ 85–115, persistent ATNR</p> <p>Exclusion criteria All children excluded at this stage were not sufficiently behind in reading ability or were below average in verbal IQ. Hearing difficulties were excluded by means of a standard audiogram</p> <p>Follow-up Post-test only</p>	<p>Intervention I: A specific movement sequence to repeat each evening</p> <p>Intensity 10 minutes every day for 12 months</p> <p>Participants n=20</p> <p>Drop-out rate Not reported</p> <p>Fidelity Adequate</p>	<p>Comparison groups C1: A set of nonspecific movements C2: No intervention</p> <p>Intensity 10 minutes every day for 12 months</p> <p>Participants C1: n=20 C2: n=20</p> <p>Drop-out rate Not reported</p> <p>Fidelity Adequate</p>	<p>Difference (95% CI) between pre- and post-test</p> <p>Word reading (composite from accuracy and speed) <i>Words</i> I: 15.3 (11.8; 18.8) C1: 7.1 (3.6; 10.5) C2: 6.3 (2.8; 9.8) p<0.001</p> <p><i>NEALE analysis of reading ability</i> I: 19.6 (16.5; 22.6) C1: 7.3 (4.9; 9.7) C2: 6.9 (5.1; 8.6) p<0.001</p> <p>Spelling I: 7.5 (1.8; 13.2) C1: 5.5 (–0.8; 11.7) C2: 6.0 (–0.7; 11.3)</p> <p>Phonemic awareness (spoonerism) I: 3.7 (0.5; 6.9) C1: 2.1 (–1.0; 5.2) C2: 2.7 (–1.0; 6.4)</p>	<p>Moderate</p> <p>Risk of bias affected by Age differences within the sample</p> <p>Other comments Reading measures only reported mixed, by sets of test batteries- and reported as reading age</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Morris et al 2012 [80] USA	<p>Design RCT, stratified by ethnicity, family socio-economic status and IQ</p> <p>Aim To evaluate effects of 3 reading remediation's</p> <p>Settings 2nd and 3rd grade in public schools in 3 large metropolitan areas</p> <p>Population n=958 children with English as primary language, assessed by a screening battery (K-BIT, WRMT-R, WRAT-3). 59% (565 children) fulfilled criteria for reading disabilities. n=279 completed 1 full year of intervention. Age: 6 years 6 months–8 years 8 months</p> <p>Inclusion criteria A Low Achievement (LA) or an Ability-Achievement Regression Corrected Discrepancy (DISC) determination of a reading disability. K-BIT >70 and reading standard score approximately 16 percentile or less. Normal hearing and vision</p> <p>Exclusion criteria Had repeated a grade. K-BIT score <70 or had a serious psychiatric/neurological illness</p> <p>Follow-up 1 year for the intervention groups. The math Classroom survival skills group (CSS) did not participate in the follow-up</p>	<p>Intervention Group reading remediation: I1: Phonological analysis and blending + CSS I2: Phonological analysis and blending + Retrieval automaticity vocabulary I3: Phonological and strategy training</p> <p>Intensity 70 hours: 1 hour/day, 5 days a week (14 weeks)</p> <p>Participants I1: n=69 I2: n=69 I3: n=73</p> <p>Drop-out rate 19,6% according to the corresponding author</p> <p>Fidelity High</p>	<p>Comparison group C: math + CSS</p> <p>Intensity 70 hours: 1 hour/day, 5 days a week (14 weeks)</p> <p>Participants n=68</p> <p>Drop-out rate: 10% left early</p> <p>Fidelity High</p>	<p>^a p<0.10; *p<0.05; **p<0.01; ***p<0.001</p> <p>Word reading (accuracy) Words (F-score, WRMT-R Word ID)</p> <p>Post-test C vs I1, I2, I3: 52.38*** I1 vs I2, I3: 5.97** (in favour of I2, I3) I2 vs I3: ns and values not reported</p> <p>Follow-up I1 vs I2, I3: 10.58** (favour of I2, I3) I2 vs I3: ns and values not reported C: N/A</p> <p>Non-words (F-score, WRMT-R Word Attack)</p> <p>Post-test C vs I1, I2, I3: 68.15*** I1 vs I2, I3: 25.36*** (in favour of I2, I3) I2 vs I3: ns and values not reported</p> <p>Follow-up I1 vs I2, I3: 10.27** (in favour of I2, I3) I2 vs I3: 3.16^a (in favour of I2) C: N/A</p> <p>Word reading (speed) Words (F-score, TOWRE, Word Reading efficiency)</p> <p>Post-test C vs I1, I2, I3: 25.02*** I1 vs I2, I3: ns and values not reported I2 vs I3: ns and values not reported</p> <p><i>Results continues on the next page</i></p>	<p>Moderate</p> <p>Risk of bias affected by: Method of compensating for drop outs but new participants were not included in growth curve analysis. Author interest in programs</p> <p>Other comments GORT not included because of attrition, (higher than other measures). Some WRAT measures not included because measures are mixed (eg letter id and reading, but see spelling)</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>					
Morris et al 2012 [80] USA				<p>Follow-up I1 vs I2, I3: 3.91* (in favour of I2, I3) I2 vs I3: ns and values not reported C: N/A</p> <p>Non-words (F-score, TOWRE, Nonword Reading efficiency)</p> <p>Post-test C vs I1, I2, I3: ns and values not reported I1 vs I2, I3: ns and values not reported I2 vs I3: ns and values not reported</p> <p>Follow-up I1 vs I2, I3: 7.19** (in favour of I2, I3) I2 vs I3: ns and values not reported C: N/A</p> <p>Spelling (F-score, WRMT-R, Spelling-mix of letter writing and spelling) C vs I1, I2, I3: 44.97*** I1 vs I2, I3: ns and values not reported I2 vs I3: 3.29^a</p> <p>Follow-up I1 vs I2, I3: 6.52* I2 vs I3: ns and values not reported C: N/A</p> <p>Reading comprehension (F-score, WRMT-R, Passage Comprehension) C vs I1, I2, I3: 24.48*** I1 vs I2, I3: 4.72* (in favour of I2, I3) I2 vs I3: ns and values not reported</p> <p>Follow-up I1 vs I2, I3: ns and values not reported I2 vs I3: ns and values not reported C: N/A</p>	

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Nag- Arulmani et al 2003 [84] India	<p>Design RCT</p> <p>Aim Compare the effects of a phonological intervention an oral proficiency intervention and special instruction as usual</p> <p>Settings 4 schools in Bangalore, 9th grade, South India</p> <p>Population 526 students were screened with Word basic reading test (WBR) and Word reading test (WRC) to group the students in reading disability and reading difficult control. The final sample, 118 students, n=90 students with reading disabilities, n=28 students with no reading disabilities</p> <p>Inclusion criteria Identified as showing difficulty with reading by the class teacher & scored below their chronological age on either WBR or WRC tests. English as 1st language, since Grade 1 and not been introduced to any other script prior to English. No sensory deficits, emotional or major health problems, reported use of at least 3 languages and rated Kannada as a dominant spoken language. Students who scored above the cut-off on a Kannada listening comprehension test were retained, to ensure minimum fluency in the dominant language Kannada</p> <p><i>Results continues on the next page</i></p>	<p>Intervention I1: Phonological intervention I2: Language exposure intervention Small groups of children: 3–8 students</p> <p>Participants I1: n=31 I2: n=31</p> <p>Intensity 10 x 90 min, 3 times a week, during 3 weeks</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Comparison group C: Non-language based intervention; focused on art and craft activities</p> <p>Intensity Same as intervention groups</p> <p>Participants n=28</p> <p>Drop-out rate Not reported</p> <p>Fidelity Not reported</p>	<p>Word reading (accuracy) Words Post-test: (T2) I1: 33.42 (5.25) p<0.001 I2: 26.77 (7.86) C: 29.57 (3.16)</p> <p>Follow-up: (T3) I1: 37.45 (4.48) p<0.001 I2: 32.48 (5.92) C: 33.12 (2.79)</p> <p>Non-words Post-test: (T2) I1: 20.81 (5.34) p<0.001 I2: 13.39 (6.27) C: 16.71 (3.65)</p> <p>Follow-up: (T3) I1:23.94 (3.99) p<0.0001 I2:17.29 (5.06) C: 33.12 (2.79)</p> <p>Spelling Post-test: I1: 26.25 (2.46) ns I2: 25.19 (2.91) C: 25.39 (2.03)</p> <p>Follow-up: I1: 28 (2.32) p<0.01 I2 : 26.19 (2.55) C: 26 (2.02)</p> <p>Reading comprehension Post-test: I1: 14.77 (3.19) ns I2: 16.32 (3.45) C: 16.14 (3.81)</p> <p><i>Results continues on the next page</i></p>	<p>Moderate</p> <p>Comments Poor information regarding drop-outs. No information regarding fidelity. All interventions focused on or used English for communication. Class teachers and parents were blind to the intervention group their children were assigned to</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i> Nag- Arulmani et al 2003 [84] India	Follow-up Study conducted in 3 phases: (T1) Start (T2) Post-test at 6 weeks (T3) Follow-up at 9 months from intervention			Follow-up: I1: 18 (2.84) ns I2: 19.03 (3.94) C: 19.12 (3.81) Phonemic awareness Post-test: I1: 118.35 (15.28) p<0.01 I2: 106.87 (16.14) C: 103.71 (18.88) Follow-up: I1: 113.64 (11.62) p<0.01 I2: 104.09 (10.40) C: 104.79 (9.78)	
Olson et al 1997 [88] USA	Design Quasi-experimental Aim Investigate the effect of 2 computer-based remedial training programs Setting Second to fifth graders Population (=participants) n=103 Inclusion criteria IQ above 90 and were in approximately the lower 10% of their class in word decoding Follow-up 1- and 2-year follow-up	Intervention I1:Phonological training Participants I1: n=58 Mean age: 8.97 years Follow-up 1: n=27 Follow-up 2: n=19 Intensity 25 hours Drop-out rate High from pre-test to follow up Fidelity Not reported	Comparison group I2: Comprehension strategies Participants I2: n=45 Age: mean 8.96 years Follow-up 1: n=20 Follow-up 2: n=17 Intensity 25 hours Drop-out rate High from pre-test to follow up Fidelity Not reported	Effect sizes at end of training, 1 year and 2 year follow-up Word reading (accuracy) <i>Words</i> Follow-up 2: I2 >I1: -0.06 Word reading <i>Non-words</i> Follow-up 2: I2 >I1: 0.38 Word reading <i>Speed</i> Follow-up 2: I2 >I1: 0.22 Phonological awareness Follow-up 2: I2 >I1: 0.59	Risk of bias Moderate Comments Only 2-years data are shown, 1-year data is also available in the article

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
O'Shaughnessy et al 2000 [72] USA	<p>Design RCT</p> <p>Aim To study the efficacy of 2 small-group research-based reading intervention</p> <p>Settings 2nd grade in 3 elementary schools with historically low achievement, in a school district in southern California</p> <p>Population n=45 (21 girls, 24 boys) Initially all second-grade children referred by their teachers. Mean age at pre-test: 7 years 8 months</p> <p>Inclusion criteria Average intelligence; scores below the 25 percentile (WRMT-R) & at least 1 year below grade level on oral fluency (CBM); scores below the 25 percentile on phonological sensitivity (TOPA)</p> <p>Exclusion criteria English as a second language extreme hyperactivity, hearing impairment, brain damage, chronic medical condition, serious emotional disturbance</p> <p>Follow-up: Pre-test-post-test only. Time for post-test not reported</p>	<p>Intervention I1: Direct, systematic training of sound blending, sound segmenting and letter-sound correspondences (PAT). Instruction and practice is decontextualised. Letter-sound units. Oral language skills emphasised. I2: Systematic training in rhyming and frequent spelling patterns via a compare/contrast decoding strategy (WAT). Instruction and practice is contextualised. Embedded within the context of reading and spelling activities. Onset-rime /whole word units. Written language skills emphasised</p> <p>Intensity 30 minutes/day, 3 times a week during 6 weeks</p> <p>Participants PAT: n=15 WAT: n=15</p> <p>Drop-out rate Not reported</p> <p>Fidelity High</p>	<p>Comparison group Mathematics training</p> <p>Intensity Same as training groups. (Control group received PAT after post-test)</p> <p>Participants n=15</p> <p>Drop-out rate Not reported</p> <p>Fidelity High</p>	<p>Data presented in raw scores. Effect size (ES)</p> <p>Word reading (accuracy) Words Word identification (WRMT-R) Mean (standard deviation) Post-test: I1: 20.9 (9.6) I2: 23.1 (6.5) C: 24.3 (9.6)</p> <p>Word reading (accuracy) Non-words Word-attack (WRMT-R) Post-test: I1: 8.2 (3.4) I2: 8.3 (3.4) C: 6.3 (5.5) ES=0.46 I1 >C ES=0.47 I2 >C</p> <p>Spelling PIAT-R Post-test: I1: 27.3 (6.1) I2: 31.1 (5.6) C: 27.3 (6.1) ES=0.42 I1 >C ES=0.65 I2 >C</p> <p><i>Results continues on the next page</i></p>	<p>Moderate</p> <p>Comments Small number of participants in each group (n=15). Short intervention period</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Continued					
O'Shaughnessy et al 2000 [72] USA				<p>Reading comprehension Passage Comprehension (WRMT-R) Post-test: I1: 14.2 (7.2) I2: 15.2 (4.3) C: 11.8 (6.1) ES=0.36 I1 >C ES=0.65 I2 >C</p> <p>Phonemic awareness Phoneme deletion Post-test: I1: 8.4 (1) I2: 7.8 (1.8) C: 4.7 (2.2)</p>	

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Rashotte et al 2001 [93] Canada	<p>Design RCT matched pairs, cross-over design</p> <p>Aim Compared effectiveness of a phonologically based, small-group reading program</p> <p>Setting 1 element school, 1–6th grade impaired readers, low socio-economic status, Newfoundland, Canada</p> <p>Population n=116 (55 girls, 61 boys) fulfilled inclusion criteria and were matched in pairs depending on decoding and word level reading for each grade level. Age: 6.1–12.8 years (mean age 8.9 years)</p> <p>Inclusion criteria Identified by their school as 2 year below grade level in reading. Average score for selected children fell at the WJ:s test. Word Attack at the 13th percentile and at the Word identification at the 19th percentile</p> <p>Exclusion criteria Not reported</p> <p>Follow-up Post-test 1 and 2, since this is a cross-over design</p>	<p>Intervention I: SRPAT: Phonemic awareness & phonics training; reading and writing for meaning in groups of 3–5 students</p> <p>Intensity 50 minutes/day; 8 weeks=35 hours</p> <p>Participants n=58</p> <p>Drop-out rate None</p> <p>Fidelity Acceptable</p>	<p>Comparison group C: Regular reading instruction</p> <p>Intensity Grade 1–3= 90 minutes/day Grade 4–6= 60+15 minutes/day</p> <p>Participants n=57</p> <p>Drop-out rate: n=1</p> <p>Fidelity Not reported</p>	<p>Cohen's d (effect size (ES) for post-test</p> <p>Word reading (accuracy) <i>Words</i> GORT-3 ES=(grade 1/2) 0.73, (grade 3/4) 1.15, (grade 5/6) 0.38</p> <p><i>Non-words</i> Word attack (WDRB) ES=(grade 1/2) 1.67, (grade 3/4) 1.81, (grade 5/6) 2.20</p> <p>Word reading (speed) <i>Words</i> Word Efficiency (TOWRE) ES=(grade 1/2) 0.19, (grade 3/4) 0.22, (grade 5/6) –0.22</p> <p><i>Non-words</i> Phonetic decoding efficiency (TOWRE) ES=(grade 1/2) 1.31, (grade 3/4) 1.91, (grade 5/6) 0.88</p> <p>Text reading fluency GORT-3 Rate ES=(grade 1/2) 0.63 (grade 3/4) 0.23, (grade 5/6) 0.92</p> <p>Spelling Schonell spelling ES=(grade 1/2) 0.75, (grade 3/4) 0.23 (grade 5/6) 0.06</p> <p>Reading comprehension GORT-3 Comprehension ES=(grade 1/2) 1.51, (grade 3/4) 0.60, (grade 5/6) 0.64</p> <p>Phonemic awareness Elision (CTOPP) ES=(grade 1/2) 0.40, (grade 3/4) 1.14, (grade 5/6) 0.51</p>	Risk of bias Moderate

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Shippen et al 2005 [73] USA	<p>Design RCT. Quasi-experimental</p> <p>Aim Comparison of 2 direct instruction reading programs</p> <p>Setting One 7th grade Urban inner-city middle school</p> <p>Population n=55 (22 girls, 33 boys) Age: 12.4–14.6 years Subjects segregated by reading level: B (lower performance in basic reading skills) and C The students were randomised to: I1: Interventions targets word attack skills, with emphasis on basic sound – symbol associations of individual letters, digraphs and blends (CRD) or I2: Teaches an overt reading decoding strategy that is faded during the final 8 lessons (REWARDS)</p> <p>Inclusion criteria Students 2 years behind in reading</p> <p>Follow-up Post-test</p>	<p>Intervention I1: CRD</p> <p>Intensity Daily 55 minutes for 30 days, for B=29 days; C=30 days</p> <p>Participants n=13 (B) + 12 (C)</p> <p>Drop-out rate None</p> <p>Fidelity Fidelity averages ranged from 67% to 89%</p>	<p>Comparison group I2: REWARDS</p> <p>Intensity Daily 55 minutes for 30 days, for B=29 days; C=30 days</p> <p>Participants n=16 (B) + 14 (C)</p> <p>Drop-out rate None</p> <p>Fidelity Fidelity averages ranged from 67% to 89%</p>	<p>Reading level group: B Mean (standard deviation) Word reading efficiency (TOWRE) I1: 0.85 (6.89) I2: 2.56 (4.03) Reading accuracy (GORT-4) I1: 0.8 (3.04) I2: 1.19 (2.51) Reading rate (GORT-4) I1: 0.15 (0.99) I2: 0.88 (1.31) Reading fluency (GORT-4) I1: -0.79 (4.04) I2: 1.19 (2.17)</p> <p>Reading level group C Mean (standard deviation) Word reading efficiency (TOWRE) I1: 5.67 (10.88) I2: 1.14 (5.16) Reading accuracy (GORT-4) I1: 0.25 (2.73) I2: 2.14 (1.61) Reading rate (GORT-4) I1: 1.91 (1.73) I2: 0.93 (1.64) Reading fluency (GORT-4) I1: 1.42 (3.02) I2: 2.79 (2.52)</p>	<p>Moderate</p> <p>Comments Attrition rate high. Small effect size. Fidelity variation may have affected outcome</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Spencer et al 2010 [96] USA	<p>Design RCT</p> <p>Aim To test the efficacy of a fluency intervention program on word-identification and reading-comprehension</p> <p>Settings 2 middle schools on the outskirts of a large urban city on the west coast. (Grades 6–8)</p> <p>Population n=60 17 girls, 43 boys. Age: 10 years 11 month–15 years. All enrolled in Special Day Programs (SDPs)</p> <p>Inclusion criteria WRMT-WI: <63; (63–78=plus GORT III)</p> <p>Follow-up Post-test (1 week after last I)</p>	<p>Intervention I: Great Leaps Reading: Fluency activities at 3 levels: sounds or individual words, short sight phrases, and connected text</p> <p>Participants n=34, (school 1, n=19, school 2, n=15)</p> <p>Intensity 10 minutes daily during 6 months. Average number sessions/week=2.1 to 4.1, overall group average 2.7 sessions per week. 47 out of the 60 students completed more than 50 sessions</p> <p>Drop-out rate 5.5%</p> <p>Fidelity Moderate</p>	<p>Comparison group C: The Skills for School Success program (general classroom & study skills)</p> <p>Participants n=26 students (school 1 n=15, school 2, n=11)</p> <p>Intensity 10 minutes daily during 6 months. Average 2.2 sessions per week=average 48.2 sessions per student</p> <p>Drop-out rate None</p> <p>Fidelity Moderate</p>	<p>Word reading (accuracy) Words Post-test: WJ, Word identification Cohen's d: 0.23, ns</p> <p>Post-test: WRMT-R/NU Sight Word Efficiency Cohen's d: 0.18, ns</p> <p>Post-test: GORT III Accuracy Cohen's d: 0.62, p<0.05</p> <p>Non-words Post-test: WRMT-R/NU Word Attack Cohen's d: 0.18, ns</p> <p>Word reading (speed) Words Post-test: GORT III Rate Cohen's d: 0.59, p<0.05</p> <p>Reading comprehension Post-test: WJ Passage comprehension Cohen's d: 0.02, ns</p> <p>Phonemic awareness Post-test: WRMT-R/NU Phonemic Decoding Efficiency Subtest Cohen's d: 0.41, p<0.05</p>	<p>Moderate</p> <p>Comments Large intervention variability, which might have an effect on the outcomes. Acceptable according to authors</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Steenbeek-Planting et al 2013 [74,87] The Netherlands	<p>Design RCT</p> <p>Aim The effects of 2 training procedures on the improvement of reading accuracy in poor readers, in relation to their initial reading level</p> <p>Settings Selected from 2 primary schools for special education</p> <p>Population n=60, 24 girls, 36 boys. Mean age: 10.5 years Children were qualified as having either a high initial reading level (HI) or a low initial reading level (LI), based on their reading soccer on the Word decoding test (WDT) prior to intervention</p> <p>Inclusion criteria The 10th percentile for their grade level on a Word decoding test. The 10th percentile for their grade level on a Lexical Decision Test or below the 12th percentile for their grade level on a non-word reading test. Classified as poor readers by their teachers, and problems were not caused by behavioral problems</p> <p>Exklusion criteria Children with mental retardation or behavioral problems were excluded from participation.</p> <p>Follow-up Post-test</p>	<p>Intervention I1: Training focusing on words the children read correctly (successes) I2: Training focusing on the words the children read incorrectly (failures)</p> <p>Participants n=9+11 (I1: LI, HI) n=10+10 (I2: LI, HI)</p> <p>Intensity 11 sessions of 20 minutes each. Children practiced individually once or twice a week</p> <p>Drop-out rate None</p> <p>Fidelity Acceptable</p>	<p>Comparison group C: No training</p> <p>Participants n=11+9</p> <p>Intensity N/A</p> <p>Drop-out rate None</p> <p>Fidelity Not relevant</p>	<p>Post-test Word reading (speed) Words Low initial level I1: 71.72 (19.40) (n=9) I2: 81.83 (11.04) (n=10) C: 60.06 (17.32) (n=11)</p> <p>High initial level I1: 85.79 (6.83) (n=11) I2: 82.14 (5.88) (n=10) C: 86.57 (14.55) (n=9)</p> <p>Non-words Low initial level I1: 46.77 (20.76) (n=9) I2: 61.62 (17.43) (n=10) C: 44.87 (16.83) (n=11)</p> <p>High initial level I1: 55.71 (16.52) (n=11) I2: 52.16 (8.76) (n=10) C: 63.51 (17.58) (n=9)</p>	<p>Moderate</p> <p>Risk of bias Randomisation procedure not described</p> <p>Comments Few participants in each group. The pre-test values are different at baseline between the different groups</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Tijms 2011 [100] The Netherlands	<p>Design Quasi-experimental</p> <p>Aim Investigate the clinical effectiveness of a treatment for children with dyslexia, as well as the moderating impact of plausible cognitive and socioeconomic factors</p> <p>Setting Children, grade 3–6, with dyslexia was recruited from the IWAL Institute in The Netherlands</p> <p>Population (=participants) n=99 Age: 9–12 years</p> <p>Inclusion criteria Either word reading rate 1.5 standard deviation or more below the mean, or word reading rate between 1 and 1.5 below the mean together with a spelling skill of 1.5 standard deviation or more below the mean. Performance on at least 2 of 4 administered phonological tasks, 1 standard deviation or more below the mean. Special remedial teaching was provided at school before entering the treatment programme</p> <p>Exclusion criteria Uncorrected sensory disabilities, broad neurological deficits and insufficient education</p> <p>Follow-up None</p>	<p>Intervention I1: Treatment programme focuses on systematic learning of letter-speech sound mappings</p> <p>Participants I1: n=75</p> <p>Intensity 43.45 treatment sessions over a period of an average 17.35 months and additional 15 minutes of home training 3 times a week</p> <p>Drop-out rate n=2</p> <p>Fidelity 86% attendance</p>	<p>Comparison group I2: Control</p> <p>Participants I2: n=21</p> <p>Intensity N/A</p> <p>Drop-out rate n=1</p> <p>Fidelity N/A</p>	<p>Effect size (area under the curve (AUC). AUC of 50% reflects a toss-up between treatment and control. An AUC of 71% corresponds to Cohen's d=0.80, usually large effect size</p> <p>Word reading <i>Speed</i> Post-test: 79%</p> <p><i>Spelling</i> Post-test: 93%</p>	<p>Risk of bias Moderate</p>

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Torgesen et al 2001 [75] USA	<p>Design RCT</p> <p>Aim Compare 2 approaches that both contain explicit instruction in word level skills but vary systematically in their depth of instruction in phonemic awareness and extent of practice in decontextualised phonemic decoding skills</p> <p>Setting Learning disability classes in 3 elementary schools in Florida</p> <p>Population n=60. Age=8–10 years (mean age: 9.8 years)</p> <p>Inclusion criteria Identified by teachers with severe difficulties. 1.5 standard deviation below on reading test. Intelligence >IQ 75. Performed low on phonological awareness. Allowed variables such as socio-economic status, race, gender, and co-morbidity of ADHD to vary</p> <p>Exclusion criteria Children that were adopted; evidence of an acquired neurological disease; experienced a perinatal encephalopathic event; sensory deficits (hearing loss, visual acuity); chronic medical illness; severe psychopathology; English second language</p> <p>Follow-up After 1 and 2 years</p>	<p>Intervention I1: Auditory discrimination in depth (ADD) program has 3 goals. Discriminate phonemes by teaching distinctive kinesthetic, auditory and visual features. Use knowledge of the distinctive features of phonemes to monitor and represent sequences of sounds in spoken syllables. Teach children self-monitoring skills</p> <p>Intensity Two 50-minutes sessions (1:1 basis) each day of the week for 8 to 9 weeks until each child has reached 67.5 hours of instructions. After that generalisation training one 50-minutes session/week for 8 weeks by the teachers that carried out the intensive training period</p> <p>Participants n=26</p> <p>Drop-out rate 20%. No differences on the major pre-test or demographic variables between the remaining sample and the original sample</p> <p>Fidelity Good</p>	<p>Comparison group I2: Embedded phonics (EP). Explicit instruction in word-level reading skills while providing extensive opportunities to read and write meaningful text</p> <p>Intensity As in the intervention group period</p> <p>Participants n=24</p> <p>Drop-out rate 20%</p> <p>Fidelity Good</p>	<p>Mean (standard deviation) Raw score;</p> <p>Word reading (accuracy) <i>Non-words</i> Post-test: I1: 89.4 (12.4) I2: 90.3 (8.3) Follow-up: I1: 93.7 (12.3) I2: 87 (standard deviation =8.9)</p> <p>Word reading (speed) <i>Word</i> Post-test: I1: 75.4 (8.2) I2: 72.1 (7.9) Follow-up: I1: 75.0 (9.3) I2: 72.1 (13.2)</p> <p>Non-words (Word attack) Post-test: I1: 96.4 (7) I2: 90.3 (8.3) Follow-up: I1: 90.7 (9.3) I2: 87 (8.9)</p> <p>Spelling (KTEA) Post-test: I1: 77.5 (5) I2: 80 (7.4) Follow-up: I1: 76.7 (5.8) I2: 74 (6.6)</p>	<p>Moderate</p> <p>Risk of bias affected by Randomisation procedure not described. No control over the medication procedure between the groups</p> <p>Other comments Within the sample there are high degree (71–81%) of comorbid attention-deficit disorder. The medication condition during pre- and post-testing was not carefully controlled in this study, nor was medication consistent over the course treatment</p>

Results continues on the next page

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>				Reading comprehension (Woodcock) Post-test: I1: 91 (9) I2: 92 (19.8) Follow-up: I1: 92.8 (8) I2: 91.5 (10.8)	
Torgesen et al 2001 [75] USA				Phonemic awareness (CTOPP) Post-test: I1: 101 (14.3) I2: 97.9 (standard deviation =12.8) Follow-up: I1: 97.9 (13.7) I2: 94.4 (14.3)	
Wexler et al 2010 [76] USA	Design RCT, with matched pairs Aim Examine the efficacy of repeated reading and wide reading practice interventions in a peer-pairing format, for high school students with severe reading disabilities Settings 11 special education and reading classes in a Metropolitan area in southwestern USA Population (=participants) n=106. Grades 9–12. Age: 13–17 years. Students were paired within classes based on median pre-test oral reading fluency scores; higher-level readers were paired with lower-level readers. Pairs were randomly assigned to interventions or control	Intervention I1. Repeated reading: one text 3 times daily, including error correction & summarisation. I2. Wide reading: 3 different texts daily, otherwise same as above Participants I1: n=33 I2: n=34 Intensity 15–20 min, 5 days a week for 10 weeks, average 12.2 hours Drop-out rate 9% all participants Fidelity 91–100%	Comparison group C: Training as usual Participants n=29 Intensity: Same as I1 and I2 Drop-out rate See Intervention Fidelity N/A	Post-test (Cohen's d) Word reading (speed) <i>Words</i> Oral reading fluency I1 vs C, Cohen's d = -0.05 ns I2 vs C, Cohen's d = -0.11 ns I1 vs I2, Cohen's d = 0.06 ns Text reading fluency TOSCRF I1 vs C, Cohen's d = 0.04 ns I2 vs C, Cohen's d = -0.23 ns I1 vs I2, Cohen's d = 0.27 ns Reading comprehension WJ-C I1 vs C, Cohen's d = -0.10 ns I2 vs C, Cohen's d = -0.20 ns I1 vs I2, Cohen's d = 0.10 ns	Moderate Risk of bias affected by Randomisation procedure not described Comments According to authors, no differential attrition across conditions

Results continues on the next page

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>	Inclusion criteria Significant reading problems based on low scores on the reading portion of TAKS				
Wexler et al 2010 [76] USA	Follow-up: Post-test				
Wise et al 2000 [101] USA	Design Quasi-experimental Aim Investigate the effect of speech-supported reading in context, with and without explicit phonological training Setting Second to fifth graders from Boulders Valley Schools Population (=participants) n= 200 Age: 7–11 years old Inclusion criteria Classroom teachers referred children with normal IQ and first language English, whom they thought were in the lower 10% of readers based on word reading Follow-up 2 years	Intervention I1: Accurate-reading-in-context Participants I1: n=91 36 girls, 55 boys Mean age: 8.92 Intensity 22 individualised computer hours reading stories, and 7 small group hours learning comprehension strategies. During 6 months Drop-out rate 23 students were dropout (I1 and I2) because of moving, scheduling problems, behavior problems Fidelity Good	Comparison group I2: Phonological-analysis Participants I2: n=109 46 girls, 63 boys Mean age: 8.96 Intensity: Phonological strategies in 7 small-group hours, and divided their computer time between phonological exercise and story reading. 27–29 hours of training in 6 months Drop-out rate See I1 Fidelity Good	Effect size (end of training) *** p<0.001 * p<0.05 Word reading Words (PIAT) I2 >I1: 0.31* Word reading Non-words (Non-word decoding) I2 >I1: 0.98*** Spelling (PIAT spelling) I2 >I1: 0.98*** Phonological awareness (Phoneme deletion) I2 >I1: 0.66*** Reading comprehension (PIAT comprehension) I2 >I1: 0.21 Test at 2-years follow-up indicated that gains in phonological decoding for the phonological-analysis treatment condition were no longer reliably greater than those for the accurate-reading-in-context treatment condition, p=0.13	Moderate

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Wise et al [85] 1995 USA	<p>Design RCT-assign in sets of 3, of similar age</p> <p>Aim Investigate the differential effects of 2 training methods</p> <p>Setting Second to fifth graders from Boulders Valley Schools</p> <p>Population (=participants) n=103 Average age: 9 years old</p> <p>Inclusion criteria Recommended by teachers to be in the lower 10% on reading ability. Verbal and performance IQ estimate of 90 or above, either by the school's testing or by their own estimate using 4 subtest of the WISC test</p> <p>Follow-up None</p>	<p>Intervention I1: Comprehension strategy training</p> <p>Participants I1: n=25 (15 girls, 30 boys) Mean age: 8.96 years (7.7–9.10)</p> <p>Intensity 7 total hours of small-group instruction (in 30-minutes and 15-minute sessions)</p> <p>Drop-out rate Total 7.2% of all participants, due to moving (n=3) and scheduling complications (n=5)</p> <p>Fidelity Not reported</p>	<p>Comparison group I2: Phonological awareness training</p> <p>Participants I2: n=58 (19 girls, 39 boys). Mean age: 8.97 years</p> <p>Intensity: Total small group sessions. 4 of the 7 hours of small group time were used to introduce sounds, letters, and articulatory labels and mouth pictures for all consonants and vowels</p> <p>Drop-out rate See intervention</p> <p>Fidelity Not reported</p>	<p>Gain scores (percentage correct/grade equivalent) for comprehension strategy and phonological analysis * p<0.05 *** p<0.001</p> <p>Gain score (grade equivalent) Word reading <i>Words</i> Post-test: I1: 0.9 I2 :1.1</p> <p>Gain score (percentage correct) Word reading <i>Non-words</i> Post-test: I1: 4.9 I2: 27.00***</p> <p>Gain score (percentage correct) Word reading <i>Speed</i> Post-test: I1:17.1* I2: 12.5</p> <p>Gain score (grade equivalent) Spelling Post-test: I1: 0.6 I2: 0.5</p>	<p>Moderate</p> <p>Comments Training groups did not differ on pre-test measures. Students were assigned in sets of 3 of similar age. We arranged sets so that within each school, half of the sets matched the other half on age, word recognition and nonsense word reading. Then the sets group were randomly assigned to condition</p>

Results continues on the next page

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
<i>Continued</i>				Gain score (percentage correct)	
Wise et al [85] 1995 USA				Phonological awareness Post-test: I1: 6.00 I2: 19.9***	
				Gain score (grade equivalent) Reading comprehension Post-test: I1: 0.4 I2: 0.6	

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Wolff 2011 [90] Sweden	<p>Design RCT</p> <p>Aim To investigate the effects of a reading & fluency training program based on phonemic awareness, RAFT</p> <p>Settings Schools in rural and urban areas in different parts of Sweden</p> <p>Population n=2 212 3rd grade students from 59 schools in 12 municipalities. Age: 9; mean age 9.25 n=117 fulfilled inclusion criteria and accepted to participate</p> <p>Inclusion criteria 1 standard deviation below mean on phonological choice task & 1 standard deviation below mean on either word decoding or orthographic choice or both</p> <p>Follow-up: Post-test and follow-up (1 year)</p>	<p>Intervention I1: RAFT One-to-one tutoring by program-trained teachers</p> <p>Participants n=58 25 girls, 33 boys</p> <p>Intensity 40 minutes daily for 12 weeks. Total: 40 hours</p> <p>Drop-out rate 1.8%</p> <p>Fidelity High</p>	<p>Comparison group Training as usual (75% were in special education classes)</p> <p>Participants n= 59 (25 girls, 43 boys)</p> <p>Intensity N/A</p> <p>Drop-out rate 7.3%</p> <p>Fidelity N/A</p>	<p>Mean (standard deviation)/Cohen's d</p> <p>Word reading (speed) <i>Words</i> Follow-up: I: 72 (21.0) C: 72 (22.8)</p> <p>Reading speed Cohen's d: 0.15</p> <p>Spelling, 8 words Post-test: I: 5.47 (1.28) C: 4.87 (1.58)</p> <p>Follow-up: I: 6.53 (1.56) C: 6.11 (1.28) Cohen's d: 0.30</p> <p>Reading comprehension Post-test: I:20.1 (7.56) C:18.4 (7.29) Follow-up: I: 22.4 (5.83) C: 20.5 (8.40) Cohen's d: 0.41</p> <p>Phonemic awareness Post-test: I:2.54 (2.49) C:2.51 (2.03) Follow-up: I: 3.09 (2.21) C: 3.39 (2.06) Cohen's d: 0.43</p>	Moderate

The table continues on the next page

Table 17.3 continued

Author Year Reference Country	Design Aim Setting Population Follow-up	Intervention Intensity Participants Drop-out rate Fidelity	Comparison group Intensity Participants Drop-out rate Fidelity	Outcome	Risk of bias Comments
Wright et al 2003 [82] UK	<p>Design RCT</p> <p>Aim Investigate the differential effects of 2 instructional methods</p> <p>Setting 3 local education authority primary schools in northwest of England</p> <p>Population (=participants) n=65. Age: 7–10 years</p> <p>Inclusion criteria Teacher concerns about poor progress in reading & spelling development + reading age/chronological age discrepancy of approx 2 years according to standard tests used in the schools</p> <p>Follow-up 12 months</p>	<p>Intervention I1: Direct instruction in procedural knowledge of phonological awareness and alphabetical principles. I2: I1 + metalinguistic concepts and meta-cognitive strategies. Groups of 4–5 children</p> <p>Participants I1: n=22 (10 girls, 12 boys) Mean age: 8.9 years (7.7–9.10) I2: n=24 (19 girls, 15 boys) Mean age: 8.9 years (7.7–9.11)</p> <p>Intensity 30 minutes twice weekly during 2 school terms. Total=40 sessions</p> <p>Drop-out rate 5=8% of all participants</p> <p>Fidelity Not reported</p>	<p>Comparison group C: Training Support as usual+1/1 oral reading with adult</p> <p>Participants n=14 (7 girls, 7 boys) Mean age: 9 years (7.4–10)</p> <p>Intensity: N/A</p> <p>Drop-out rate See intervention</p> <p>Fidelity N/A</p>	<p>The result is presented in reading age and spelling age scores in relation to chronological age</p> <p>Word reading (speed) Words Graded Word Reading test Pre-test I1: 6.9 (.98) I2: 7 (.35) C: 7.4 (.36)</p> <p>Post-test: I1: 7.7 (0.54) I2: 7.9 (0.79) C: 8.4 (0.66)</p> <p>Follow-up: I1: 7.8 (0.72) I2: 8.2 (0.47) C: 8.4 (0.66)</p> <p>Spelling Graded word spelling test Pre-test: I1: 6.7 (1.7) I2: 6.6 (.87) C: 7.4 (.28)</p> <p>Post-test: I1: 7 (0.49) I2: 7.3 (0.72) C: 7.7 (0.14)</p> <p>Follow-up: I1: 7.6 (0.55) I2: 7.5 (0.46) C: 8.2 (0.28)</p>	<p>Moderate</p> <p>Comments Results hard to interpret e.g no reports regarding standard deviation on most of the reading and writing tests. No info on fidelity. Comparison group is not randomly assigned as the 2 intervention groups</p>

ADHD = Attention deficit hyperactive disorder; AIMSWeb = Oral reading fluency; BAS II = British Ability Scales II; CELF = Clinical Evaluation of Language Fundamentals; CTOPP = Comprehension Test of Phonological Processing; DO = Dyslexics without training; DW = Dyslexics with training; EWR Test = Early Word Recognition Test; GFW = Goldman Fristoe Woodcock Sound Symbol Tests; GORT = Gray Oral Reading Test; GORT III = Gray Oral Reading Test-III; GORT-4 = Gray Oral Reading Test-4; KIB-T = Kaufman Brief Intelligence Test, LPA = letter sound and phonological awareness; N/A = not applicable; NRST = Non-word Reading and Spelling Test; ns = not significant; PB = Phonics based reading program; PCF = phonemic awareness test; PIAT = Peabody Individual Achievement Test; PICTO = Pictogram; RAFT = Multicomponent intervention, Phonemic awareness/decoding (60%), comprehension strategies, fluency training (40%); RDC = Reading difficult control; SS = Standard Score (i.e norm referenced score around a mean of 100, SD 15); ST = Linguistics Skills Training; TAKS = Texas Assessment on Knowledge & Skill; TORF = Test of Oral Reading Fluency, Graded Nonword Reading and Spelling Test; TOSCRF = Test of silent contextual fluency; TOWRE = Test of Word reading efficiency; TRAD = Traditional method; SRPAT = Spell Read Phonological Auditory Training; WISC = Wechsler Intelligence Scale for Children-Revised; WIST = Word Identification Strategy Training program; WJ III = Woodcock-Johnsons Test of Achievement III; WJ = Woodcock-Johnsson; WJ-C = Woodcock-Johnson comprehension; WORD = Wechsler objective reading dimensions; WRAT-3 = Wide Range Achievement Test-3; WRC= WORD Reading Test; WRMT = Woodcock Reading Mastery Test; WRMT-R = Woodcock Reading Mastery Test-Revised; WRMT-T = Woodcock Reading Mastery Test, Third Edition; WRMT-WI = Woodcock Reading Mastery Test

18. Referenser

1. Lundberg I. Läsningens psykologi och pedagogik. Stockholm: Natur & Kultur, 2010.
2. SBU. Utvärdering av metoder i hälso- och sjukvården. En handbok. www.sbu.se/sv/Evidensbaserad-varld/Utvardering-av-metoder-i-halso-och-sjukvarden--En-handbok/.
3. Høien T, Lundberg, I. Dyslexi. Från teori till praktik. Stockholm: Natur & Kultur, 1999.
4. Lundberg I TF, Austad I. Towards a Sharper Definition of Dyslexia. *Dyslexia: Advances in Theory and Practice*. Kluwer Academic Publishers 1999;9-29.
5. Snowling M. Developmental dyslexia. *Curr Pediatr* 2001;11:10-13.
6. Harlaar N, Spinath, EM, Dale, PS, Plomin, R. Genetic influences on word recognition abilities and disabilities: A study of 7 years old twins. *J Child Psychol Psychiatry* 2006;46:373-86.
7. Wadsworth SJ, Olson RK, Pennington BF, DeFries JC. Differential genetic etiology of reading disability as a function of IQ. *J Learn Disabil* 2000;33:192-9.
8. Morgan W. A case of congenital word-blindness *BMJ* 1896;2:(1378).
9. Hinshelwood J. Congenital word-blindness. *Lancet* 1900;1:1506-08.
10. Orton S. Reading, writing and speech problems in children. Chapman & Hall, London, 1937.
11. Vellutino F. The validity of perceptual deficit explanations of reading disability: A replay to Fletcher and Satz. *J Learn Disabil* 1979;12:160-70.
12. Stanovich K, Siegel, LS. Phenotypic performance of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *J Educ Psychol* 1994;86:24-53.
13. Vellutino FR, Scanlon DM, Sipay ER, Small SG, Pratt A, Chen R, Denckla MB. Cognitive profiles of difficult to remediate and readily remediated poor readers: Toward distinguishing between constitutionally and experientially based causes of reading disability. *J Educ Psychol* 1996;88:601-38.
14. Stuebing KK, Barth AE, Molfese PJ, Weiss B, Fletcher JM. IQ is not strongly related to response to Reading Instruction: A Meta-Analytic Interpretation. *Except Child* 2009;76:31-51.
15. Willcutt EG, Pennington BF. Psychiatric comorbidity in children and adolescents with reading disability. *J Child Psychol Psychiatry* 2000;41:1039-48.
16. Bruce B, Thernlund G, Nettelbladt U. ADHD and language impairment: A study of the parent questionnaire FTF

- (Five to Fifteen). *Eur Child Adolesc Psychiatry* 2006;15:52-60.
17. Boada R, Willcutt, EG, Pennington BF. Understanding the Comorbidity between Dyslexia and Attention-Deficit/Hyperactivity Disorder. *Top Lang Disord* 2012;32:264-84.
 18. Bosse M-L, Tainturier MJ, Valdois S. Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition* 2007;104:198-230.
 19. Pennington BF, Santerre-Lemmon L, Rosenberg J, MacDonald B, Boada R, Friend A, et al. Individual prediction of dyslexia by single versus multiple deficit models. *J Abnorm Psychol* 2012;121:212-24.
 20. Snowling MJ. Specific disorders and broader phenotypes: The case of dyslexia. *Q J Exp Psychol* 2008;61:142-56.
 21. Läs- och skrivsvårigheter/Dyslexi i skolan. Skolinspektionens rapport 2011:8 <http://www.skolinspektionen.se/Documents/Kvalitetsgranskning/dyslexi/kvalgr-dyslexi-slutrappport.pdf>.
 22. Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R. Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *NEJM* 1992;326:145-50.
 23. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
 24. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10.
 25. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6.
 26. Catts HW. Early identification of dyslexia: Evidence from a follow-up study of speech-language impaired children. *Ann Dyslexia* 1991;41:163-77.
 27. Elbro C, Borstrøm I, Petersen DK. Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Read Res Q* 1998;33:36-60.
 28. Furnes B, Samuelsson S. Predicting reading and spelling difficulties in transparent and opaque orthographies: A comparison between Scandinavian and US/Australian children. *Dyslexia: An International Journal of Research and Practice* 2010;16:119-42.
 29. Helland T, Plante E, Hugdahl K. Predicting dyslexia at age 11 from a risk index questionnaire at age 5. *Dyslexia: An International Journal of Research and Practice* 2011;17:207-26.
 30. Lamm O, Share DL, Shatil E, Epstein R. Kindergarten dichotic listening as a predictor of first-grade reading achievement. *Dyslexia: An International Journal of Research and Practice* 1999;5:138-54.
 31. Niemi P, Nurmi JE, Lyyra AL, Lerkkanen MK, Lepola J, Poskiparta E,

- Poikkeus AM. Task avoidance, number skills and parental learning difficulties as predictors of poor response to instruction. *J Learn Disabil* 2011;44:459-71.
32. O'Connor RE, Jenkins JR. Prediction of reading disabilities in kindergarten and first grade. *Sci Stud Read* 1999; 3:159-97.
33. Puolakanaho A, Ahonen T, Aro M, Eklund K, Leppänen PHT, Poikkeus A-M, et al. Very early phonological and language skills: Estimating individual risk of reading disability. *J Child Psychol Psychiatry* 2007;48:923-31.
34. Burke Mack D, Hagan-Burke S. Concurrent criterion-Related validity of early literacy indicators for middle of first grade. *Assess Eff Interv* 2007; 32:66-77.
35. Clarke AT, Power TJ, Blom-Hoffman J, Dwyer JF, Kelleher CR, Novak M. Kindergarten Reading Engagement: An Investigation of Teacher Ratings. *Journal of Applied School Psychology* 2003;20:131-44.
36. Fien H, Baker SK, Smolkowski K, Smith JLM, Kame'enui EJ, Beck CT. Using nonsense word fluency to predict reading proficiency in kindergarten through Second grade for English learners and native English speakers. *School Psych Rev* 2008;37:391-408.
37. Hagan-Burke S, Burke Mack D, Crowder C. The convergent validity of the Dynamic Indicators of Basic Early Literacy Skills and the Test of Word Reading Efficiency for the beginning of first grade. *Assess Eff Interv* 2006; 31:1-15.
38. McBride James R, Ysseldyke J, Milone M, Stickney E. Technical adequacy and cost benefit of four measures of early literacy. *Canadian Journal of School Psychology* 2010;25:189-204.
39. Daub D, Colarusso Ronald P. The validity of the WJ-R, PIAT-R, and DAB-2 Reading subtests with students with learning disabilities. *Learn Disabil Res Pract* 1996;11:90-95.
40. Elliott J, Lee Steven W, Tollefson N. A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills-Modified. *School Psych Rev* 2001;30:33-49.
41. Erford BT, Dutton JL. Technical Analysis of the Slosson Phonics and Structural Analysis Test. *Educ Psychol Meas* 2005;65:1011-25.
42. Gotch CM, French BF. The factor structure of the CIBS-II-Readiness assessment. *J Psychoeduc Assess* 2011;29:249-60.
43. Havey JM, Story N, Buker K. Convergent and concurrent validity of two measures of phonological processing. *Psychol Sch* 2002;39:507-14.
44. Hintze JM, Ryan AL, Stoner G. Concurrent Validity and Diagnostic Accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing. *School Psych Rev* 2003; 32:541-56.
45. Keenan Janice M, Betjemann Rebecca S, Olson Richard K. Reading comprehension tests vary in the skills they assess: Differential dependence on

- decoding and oral comprehension. *Sci Stud Read* 2008;12:281-300.
46. Keenan J, Meenan CE. Test Differences in Diagnosing Reading Comprehension Deficits. *J Learn Disabil*, 2014;47:125-35.
 47. Kilpatrick DA. Phonological segmentation assessment is not enough: A comparison of three phonological awareness tests with first and second graders. *Canadian Journal of School Psychology* 2012;27:150-65.
 48. Mather N, Vogel SA, Spodak RB, McGrew KS. Use of the Woodcock-Johnson-Revised writing tests with students with learning disabilities. *J Psychoeduc Assess* 1991;9:296-307.
 49. Prewett P, Giannuli M. The relationship among the reading subtests of the WJ-R, PIAT-R, K-TEA, and WRAT-R. *J Psychoeduc Assess* 1991;9:166-74.
 50. Shull-Senn S, Weatherly M, Morgan SK, Bradley-Johnson S. Stability Reliability for Elementary-Age Students on the Woodcock-Johnson Psychoeducational Battery-Revised (Achievement Section) and the Kaufman Test of Educational Achievement. *Psychol Sch* 1995;32:86-92.
 51. French B, Glascoe FP. CIBS-II standardization and validation manual: Brigance Diagnostic Comprehensive Inventory of Basic Skills-II. North Billerica, MA: Curriculum Associates, 2010.
 52. Semel E, Wiig EH, Secord WA. Clinical evaluation of language fundamentals, fourth edition-Screening test (CELF-4 screening test). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company, 2004.
 53. Wagner R, Torgesen JK, Rashotte CA. The Comprehensive Test of Phonological Processing (CTOPP). Austin, TX:Pro-Ed, 1999.
 54. Good R, Kaminski RA. Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *Sch Psychol Q* 1996;11:326-36.
 55. MacGinitie WH, MacGinitie RK, Maria K, Dreyer LG, Hughes KE. Gates-MacGinitie reading tests- 4th Ed. (GMRT-4). Riverside Publishing; 2000.
 56. Wiederholt L, Bryant B. Examiner's manual: Gray Oral Reading Test-3. Austin, TX:Pro-Ed, 1992.
 57. Fugate M. Review of the Group Reading Assessment and Diagnostic Evaluation. *Mental Measurements Yearbook*, 15, 2002.
 58. Badian N. The prediction of good and poor reading before kindergarten entry: A four year follow-up. *Education* 1982;16:308-18.
 59. Clarke A, Power TJ, Blom-Hoffman J, Dwyer JF, Kelleher CR, Novak M. Kindergarten reading engagement: An investigation of teacher ratings. *Journal of Applied School Psychology* 2003;20:131-44.
 60. Markwardt F, Jr. Manual. Peabody Individual Achievement Test-Revised. Circle Pines, MN: American Guidance Service, 1989.

61. Karlsen B, Gardner E. Stanford Diagnostic Test. San Francisco: Harcourt Brace, 1994.
62. Torgesen J, Bryant BR. TOPA – 2 + Test of Phonological Awareness Second Edition: PLUS. Austin, Texas: pro-ed, 2004.
63. Torgesen J, Wagner R, Rashotte C. Test of Word Reading Efficiency (TOWRE), 1999. ProEd Publishing Co, 8700 Shoal Creek Blvd, Austin, TX 78757-6897.
64. Texas Education Agency, 2010. The University of Texas System & The University of Houston System. Brookes Publishing Co.
65. Wechsler D. Wechsler Individual Achievement Test. San Antonio, TX: NCS Pearson, 2001.
66. Woodcock R, Johnson M, Woodcock-Johnson Psychoeducational Battery-Revised. Allen, TX: DLM, 1989.
67. Woodcock R, McGrew KS, Mather N. Woodcock-Johnson III. Itasca, IL: Riverside Publishing, 2001.
68. Wilkinson G, Robertson GJ. Wide Range Achievement Test 4 professional manual. Lutz, FL: Psychological Assessment Resources, 2006.
69. York Assessment of Reading for Comprehension (YARC). <http://www.yarcsupport.co.uk/>
70. Woodcock R, Mather N, Wendling B. Essentials of WJ III Tests of Achievement Assessment. John Wiley & Sons. New York, NY, 2001.
71. Das J, Hayward D, Georgiou G, Janzen T, Boora N. Comparing the effectiveness of two reading intervention programs for children with reading disabilities. *Educación y Diversidad*, 2009;ISSN: 1888-4857:15-36.
72. O'Shaughnessy Tam E, Swanson HL. A comparison of two reading interventions for children with reading disabilities. *J Learn Disabil* 2000;33:257-77.
73. Shippen Margaret E, Houchins David E, Steventon C, Sartor D. A Comparison of Two Direct Instruction Reading Programs for Urban Middle School Students. *Remedial and Special Education* 2005;26:175-82.
74. Steenbeek-Planting EG, van Bon, WHJ, Schreuder, R. Improving the reading of bisyllabic words that involve context-sensitive spelling rules: focus on successes or on failures? *Springer* 2013;26:1437-58.
75. Torgesen JK, Alexander AW, Wagner RK, Rashotte CA, Voeller KK, Conway T. Intensive remedial instruction for children with severe reading disabilities: immediate and long-term outcomes from two instructional approaches. *J Learn Disabil* 2001;34:33-58, 78.
76. Wexler J, Vaughn S, Roberts G, Denton Carolyn A. The efficacy of repeated reading and wide reading practice for high school students with severe reading disabilities. *Learn Disabil Res Pract* 2010;25:2-10.
77. del Rosario Ortiz GM, Espinel AI, Rosquete RG. Remedial interventions for children with reading disabilities: speech perception - an effective compo-

- ment in phonological training?
J Learn Disabil 2002;35:334-42.
78. Fälth L, Gustafson S, Tjus T, Heimann M, Svensson I. Computer-assisted interventions targeting reading skills of children with reading disabilities - a longitudinal study. *Dyslexia* 2013;19:37-53.
79. Kerstholt Mariette T. Training in Phonemic Segmentation: The Effects of Visual Support. *Read Writ: An Interdisciplinary Journal* 1994;6:361-85.
80. Morris R, Lovett M, Wolf M, Sevcik R, Steinbach K, Frijters J, Shapiro M. Multiple-component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *J Learn Disabil* 2012;45:99-127.
81. Macaruso P, Rodman A. Benefits of computer-assisted instruction for struggling readers in middle school. *Eur J Spec Needs Educ* 2009;24:103-13.
82. Wright J, Jacobs B. Teaching phonological awareness and metacognitive strategies to children with reading difficulties: A comparison of the two instructional methods. *Educ Psychol* 2003;23:17-45.
83. Mavrommati TD, Miles TR. A Pictorial Method for Teaching Spelling to Greek Dyslexic Children. *Dyslexia* 2002;8:86-101.
84. Nag-Arulmani S, Reddy V, Buckley S. Targeting phonological representations can help in the early stages of reading in a non-dominant language. *J Res Read* 2003;26:49-68.
85. Wise BW, Olson RK. Computer-Based Phonological Awareness and Reading Instruction. *Ann Dyslexia* 1995;45:99-122.
86. Calhoon MB. Effects of a Peer-Mediated Phonological Skill and Reading Comprehension Program on Reading Skill Acquisition for Middle School Students with Reading Disabilities. *J Learn Disabil* 2005;38:424-33.
87. Steenbeek-Planting EG, van Bon WH, Schreuder R. Improving word reading speed: individual differences interact with a training focus on successes or failures. *Read Writ* 2012;25:2061-89.
88. Olson RK, Wise B, Ring J, Johnson M. Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the posttraining development of word recognition. *Sci Stud Read* 1997;1:235-53.
89. Calhoon MB, Sandow A, Hunter CV. Reorganizing the instructional reading components: could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Ann Dyslexia* 2010;60:57-85.
90. Wolff U. Effects of a randomised reading intervention study: An application of structural equation modelling. *Dyslexia: An International Journal of Research and Practice* 2011;17:295-311.
91. Blachman B, Schattschneider C, Fletcher J, Francis D, Clonan S, Shaywitz B, Shaywitz S. Effects of Intensive Reading Remediation for

- Second and Third Graders and a 1-year Follow-up. *J Educ Psychol* 2004;96:444-61.
92. Kast M, Meyer M, Vogeli C, Gross M, Jancke L. Computer-based multi-sensory learning in children with developmental dyslexia. *Restor Neurol Neurosci* 2007;25:355-69.
 93. Rashotte C, MacPhee K, Torgesen J. The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learn Disabil Q* 2001;24:119-34.
 94. Hatcher P, Hulme C, Miles JNV, Carroll JM, Hatcher J, Gibbs S, Smith G, Bowyer-Crane C, Snowling MJ. Efficacy of small group reading intervention for beginning readers with reading-delay: a randomised controlled trial. *J Child Psychol Psychiatr* 2006; 47:820-27.
 95. Gustafson S, Samuelsson S, Ronnberg J. Why Do Some Resist Phonological Intervention? A Swedish Longitudinal Study of Poor Readers in Grade 4. *Scand J Educ Res* 2000;44:145-62.
 96. Spencer S, Manis F. The effects of a fluency intervention program on the fluency and comprehension outcomes of middle-school students with severe reading deficits. *Learn Disabil Res Pract* 2010;25:76-86.
 97. Bhat P, Griffin C, Sindelar P. Phonological awareness instruction for middle school students with learning disabilities. *Learn Disabil Q* 2003; 26:73-87.
 98. Gelzheiser L, Scanlon D, Vellutino F, Hallgren-Flynn L, Schatschneider C. Effects of the Interactive Strategies Approach—Extended. A responsive and comprehensive intervention for intermediate-grade struggling readers. *Elem Sch J* 2011;112:280-306.
 99. Lovett MW, Lacerenza L, De Palma M, Frijters JC. Evaluating the efficacy of remediation for struggling readers in high school. *J Learn Disabil* 2012;45: 151-69.
 100. Tijms J. Effectiveness of computer-based treatment for dyslexia in a clinical care setting: Outcomes and moderators. *Educ Psychol* 2011;31: 873-96.
 101. Wise BW, Ring J, Olson RK. Individual differences in gains from computer-assisted remedial reading. *J Exp Child Psychol* 2000;77:197-235.
 102. Cogo-Moreira H, Andriolo RB, Yazigi L, Ploubidis GB, Brandao de Avila CR, Mari JJ. Music education for improving reading skills in children and adolescents with dyslexia. *The Cochrane database of systematic reviews* 2012;8:CD009133.
 103. Bayles M. *Professional Ethics*. Belmont: Wadsworth, 1989.
 104. Lärarförbundet och Lärarnas Riksförbund, 2001. <http://www.lr.se/duidinyrkesroll/yrkesetik/studieoch-yrkesvagledaresyrkesetik/deetiskarikt-linjerna.4.12dcace212693f1e8748000-2966.html> och <http://www.lararforbundet.se/web/ws.nsf/documents/00377D31?OpenDocument>.
 105. Council for Learning Disabilities. <http://www.council-for-learning-disabilities.org/>

106. Taylor L, Hume IR, Welsh N. Labeling and self-esteem: the impact of using specific vs generic labels. *Educ Psychol* 2010;30:191-202.
107. Glazzard J. The impact of dyslexia on pupils self-esteem. *Support for learning* 2010;25:63-69.
108. Eriksson Gustavsson A-L. Det är tufft att plugga...men jag känner att jag klarar det. En studie om akademiska studier och skriftspråkliga svårigheter. LiU-PEK-R-257, Linköpings universitet, Institutionen för beteendevetenskap och lärande, 2011.
109. Eriksson Gustavsson A-L. Att hantera läskrav i arbetet. Om inustringarbetare med läs- och skrivsvårigheter. *Linköping Studies in Education and Psychology* nr 88, 2002. Linköping university
110. Ingesson S. Growing up with dyslexia: Interviews with teenagers and young adults. *Sch Psychol Int* 2007;28:574-91.
111. McNulty MA. Dyslexia and the life course. *J Learn Disabil* 2003;36:363-81.
112. STAR Early Literacy (SEL). Renaissance Learning UK Ltd, 32 Harbour Exchange Square, London, E14 9GE. <http://www.renaissance.com/products/star-assessments/star-early-literacy>.
113. Hakkaart-van Roijen L, Goettsch WG, Ekkebus M, Gerretsen P, Stolk EA. The cost-effectiveness of an intensive treatment protocol for severe dyslexia in children. *Dyslexia: An International Journal of Research and Practice* 2011; 17:256-67.
114. Boyle J, McCartney E, Forbes J, O'Hare A. A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technol Assess (Winchester, England)* 2007;11: iii-iv, xi-xii, 1-139.
115. Dyslexi utreds tre år försent. *Specialpedagogik* 2012. <http://www.lararnasnyheter.se/specialpedagogik/2012/09/12/dyslexi-utreds-tre-ar-sent>.
116. WHO. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*. 2007.
117. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 2013. ISBN 978-0-89042-554-1.
118. Barns och elevers utveckling mot målen. *Svensk författningssamling*, 2010:800. Skollag, 3 kap (2010:800).
119. Pedagogisk utredning av läs- och skrivsvårigheter/dyslexi. *Specialpedagogiska skolmyndigheten (SPSM)*, 2013. <http://www.spsm.se/sv/Stod-i-skolan/Funktionsnedsattningar/Las--och-skrivsvarigheter--dyslexi/Pedagogisk-utredning-av-las--och-skrivsvarigheter-dyslexi/>
120. Adler B. *Adler Färdighetstest i Läsnig* 2010. Kognitivt Centrum, Förslöv.
121. Frylmark A. *Bedömning av språk*, B.A.S. OrdAf, Domherrevägen 16, 871 65 Härnösand. Utgivningsår 2002, 2006, 2010.

122. Magnusson E, Nauc ler, K. Bed mning av spr klig medvetenhet hos f rskolebarn och skolbarn. Pedagogisk Design, L ddek pinge, 1993.
123. Franz n L. Diagnostisering av elever med l s- och skrivsv righeter. Gleerups f rlag, Lund, 1996.
124. Herrstr m M. Djur - l stest. Ordfabriken, 1998.
125. J rpsten B. DLS bas f r skol r 1 och 2. Hogrefe Psykologif rlaget, Stockholm, 2005.
126. J rpsten B. DLS f r skol r 2 och 3 - reviderad version. Hogrefe Psykologif rlaget, Stockholm, 2013.
127. J rpsten B, Taube, K. DLS f r skol r 4–6. Hogrefe Psykologif rlaget, Stockholm, 2010.
128. J rpsten B. DLS f r skol r 7–9 och 1 i gymnasiet. Hogrefe Psykologif rlaget, Stockholm, 2002.
129. Bj rkqvist L, J rpsten B, Stadler E. DLSM – Diagnostiska l s- och skrivmaterial f r mellanstadiet. Hogrefe Psykologif rlaget, Stockholm, 1983.
130. Lundberg I, Wolff U. DUVAN. Hogrefe Psykologif rlaget, Stockholm, 2003.
131. Hemmingsson I, Olofsson  . Materialet Fonolek. LPC,  stersunds kommun,  stersund, 1994.
132. Lindahl R. Lindahls h gl sningsprov H4. Rickard Lindahl F rlag, Hagens stationsv g 6, 421 71 V stra Fr lunda, Sweden, 1954.
133. H ggstr m I. Hur l ter orden. Ing-Read AB, Fridtunagatan 19, 582 12 Link ping, Sweden. 2008.
134. Frost J, Nielsen JC. IL-basis, individuella prov. Hogrefe Psykologif rlaget, Stockholm, 1996.
135. Frost J, Nielsen JC. IL-basis, gruppprov. Hogrefe Psykologif rlaget, Stockholm, 1996.
136. Johansson MG. Kvalitativa diagnoser.  stersund: MG L s- och Skriv AB.
137. Wolff U. Lilla DUVAN. Hogrefe Psykologif rlaget, Stockholm 2010.
138. H jjen T. LOGOS. Utgivnings r under 2000-talet. www.logometrica.se
139. Johansson M. LS - Reviderad. Klassdiagnoser i l sning och skrivning f r h gstadiet och gymnasiet. Hogrefe Psykologif rlaget, Stockholm, 2004.
140. Frylmark A. L sa eller lyssna. OrdAf, Domherrev gen 16, 871 65 H rn sand, 2008.
141. Johansson M. L sEttan. Natur & Kultur, 2009.
142. Jacobsson C. L skedjor. Hogrefe Psykologif rlaget, Stockholm, 2001.
143. Allard B, Askeljung M, M hl P. L sstandarder: diagnoser i  verg ng mellan  r 3 och  r 4 relaterade till kursplanem l. Sanoma Utbildning AB, Sveav gen 56, 111 34 Stockholm, 2010.
144. Elw r  , Fridolfsson I, Samuelsson S, Wiklund C. L St. Hogrefe

- Pyskologiförlaget, Stockholm, 2009.
145. Madison S. Madisons ordavkodning. Madison undervisning, 1986-2003.
146. Madison S. Madisons ordförståelse. Madison undervisning, 1996-1999.
147. Madison S. Madisons rättskrivningsprov alla stadier. Madison undervisning, 1987-1994.
148. Madison S. Madisons tystläsning med förståelsekontroll. Madison undervisning, 1997-2009.
149. Witting M. Metod för läs- och skrivinlärning. Verti AB, 1985.
150. Johansson M. MG-kedjor. MG Läs- och skrivkonsult AB, 1998.
151. Hemmingsson I. Materialet Morfolek. LPC, Östersunds kommun, Östersund, 1999.
152. Wolff U. MiniDUVAN. Test av fonologisk medvetenhet hos förskolebarn. Hogrefe Pyskologiförlaget, Stockholm, 2013. ISBN 978-91-86393-39-7.
153. Frylmark A. Observation av språk, O.A.S. OrdAf, Domherrevägen 16, 871 65 Härnösand, 2008.
154. Magnusson E, Nauclet K. OLAF testblanketter. Pedagogisk Design, 2010. www.studentlitteratur.se
155. Olofsson Å. Materialet Ordavkodning. LPC, Östersunds kommun, Östersund, 1994.
156. Olofsson Å, Söderberg Juhlander P. Ord & bild - Bild & ord. Natur & Kultur, 2013.
157. Hemmingsson I. Materialet Ordlek LPC, Östersunds kommun, Östersund, 1996.
158. Göteborgs universitet. Ordracet, individualprov. Hargdata AB, Linköping, 1996.
159. Soegård A, Hesselholdt S, Kreiner S, Poulsen Nielsen JC. OS64 och OS120. Pedagogisk Design, 1987. www.studentlitteratur.se
160. Mårtens M. Provia, datoriserat individualprov, 15 språkliga områden. Logopedkonsult, 2008.
161. Madison S. Screening, datorprogram. Madison undervisning, 1997-2009.
162. Nielsen J, Kreiner S, Poulsen A, Soegård A. SL40 och SL60. Pedagogisk Design, 1986. www.studentlitteratur.se.
163. Ellneby Y. Språkresan - en väg till ett andra språk. Hogrefe Pyskologiförlaget, Stockholm, 2012.
164. Rockberg S, Johansson MG. Stavning L, M, H. LPC i Östersund, 1994.
165. Nilsson E, Pettersson K. Stavux - stavningsprov för vuxna. OrdAf AB, Astrid Frylmark, Domherrevägen 16, 871 65 Härnösand, 2011.
166. de Jounge-Sturesson C, H-Holmberg B, Thelberg G. SVAPP: 6-årstest av substantiv, verb, adjektiv, preposi-

- tioner i det passiva ordförrådet. Natur & Kultur, 2006.
167. Palmkvist A. Trulle: Testa språket med Trulle. Gleerups förlag, Lund, 2008.
168. Dalgaard K, Krogshøj Lund Y, Otzen E. Testbatteriet, individualprov. HODAB, 2008.
169. Taube K, Tornéus M, Lundberg I. UMESOL. Hogrefe Psykologiförlaget, Stockholm, 1984.
170. Lundberg I. Vilken bild är rätt? Natur & Kultur, 2001.
171. Lindgren S-A, Laine M. The adaptation of an adult group screening test for dyslexia into Finland - Swedish: normative data for university students and the effects of language background on test performance. *Scand J Psychol* 2007;48:419-32.
172. Wolff U, Lundberg I. A Technique for Group Screening of Dyslexia among Adults. *Ann Dyslexia* 2003; 53:324-39.
173. Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association - the American Psychological Association - and the National Council on Measurement in Education, 1999. *Standards for Educational and Psychological Testing*. Washington: American Educational Research Association.
174. Vacha-Haase T, Kogan LR, Thompson B. Sample compositions and variabilities in published studies versus those in test manuals: validity of score reliability inductions. *Educ Psychol Meas* 2000;60:509-22.
175. King T. *Assistive technology*. Boston: Allyn & Bacon, 1999. (DLC) 98008342, (OCOLC)39235545.
176. Singleton C. *Intervention for Dyslexia. A review of published evidence on the impact of specialist dyslexia teaching*. University of Hull, 2009.
177. Rose J. *Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties. An independent report from Sir Jim Rose to the Secretary of State for Children, Schools and Families June 2009*. <http://webarchive.nationalarchives.gov.uk/20130401151715/> <https://www.education.gov.uk/publications/eOrderingDownload/00659-2009DOM-EN.pdf>.
178. Simpson J, Everatt J. Reception class predictors of literacy skills. *Br J Educ Psychol* 2005;75:171-88.
179. Guerin D, Griffin J, Gottfried A, Christenson G. Concurrent validity and screening efficiency of the Dyslexia Screener. *Psychol Assess* 1993;5:369-73.
180. Everatt J. Research methods in dyslexia. *Dyslexia* 2007;13:231-3.
181. Gordon HW. Cerebral organization in bilinguals. I. Lateralization. *Brain Lang* 1980;9:255-68.
182. Lyytinen H, Ahonen T, Eklund K, Guttorm TK, Laakso ML, Leinonen S et al. Developmental pathways of children with and without familial

- risk for dyslexia during the first years of life. *Dev Neuropsychol* 2001;20: 535-54.
183. Lyytinen H, Aro M, Eklund K, Erskine J, Guttorm T, Laakso ML, et al. The development of children at familial risk for dyslexia: birth to early school age. *Ann Dyslexia* 2004;54:184-220.
184. Lyytinen H, Erskine J, Tolvanen A, Torppa M, Poikkeus A-M, Lyytinen P. Trajectories of reading development: A follow-up from birth to school age of children with and without risk of dyslexia. *Merrill Palmer Q* 2006;52:514-46.
185. Olson RK. Genes, environment, and dyslexia. The 2005 Norman Geschwind Memorial Lecture. *Ann Dyslexia* 2006;56:205-38.
186. Heimann M, Lundälv M, Tjus T, Nelson KE. Omega-Interactive sentences, a multimedia software for language exploration and play. Göteborg: Topic Data & Språkbehandling HB, 2004. (Available from: www.omega-is.com).
187. Ferreira J, Gustafson S, Rönnerberg J. COMPHOT: Computerized phonological training. Linköping: Institutionen för beteendevetenskap/ Institutet för handikappvetenskap, Linköpings universitet, 2003.
188. Kairaluoma L, Narhi V, Ahonen T, Westerholm J, Aro M. Do fatty acids help in overcoming reading difficulties? A double-blind, placebo-controlled study of the effects of eicosapentaenoic acid and carnosine supplementation on children with dyslexia. *Child Care Health Dev* 2009;35:112-19.
189. Häyrynen T, Serenius-Sirve S, Korkman M. Reading and writing test designed for and normalized in Finnish elementary school. *Psykologien Kustannus Oy*, Helsinki, Finland, 1999.
190. Lovett M, Lacerenza L, Borden Susan L, Frijters J, Steinbach K, De Palma M. Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *J Educ Psychol* 2000;92:263-83.
191. McPhillips M, Hepper PG, Mulhern G. Effects of replicating primary-reflex movements on specific reading difficulties in children: a randomised, double-blind, controlled trial. *Lancet* 2000;355:537-41.

Rapporter publicerade av SBU

Gula rapporter (2010–2014)

Dyslexi hos barn och ungdomar – tester och insatser (2014), nr 225
Arbetsmiljöns betydelse för symtom på depression och utmattningssyndrom (2014), nr 223
Tidig upptäckt av symtomgivande cancer (2014), nr 222
Omhändertagande av äldre som inkommer akut till sjukhus (2013), nr 221
Behandling av urininkontinens hos äldre och sköra äldre (2013), nr 219
Mat vid fetma (2013), nr 218
ADHD – Diagnostik och behandling, vårdens organisation och patientens delaktighet (2013), nr 217
Arbetsmiljöns betydelse för sömnstörningar (2013), nr 216
Autismspektrumtillstånd – diagnostik och insatser, vårdens organisation och patientens delaktighet (2013), nr 215
Skattning av njurfunktion (2013), nr 214
Schizofreni – läkemedelsbehandling, patientens delaktighet och vårdens organisation (2012), nr 213
Diagnostik och uppföljning av förstämningssyndrom (2012), nr 212
Implementeringsstöd för psykiatrisk evidens i primärvården (2012), nr 211
Arbetets betydelse för uppkomst av besvär och sjukdomar – nacken och övre rörelseapparaten (2012), nr 210
Godartad prostataförstoring med avflödes hinder (2011), nr 209
Treatment of Hemophilia A and B and von Willebrand Disease (2011), nr 208E
Medicinska och psykologiska metoder för att förebygga sexuella övergrepp mot barn (2011), nr 207
Blödande magsår (2011), nr 206
Tandförluster (2010), nr 204
Rotfyllning (2010), nr 203
Program för att förebygga psykisk ohälsa hos barn (2010), nr 202
Mat vid diabetes (2010), nr 201
Antibiotikaprofylax vid kirurgiska ingrepp (2010), nr 200
Behandling av sömnbesvär hos vuxna (2010), nr 199
Rehabilitering vid långvarig smärta (2010), nr 198
Triage och flödesprocesser på akutmottagningen (2010), nr 197

SBU Alert-rapporter (2010–2014)

Aromatshämmande läkemedel vid behandling av postmenopausala kvinnor med tidig ER-positiv bröstcancer, nr 2014-02

Bilddiagnostik vid misstänkt prostatacancer, nr 2014-01

Behandlingstid vid borreliainfektion, nr 2013-05

Kontinuerlig subkutan glukosmätning vid diabetes, nr 2013-04

Insulinpumpar vid diabetes, nr 2013-03

Internetförmedlad psykologisk behandling vid ångest- och förstämningssyndrom, nr 2013-02

Transient elastografi vid misstänkt leverfibros och levercirros, nr 2013-01

Blodprov för tidig diagnostik av Alzheimers sjukdom, nr 2012-01

Vakuumassisterad sårbehandling, nr 2011-09

Perifert inlagd central venkateter (PICC), nr 2011-08

Analys av foster-DNA i kvinnans blod: icke-invasiv fosterdiagnostik för blodgrupps- eller könsbestämning, nr 2011-07

Molekylärdiagnostiska test för män med ökad sannolikhet för prostatacancer, nr 2011-06

Datorassisterad granskning inom mammografiscreening (CAD), nr 2011-05

Dabigatran för att förebygga stroke vid förmaksflimmer, nr 2011-04

Datortomografi för misstänkt kranskärlssjukdom, nr 2011-03

Perkutan vertebroplastik och ballongkyfoplastik vid ryggsmärta pga kotkompression som orsakats av osteoporos, nr 2011-02

Lasermedierad värmebehandling av levermetastaser, nr 2011-01

Kateterburen ablationsbehandling vid förmaksflimmer, nr 2010-06

Urinprov vid diagnostik av klamydia hos kvinnor, nr 2010-05

Hem blodtrycksmätning, nr 2010-04

Tidig och riktad ultraljudsundersökning efter fysiskt trauma, nr 2010-03

Silverförband vid behandling av kroniska sår, nr 2010-02

Cilostazol vid behandling av fönstertittarsjuka (claudicatio intermittens), nr 2010-01

Vita rapporter (2006–2014)

Immunologiska reaktioner vid användande av proteinbaserade läkemedel ESA (2013), nr 224

Immunologiska reaktioner vid användande av somatropin biosimilar (2013), nr 220
Volym och resultat (2011), nr 205

Behandling med vitamin D och kalcium (2006), nr 178
