

Rapport om projekt vedrørende tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)

Anna Steenberg Gellert og Carsten Elbro

Center for Læseforskning

Københavns Universitet

2016

Indholdsfortegnelse

Indholdsfortegnelse	2
Resumé	3
Forord	5
1. Indledning	6
2. Baggrund	7
2.1. Læsning og læsevanskeligheder.....	7
2.2. Tidlig identifikation af elever i risiko for afkodningsvanskeligheder.....	7
2.3. Dynamisk testning.....	9
3. Formål og fokus	12
4. Metode	13
4.1. Design.....	13
4.2. Deltagere.....	13
4.3. Testmaterialer.....	15
5. Resultater	26
5.1. Testresultater.....	26
5.2. Indbyrdes sammenhænge mellem testresultater.....	28
5.3. Kategorisering af elever med afkodningsvanskeligheder i slutningen af 2. klasse.....	30
5.4. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra efteråret 0. klasse.....	33
5.5. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra slutningen af 0. klasse.....	36
5.6. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra midten af 1. klasse.....	41
5.7. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra slutningen af 1. klasse.....	43
5.8. Sammenfatning og diskussion vedr. forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testning i 0. og 1. klasse.....	45
6. Mulige afgrænsninger og deres implikationer	51
6.1. Præcision i tidlig identifikation af elever i risiko for alvorlige afkodningsvanskeligheder ved forskellige niveauer af sensitivitet.....	51
6.2. Mulige øvre grænseværdier mhp. udpegning af elever i risiko for alvorlige afkodningsvanskeligheder.....	54
6.3. Det videre arbejde med afgrænsning af risikoelever.....	61
7. Undervisningsministeriets beslutninger	64
Litteraturhenvisninger	65
Bilag 1: Testoversigt	70
Bilag 2: Forudsigelse af afkodningsvanskeligheder i 1. klasse på basis af testresultater i 0. klasse	75
Bilag 3: Præcision i tidlig identifikation af risikoelever med kategorisering i 2. klasse på basis af Elbros ordlister og DVO-ordblindetesten ved forskellige niveauer af sensitivitet	78

Resumé

Denne rapport beskriver et projekt, som blev gennemført ved Center for Læseforskning på Københavns Universitet i perioden 2012-2015 og blev støttet af Trygfonden, Ministeriet for Børn, Undervisning og Ligestilling (herefter benævnt Undervisningsministeriet) samt Socialstyrelsen.

Det primære formål med projektet var at undersøge, hvorvidt såkaldte dynamiske test i 0. og 1. klasse kan bidrage til at forudsige senere alvorlige vanskeligheder med ordafkodning i læsning, når man allerede kender elevernes præstationer på traditionelle test af læseforudsætninger og læsefærdigheder. Sigtet med at anvende dynamisk testning er at afdække, hvor let en elev har ved at lære sig noget nyt inden for et bestemt område – snarere end at afdække elevens aktuelle færdighedsniveau som ved traditionel testning. Det sekundære formål med projektet var at skabe datagrundlag for nogle mulige afgrænsninger med henblik på tidlig identifikation af elever i risiko for udvikling af alvorlige vanskeligheder med ordafkodning i læsning (identifikation af de skrevne ord).

Denne rapport fremlægger resultater af en langtidsundersøgelse med 158 elever fra 0. til 2. klasse, som udgjorde den centrale del af projektet. Rapporten redegør for, hvorvidt man på basis af testning gennemført på langtidsundersøgelsens fire første testtidspunkter (hhv. efteråret 0. klasse, slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse) kunne forudsige alvorlige vanskeligheder med ordafkodning i læsning hos eleverne i slutningen af 2. klasse.

Elevernes færdigheder i *slutningen af 2. klasse* blev vurderet med tre forskellige testmaterialer. For det første blev eleverne testet med et materiale til identifikation af elever i risiko for dysleksi fra Dansk Videnscenter for Ordblindhed. For det andet blev der anvendt højtlesning af nogle lister med rigtige ord og nye ord (vrøvleord), som har indgået i tidligere forskningsprojekter og er alment udbredte i praksis. For det tredje gennemførte eleverne på Undervisningsministeriets opfordring som noget ekstra den it-baserede ordblindetest, som Undervisningsministeriet udsendte tidligere i 2015. Den aktuelle undersøgelse viste, at der var en høj grad af overensstemmelse mellem elevernes præstationer på disse tre testmaterialer i slutningen af 2. klasse.

I *efteråret 0. klasse* deltog eleverne i test hentet fra det sprogvurderingsmateriale, som er obligatorisk for børnehaveklasseelever og som bl.a. indeholder opgaver mhp. vurdering af elevernes lydlige kompetencer. Eleverne gennemførte desuden flere andre test, herunder en traditionel test af sproglydsopmærksomhed med fokus på identifikation af den første lyd i ord samt en supplerende dynamisk test af sproglydsopmærksomhed. Undersøgelsen viste, at elevernes testresultater inden for området lydlige kompetencer på det obligatoriske sprogvurderingsmateriale i kombination med den traditionelle test af sproglydsopmærksomhed i efteråret 0. klasse gav den bedst mulige forudsigelse af, hvilke elever der i slutningen af 2. klasse ville udvise alvorlige afkodningsvanskeligheder. Den supplerende dynamiske test af sproglydsopmærksomhed kunne i efteråret 0. klasse ikke derudover bidrage til en mere præcis forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse.

I *slutningen af 0. klasse* var en dynamisk test af afkodning den bedste af de anvendte test til at forudsige elevernes afkodningsfærdigheder i slutningen af 2. klasse. Dette gjaldt ikke blot for elever med dansk som modersmål, men også for elever med dansk som andetsprog, hvor det ellers ofte er ekstra vanskeligt at vurdere, om de er i risiko for udvikling af afkodningsvanskeligheder. I den dynamiske afkodningstest blev eleverne præsenteret for et lille, konstrueret alfabet i form af tre nye symboler med tilhørende sproglyde. Eleverne skulle i løbet af testen forsøge at lære at læse nye ord (vrøvleord) skrevet med det konstruerede alfabet og kunne undervejs få hjælp i form af gentagelser og feedback. Selv når man i forvejen kendte elevernes resultater på den slags test, som man traditionelt bruger til at udpege

elever i risiko for udvikling af afkodningsvanskeligheder, bidrog den dynamiske afkodningstest med ekstra forudsigelsesværdi. Undersøgelsen viste altså, at hvis man allerede i slutningen af 0. klasse forsøger at udpege de elever, der har størst risiko for at udvise alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, bliver udpegningen mere sikker, hvis man også inddrager elevernes præstationer på den dynamiske afkodningstest. Samtidig viste undersøgelsen imidlertid også, at den dynamiske afkodningstest ikke gjorde alle de traditionelle test overflødige. Forudsigelsen af, hvorvidt eleverne i slutningen af 2. klasse ville udvise alvorlige afkodningsvanskeligheder eller ej, blev således mere præcis, hvis man ud over den dynamiske afkodningstest i slutningen af 0. klasse også inddrog en traditionel test af bogstavkendskab.

Når man i *midten af 1. klasse* anvendte elevernes resultater på traditionelle afkodningstest med højtlesning af lister med rigtige ord og nye ord (vrøvleord) i kombination med deres resultater på en traditionel bogstavkendskabstest, kunne man i langt hovedparten af tilfældene korrekt forudsige, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse.

I *slutningen af 1. klasse* gav traditionelle afkodningstest med højtlesning af lister med rigtige ord og nye ord i sig selv en meget høj grad af forudsigelse, og supplerende test kunne på dette tidspunkt ikke bidrage til en mere præcis forudsigelse af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse.

Resultaterne af langtidsundersøgelsen danner grundlag for anbefalingerne i *Vejledning til tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)*, som sammen med tilhørende testmateriale stilles til rådighed for skolerne gennem Undervisningsministeriet. Ministeriet har fastlagt de konkrete grænseværdier på de test, som i denne vejledning anbefales til brug for udpegning af risikoelever.

Forord

Denne rapport beskriver en langtidsundersøgelse af elever fra 0. til 2. klasse, som blev gennemført ved Center for Læseforskning på Københavns Universitet i perioden 2012-2015. Resultaterne af denne langtidsundersøgelse danner grundlag for anbefalingerne i *Vejledning til tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)*.

Forsknings- og udviklingsprojektet er finansieret primært af Trygfonden og Undervisningsministeriet. Desuden har vi modtaget støtte fra Socialstyrelsen (Videnscenter for Handicap, Hjælpe midler og Socialpsykiatri), som har bevilget arbejdstid til projektrelaterede koordinerende og formidlende aktiviteter udført af flere af styrelsens konsulenter (Mia Finnemann Schultz, Stine Weiersøe Villadsen og Birgit Dilling Jandorf). Vi takker for al denne støtte, som har gjort det muligt at gennemføre projektet.

I langtidsundersøgelsen deltog fra hhv. Herlev, København og Næstved kommuner følgende seks skoler: Blågård Skole, Kildegårds skolen, Sydbyskolen, Tingbjerg Heldagsskole, Tove Ditlevsens Skole og Ålholm Skole. Der rettes en stor tak til alle de elever, som deltog i langtidsundersøgelsen. Også tak til Herlev Byskole, som løbende har deltaget i pilotafprøvning af test før afviklingen af testningen i hovedundersøgelsen. Vi takker endvidere alle forældre, lærere, børnehaveklasseledere, skoleledere, konsulenter og vejledere, som har hjulpet ved at give os tilladelse til at teste eleverne og/eller bistået med at overvinde praktiske udfordringer. Ligeledes takkes projektets følgegruppe bestående af læsekonsulenter fra de involverede kommuner, læsevejledere fra de involverede skoler samt repræsentanter fra hhv. Undervisningsministeriet, Videnscenter for Handicap, Hjælpe midler og Socialpsykiatri (Socialstyrelsen) samt Nationalt Videnscenter for Læsning: Mette Anker, Annette Breinholt, Henriette Vestergaard Christensen, Marianne Ellingsen, Lisbeth Hansen, Joan Henriksen, Lene Leinum, Karen Maass, Vanja Milanovic, Charlotte Møller, Trine Nobelius, Lis Pøhler, Lene Rasmussen og Line Knudsen Spang.

Desuden takker vi Hanne Trebbien Daugaard fra Center for Læseforskning, Steen Linke Larsen fra ITMEDIA samt de filmede børnehaveklassebørn for arbejdet med den instruktionsvideo, der sammen med de øvrige testmaterialer og den tilhørende vejledning stilles til rådighed for skolerne gennem Undervisningsministeriet.

Sidst, men ikke mindst, takkes alle studentermedarbejdere og praktikanter ved Center for Læseforskning for uundværlig hjælp med testning af elever og bearbejdning af testresultater.

Denne projektrapport er udarbejdet af undertegnede i samarbejde med Carsten Elbro fra Københavns Universitet. Den første version af rapporten samt de anbefalede testmaterialer med tilhørende vejledning blev som planlagt afleveret til Undervisningsministeriet i slutningen af 2015. Ministeriet tog i sommeren 2016 de nødvendige politiske beslutninger og fastlagde på grundlag af de rapporterede resultater fra langtidsundersøgelsen nogle vejledende afgrænsninger til brug for udpegning af risikoelever. Herefter blev de ministerielle beslutninger indarbejdet i kapitel 7 i projektrapporten og i *Vejledning til tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)*.

Anna Steenberg Gellert

Center for Læseforskning
Københavns Universitet

1. Indledning

Denne rapport beskriver en dansk langtidsundersøgelse af elever fra 0. til 2. klasse, som blev gennemført ved Center for Læseforskning på Københavns Universitet. Der fokuseres i denne rapport på at undersøge, hvor godt man kan forudsige afkodningsvanskeligheder hos elever i slutningen af 2. klasse på basis af testresultater hos de samme elever på forskellige tidligere tidspunkter i indskolingen med særligt henblik på at belyse, om såkaldt dynamisk testning kan bidrage til forudsigelsen, når man i forvejen kender elevernes resultater på den type test, man traditionelt bruger til at udpege elever i risiko for udvikling af afkodningsvanskeligheder (herunder ordblindhed).

Rapporten er opbygget som følger: I kapitel 2 redegør vi for den praktiske og forskningsmæssige baggrund for langtidsundersøgelsen og giver således en indføring i områderne læsning og læsevanskeligheder, problematikken omkring tidlig identifikation af børn i risiko for udvikling af afkodningsvanskeligheder samt dynamisk testning. I kapitel 3 beskriver vi undersøgelsens formål og fokus. I kapitel 4 redegør vi for undersøgelsens metode, herunder design, deltagere og testmaterialer. I kapitel 5 fremlægger vi hovedresultaterne, idet vi redegør for forudsigelsesværdien af udvalgte test gennemført på langtidsundersøgelsens fire første testtidspunkter (hhv. efteråret 0. klasse, slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse) i forhold til forekomst af alvorlige afkodningsvanskeligheder hos deltagerne på langtidsundersøgelsens femte og sidste testtidspunkt (slutningen af 2. klasse). I kapitel 6 gives der på basis af resultaterne af undersøgelsen eksempler på mulige afgrænsninger og demonstration af disse afgrænsningers implikationer. Kapitel 7 indeholder et resumé af Undervisningsministeriets beslutninger, herunder de vejledende afgrænsninger, som ministeriet fastlagde på grundlag af resultaterne fra langtidsundersøgelsen og som indgår i *Vejledning til tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)*.

I den ovenfor skitserede del af rapporten inddrages af hensyn til overskueligheden i fremstillingen kun udvalgte test, som deltagerne i langtidsundersøgelsen har gennemført. Bilag 1 indeholder imidlertid en fuldstændig oversigt over de anvendte test i undersøgelsen.

Der lægges i denne rapport størst vægt på at redegøre for, hvorvidt man på basis af test i hhv. 0. og 1. klasse kan forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse (altså ved langtidsundersøgelsens sluttidspunkt), da elevernes læseudvikling på dette tidspunkt i undersøgelsen er nået længst og giver bedst grundlag for at vurdere, om der er tale om vedvarende afkodningsvanskeligheder. Bilag 2 giver dog en kortfattet redegørelse for, hvorvidt man på basis af langtidsundersøgelsens test i 0. klasse kan forudsige afkodningsvanskeligheder i 1. klasse, samt henvisninger til artikler, hvor man kan læse nærmere om dette.

Det skal understreges, at denne rapport *ikke* træffer beslutninger. Den giver en del af et *empirisk grundlag* for at træffe beslutninger vedr. afgrænsninger af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder. At fastlægge disse afgrænsninger er *ikke* en forskningsmæssig opgave, men derimod en politisk beslutning, som derfor blev overladt til Undervisningsministeriet.

2. Baggrund

2.1. Læsning og læsevanskeligheder

Læsning består af to hovedkomponenter: afkodning og sprogforståelse. Afkodningen betegner de læsefærdigheder, som bruges til at identificere skrevne ord. Sprogforståelsen betegner blandt andet opfattelsen af, hvad de enkelte ord og sætninger betyder. Sprogforståelsen er i modsætning til afkodningen ikke noget særligt for læsning, for man bruger også sin sprogforståelse, når man lytter til talt sprog (jf. Elbro, 2014; Gough, Hoover & Peterson, 1996).

Det at afkode skrevne ord er det nye, børn skal lære for at kunne læse. I alfabetiske skriftsystemer som vores må begynderlæsere lære at udnytte skriftens lydprincip, altså de grundlæggende regler for, hvordan bogstaver og bogstavfølger omsættes til sproglyde. Det er netop her – ved tilegnelsen af det fundamentale i den alfabetiske skrift – at nogle børn har vanskeligheder (Elbro, 2007). Nogle børn har således svært ved at lære at læse, fordi de har svært ved at forbinde bogstaver og bogstavfølger med deres almindelige lyde. Denne grundlæggende vanskelighed med at lære at udnytte skriftens fundamentale lydprincip giver en usikker læsning især af nye ord; og den giver problemer med at stave. Det er vedvarende store vanskeligheder på dette felt, der er kernen i ordblindhed (dysleksi) (Lyon m.fl., 2003; Elbro, 2007; Elbro, 2012).

Vanskeligheder med at lære at læse påvirker ofte elevers selvopfattelse og kan fra et tidligt tidspunkt i skoleforløbet gå ud over deres motivation (Poskiparta m.fl., 2003; Swalander, 2012). Efterhånden som eleverne bliver ældre, vil basale vanskeligheder med at afkode ord begrænse forståelsen af de tekster, de skal læse, lige som store stavevanskeligheder vil gå ud over den skriftlige fremstilling. Ordblindevanskeligheder kan således få negative konsekvenser for elevers faglige udvikling i mange skolefag. Derfor er det vigtigt, at elever med forøget risiko for udvikling af sådanne vanskeligheder opdages tidligt, så der kan sættes ind med målrettede og særligt tilrettelagte indsatser, der bedst muligt kan støtte og fremme disse elevers udvikling.

Der findes andre typer læsevanskeligheder end ordblindhed. Således kan elever have vanskeligheder med tekstforståelse primært på grund af sprogforståelsvanskeligheder, fx hvis eleven har et andet modersmål end dansk og begrænset kendskab til danske ord. Elever med dansk som andetsprog er som udgangspunkt ikke i særlig risiko for at få alvorlige afkodningsvanskeligheder, men begrænsninger i dansk ordforråd og specifik baggrundsviden kan lede til alvorlige forståelsvanskeligheder og dermed til vanskeligheder med tekstforståelse (fx Droop & Verhoeven, 1998; 2003; Nielsen, 1998; Rønberg, 2014). Andre eksempler på grupper med forøget risiko for læsevanskeligheder er elever med nedsat hørelse og elever med opmærksomhedsforstyrrelser (fx ADHD). Elever med sådanne andre typer af vanskeligheder har også brug for ekstra støtte i forbindelse med læseundervisningen. Det projekt, som beskrives i denne rapport, fokuserer imidlertid på tidlig identifikation af elever med forøget risiko for alvorlige afkodningsvanskeligheder (herunder ordblindhed).

2.2. Tidlig identifikation af elever i risiko for afkodningsvanskeligheder

Man kan forebygge og reducere afkodningsvanskeligheder væsentligt gennem en forstærket tidlig indsats (fx Fien m.fl., 2015; Harn m.fl., 2008). Derfor er det relevant allerede i indskolingen at forsøge at identificere de elever, der har størst risiko for at udvikle sådanne vanskeligheder og dermed har brug for en forstærket tidlig indsats med fx mere opmærksomhed og støtte i forbindelse med læseundervisningen. En relativt tidlig forudsigelse af, hvilke elever der senere får afkodningsvanskeligheder, vil imidlertid aldrig være 100 % sikker. Dette hænger sammen med, at et af må-

lene med undervisningen netop er at *modarbejde* forudsigelsen, således at elever med svage forudsætninger får ekstra støtte og opmærksomhed for at undgå, at sådanne elever faktisk udvikler alvorlige vanskeligheder.

Når eleverne har modtaget formel læseundervisning i nogle måneder, er det relevant at afdække deres afkodningsfærdigheder ved at lade dem forsøge at højtlæse enkeltord og udtalelige nonord (vrøvleord). En ny dansk undersøgelse tyder således på, at sådan afdækning fra omkring midten af 1. klasse kan danne grundlag for en vurdering af, hvorvidt en elev er i risiko for at udvikle afkodningsvanskeligheder (Poulsen m.fl., 2015). Jo længere læseundervisningen har stået på, jo mere sikkert bliver grundlaget for vurderingen af, hvorvidt eleverne vil udvikle alvorlige afkodningsvanskeligheder eller ej. Omvendt gælder det som regel, at jo tidligere man forsøger at udpege elever med risiko for udvikling af alvorlige afkodningsvanskeligheder, jo mere usikkert bliver grundlaget for vurderingen (jf. Catts m.fl., 2009; Johnson m.fl., 2009; Nielsen & Poulsen, 2012; Poulsen m.fl., 2015). Således er det typisk forbundet med større usikkerhed at forsøge at identificere elever med forøget risiko for udvikling af afkodningsvanskeligheder før påbegyndelsen af den formelle læseundervisning end efter.

En af de væsentligste grunde til, at det er vanskeligt at udpege elever i risikogruppen før eller omkring påbegyndelsen af den formelle læseundervisning er, at traditionelle læsetest ikke i sig selv er så brugbare til tidlig identifikation af elever i risikogruppen på dette tidspunkt. Det kan naturligvis godt lade sig gøre allerede i 0. klasse at give eleverne en traditionel test af læsefærdigheder, fx ved at lade eleverne forsøge at læse nogle lister med enkeltord højt. Og elever, der allerede på dette tidspunkt kan læse mange enkeltord korrekt, er sandsynligvis ikke i risiko for senere at udvikle afkodningsvanskeligheder. Problemet er imidlertid, at en meget stor del af eleverne før og omkring påbegyndelsen af den formelle læseundervisning af gode grunde slet ikke kan læse nogen ord, og at langt fra alle ikke-læserne er i risiko for at udvikle afkodningsvanskeligheder. Det er derfor hensigtsmæssigt at forsøge at udpege en mindre gruppe af elever, som reelt kan formodes at have forøget risiko for at udvikle alvorlige afkodningsvanskeligheder og som derfor har mest brug for en forstærket indsats i forbindelse med den tidlige læseundervisning.

Da traditionelle læsetest altså er for svære til at kunne bruges til udpegning af egentlige risikoelever før og omkring påbegyndelse af læseundervisningen, forsøger man i stedet typisk at vurdere elevernes forudsætninger for udvikling af læsefærdigheder, bl.a. ved hjælp af test af elevernes bogstavkendskab og fonologiske opmærksomhed (dvs. opmærksomhed på de talte lyde i sproget). En sådan vurdering er relevant, da mange undersøgelser har vist, at der på gruppeniveau er en relativt stærk sammenhæng mellem børns præstationer på sådanne test af forudsætninger før læseundervisningen og deres senere faktiske læsefærdigheder (jf. National Early Literacy Panel, 2008). Imidlertid er denne sammenhæng mellem børns tidlige forudsætninger og deres senere faktiske læsefærdigheder langt fra fuldstændig og giver ikke på individniveau nogen særlig præcis udpegning af netop de elever, der faktisk senere udvikler afkodningsvanskeligheder. Således er der en del elever, som før eller omkring påbegyndelsen af den formelle læseundervisning klarer sig dårligt på traditionelle test af forudsætninger og altså umiddelbart kunne formodes at være i risiko for udvikling af afkodningsvanskeligheder, men som alligevel senere viser sig at udvikle upåfaldende afkodningsfærdigheder (fx Catts m.fl., 2009; Nelson, 2008; Nielsen & Poulsen, 2012; Poulsen m.fl., 2015).

I nogle tilfælde kan problemet forstærkes ved, at elever med svag sprogforståelse (fx elever med et andet modersmål end skolesproget) kan klare sig dårligt på nogle test af fx fonologisk opmærksomhed primært på grund af begrænset kendskab til de ord, der indgår i testene. I sådanne tilfælde kan elever, der faktisk har god fonologisk opmærksomhed, blive undervurderet og fejlagtigt udpeget som elever i risiko for udvikling af afkodningsvanskeligheder.

Udfordringen består i så vidt muligt tidligt at identificere de elever, der faktisk har forøget risiko for at udvikle afkodningsvanskeligheder, *uden* at man samtidig fejlagtigt udpeger en stor gruppe risikoelever, som viser sig at få en problemfri læseudvikling. Der er således to aspekter, der er vigtige, når man vurderer en test (eller et testbatteri), som skal anvendes til at forudsige, hvorvidt elever har forøget risiko for at udvikle vanskeligheder. Det er

- 1) testens *sensitivitet*, dvs. hvor god testen er til at udpege elever, der rent faktisk får vanskeligheder
- 2) testens *specificitet*, dvs. hvor god testen er til at udpege elever, der *ikke* får vanskeligheder.

Man kan øge sensitiviteten i identifikationen af risikoelever ved på en given test at sænke kriteriet for, hvornår elever er i risikogruppen. Så får man naturligvis udpeget flere elever som risikoelever og er dermed mere sikker på, at man ikke overser ret mange af de elever, der faktisk har mest brug for ekstra støtte og opmærksomhed. Men det vil typisk have den omkostning, at flere elever samtidig fejlagtigt udpeges til risikogruppen, således at der bliver flere såkaldte ”falsk positive” (dvs. falske alarmer), som ikke får vanskeligheder (jf. Nielsen & Poulsen, 2012; Poulsen m.fl., 2015). Der gives senere i rapporten på basis af resultater fra den aktuelle langtidsundersøgelse konkrete eksempler på denne problematik (jf. afsnit 6.1 og 6.2). Hvis (for) mange elever udpeges som risikoelever, bliver det vanskeligt at tildele ret mange ekstra ressourcer til netop de elever, der har størst behov.

Forholdet mellem sensitivitet og specificitet kan vises i en såkaldt receiver operating characteristics (ROC) kurve, der kombinerer værdierne for sensitivitet og specificitet ved alle tænkelige cut-offs (Metz, 1978). Høje værdier af sensitivitet og specificitet medfører et stort areal under kurven (area under the curve = AUC). Arealet af denne kurve, som angives på en skala fra 0 til 1, har de seneste år vundet indpas som et samlet mål for, hvor præcise test (evt. kombinationer af flere test) er til at forudsige fx senere læsevanskeligheder. Internationalt anbefales det, at den samlede AUC-værdi for testbatterier mhp. tidlig identifikation af elever i risiko for udvikling af læsevanskeligheder skal være på 0,80 eller derover (jf. Catts m.fl., 2009; Catts m.fl., 2015). AUC-værdier vil i denne rapport blive angivet og inddraget i forbindelse med vurderingen af præcisionen af test (og kombinationer af test) med henblik på tidlig identifikation af elever i risiko for udvikling af afkodningsvanskeligheder (jf. kapitel 5).

2.3. Dynamisk testning

Traditionelt bruger man primært såkaldt statiske test til at undersøge elevers sprog og læsning. For eksempel kan man ved hjælp af en traditionel, statisk test af ordkendskab forsøge at afdække, hvor mange ord en elev kender (altså elevens aktuelle ordkendskabsniveau). Imidlertid er en anden type testning, dynamisk testning, ved at vinde frem. Hovedformålet med dynamisk testning er at undersøge, hvor let en elev har ved at lære noget nyt – og altså ikke elevens aktuelle færdighedsniveau som ved traditionel, statisk testning (Grigorenko & Sternberg, 1998; Grigorenko, 2009). Således har en dynamisk ordkendskabstest primært til formål at afdække, *hvor let en elev har ved at lære nye ord*, snarere end elevens aktuelle ordkendskabsniveau (fx Gellert & Elbro, 2013). Tilsvarende kan man med en traditionel, statisk test af læsning eksempelvis undersøge, hvor mange ord en elev kan læse korrekt (altså elevens aktuelle læsefærdighedsniveau), hvorimod formålet med en dynamisk læsetest kan være at undersøge, *hvor let en elev har ved at lære at læse*.

I et hyppigt anvendt dynamisk testformat hjælpes eleven ved fejlsvar hen imod løsning af opgaven i form af feedback eller cues. Det opgøres så, hvor megen hjælp eleven skal have for at kunne løse opgaven. En elevs præstation på en dynamisk test antages at være mindre afhængig af hans eller hendes forudgående erfaringer og tidligere erhvervede færdigheder end præstationen på en traditionel, statisk test. Dynamiske test antages i højere grad at give mål for læ-

ringspotentiale og dermed at kunne forudsige elevers fremtidige udvikling inden for testens område (Sternberg & Griorenko, 2002).

Resultaterne af flere udenlandske langtidsundersøgelser tyder på, at inddragelse af dynamiske test af fonologisk opmærksomhed (sproglydsopmærksomhed) i begyndelsen af børnehaveklassen kan give bedre muligheder for at identificere de børn, der omkring et halvt år senere viser tegn på vanskeligheder i den tidlige fase i læseudviklingen (fx Bridges & Catts, 2011; Spector, 1992). I modsætning hertil viste en ny amerikansk langtidsundersøgelse, der har fulgt elevernes læseudvikling længere frem til slutningen af 1. klasse, at en dynamisk test af fonologisk opmærksomhed i begyndelsen af børnehaveklassen faktisk ikke kunne bidrage selvstændigt til at forudsige, hvilke elever der i slutningen af 1. klasse ville udvise tegn på afkodningsvanskeligheder, når man først havde taget højde for elevernes tidlige resultater på traditionelle test af bl.a. fonologisk opmærksomhed og bogstavkendskab (Catts m.fl., 2015). I den langtidsundersøgelse, der beskrives i denne rapport, blev det i en dansk kontekst afdækket, om en dynamisk test af fonologisk opmærksomhed kan bidrage til allerede i efteråret 0. klasse at identificere de elever, der senere vil udvise alvorlige afkodningsvanskeligheder. I den danske, dynamiske test af fonologisk opmærksomhed blev eleven (ved fejlsvaret) gradvis hjulpet hen imod løsning af opgaven med gradvist øget støtte i form af et systematisk hierarki af feedback og støtte. En mere detaljeret beskrivelse af testen gives i afsnit 4.3.1.2 ”Dynamisk test af fonologisk opmærksomhed (forlydsidentifikation)”.

Endvidere har udenlandske forskergrupper udviklet dynamiske test af afkodning og undersøgt, om sådanne test kunne bidrage til at forudsige læseudviklingen hos børn. I en amerikansk langtidsundersøgelse gennemførte børnehaveklassebørn før læseundervisningens påbegyndelse både dynamiske afkodningstest og traditionelle, statiske test af læsning og læseforudsætninger og blev testet igen med traditionelle læsetest i slutningen af 1. klasse (Petersen, Allen & Spencer, 2016). Undersøgelsen viste, at i børnehaveklassen var de dynamiske afkodningstest betydeligt bedre til at forudsige læsevanskeligheder i slutningen af 1. klasse end de traditionelle, statiske test var, både blandt elever med engelsk som modersmål og engelsk som andetsprog. Hvis man tog højde for børnenes resultater på de dynamiske afkodningstest, bidrog de traditionelle test således slet ikke med nogen yderligere forudsigelsesværdi. Disse resultater skal dog ses i lyset af, at de konkrete traditionelle test, som de dynamiske test blev sammenlignet med i Petersen og kollegers undersøgelse, faktisk ikke var særlig gode til at forudsige senere læsevanskeligheder, hvilket også andre undersøgelser har påvist (fx Catts m.fl., 2009; Nelson, 2008). Det er uvist, om de dynamiske afkodningstest ville vise sig lige så overlegne, hvis disse dynamiske test blev sammenlignet med traditionelle test med en bedre prædiktionsværdi end dem, Petersen m.fl. anvendte i deres undersøgelse. I en anden amerikansk langtidsundersøgelse gennemførte eleverne først i efteråret i 1. klasse en dynamisk afkodningstest foruden en række statiske test af færdigheder og forudsætninger, som tidligere har vist sig at være gode prædiktorer af læseudviklingen (Fuchs m.fl., 2011). I slutningen af 1. klasse blev børnenes læsefærdigheder afdækket ved hjælp af traditionelle, statiske test. Undersøgelsen viste, at efter der var taget højde for børnenes resultater på alle de statiske test i efteråret 1. klasse, kunne resultaterne på den dynamiske afkodningstest bidrage ganske lidt, men dog signifikant til forudsigelse af forskellene i børnenes læsefærdigheder i slutningen af 1. klasse. I den langtidsundersøgelse, der beskrives i denne rapport, blev det i en dansk kontekst afdækket, om en dynamisk test af afkodning efter skrap kontrol med stærke, traditionelle prædiktorer kunne bidrage til hhv. i slutningen af 0. klasse og midten af 1. klasse at forudsige, hvilke elever der senere ville udvise alvorlige afkodningsvanskeligheder.

I den aktuelle danske langtidsundersøgelse tog vi udgangspunkt i en dynamisk afkodningstest, som vi tidligere har udviklet ved Center for Læseforskning på Københavns Universitet (Daugaard, Elbro & Gellert, 2011a). I denne test præsenteres deltageren for et lille, konstrueret alfabet bestående af tre nye bogstaver i form af symboler (ϣ, ϕ og ϗ) med tilhørende sproglyde (/s/, /m/ og /a/). I den første deltest skal deltageren lære og huske koblingen mellem disse symboler og deres lyde. Derefter skal deltageren i den anden deltest lære at læse helt simple nonord skrevet med det konstruerede alfabet. Den lydige syntese bliver illustreret visuelt ved hjælp af bogstavbrikker, der gradvist flyttes sammen. Deltageren får korrigerende feedback på sin læsning. Hvis deltageren lærer at læse de simple ord inden for et vist antal forsøg, går man videre til læsning af lidt længere og sværere nonord i den tredje deltest. En mere detaljeret beskrivelse af testen gives i afsnit 4.3.2.1. Denne dynamiske afkodningstest blev oprindeligt konstrueret for at begrænse betydningen af deltagerens tidligere læseerfaringer og generelle sprogfærdigheder i forbindelse med vurdering af afkodningsfærdigheder. Det er nemlig typisk vanskeligt at skelne mellem egentlige afkodningsvanskeligheder og mere generelle sproglige vanskeligheder ved undersøgelse af eventuel ordblindhed blandt deltagere med et begrænset kendskab til det sprog, testningen foregår på (Everatt m.fl., 2010; Gellert, 2009). Derfor er der en tendens til, at almindelige læsetest overdiagnosticerer ordblindhed blandt deltagere med fx dansk som andetsprog (Dansk Videnscenter for Ordblindhed, 2006; Elbro, Daugaard & Gellert, 2012). Det konstruerede alfabet i den dynamiske afkodningstest blev således valgt for at mindske indflydelsen af deltagerens erfaring med læsning på det latinske alfabet. Desuden bruges der i testen kun non-verbale (ordløse) instruktioner for at minimere risikoen for, at forskelle i deltagernes sprogforståelse får betydning for testresultatet. På disse punkter adskiller vores dynamiske afkodningstest sig væsentligt fra de dynamiske afkodningstest, som andre forskergrupper har udviklet og afprøvet (Fuchs m.fl., 2011; Petersen & Gillam, 2015; Petersen, Allen & Spencer, 2016).

I en første undersøgelse blandt voksne med dansk som enten modersmål eller andetsprog fandt vi, at vores dynamiske afkodningstest både kan bruges til at identificere ordblinde og samtidig er mere uafhængig af generelle sprogfærdigheder og tidligere læseerfaringer end traditionelle, statiske læsetest (Daugaard, Elbro & Gellert, 2011b; Elbro, Daugaard & Gellert, 2012). Blandt de voksne med dansk som modersmål indgik en gruppe danskere, som allerede var diagnosticeret som ordblinde ved hjælp af gængse ordblindetest, samt en gruppe danskere uden ordblindvanskeligheder. Resultaterne af afprøvningen viste, at den dynamiske afkodningstest i høj grad kunne skelne mellem de deltagende ordblinde og ikke-ordblinde voksne med dansk som modersmål og således var en gyldig ordblindetest. En række mere traditionelle, statiske sprog- og læsetest indgik også i afprøvningen, så de voksnes resultater på den dynamiske test kunne sammenlignes med resultaterne på disse test. Det viste sig, at resultaterne på den dynamiske afkodningstest blandt de voksne med dansk som andetsprog var mindre afhængige af deres danske ordforråd og deres tidligere skolegang, end disse voksnes resultater på traditionelle læsetest var. Den første undersøgelse tydede altså på, at den dynamiske afkodningstest opfyldte sit formål og er bedre egnet til at udpege ordblinde blandt voksne indvandrere, end traditionelle læsetest er.

På baggrund af resultaterne fra den første afprøvning af den dynamiske afkodningstest var der grund til at formode, at denne test også kunne være anvendelig i indskolingen, ikke blot i forhold til elever med dansk som modersmål, men også med elever fra sproglige minoritetsgrupper. Som nævnt består kernen i ordblindhed netop af vedvarende store vanskeligheder med at udnytte skriftens lydprincip. Hvis elever – uanset sproglig baggrund – har meget svært ved at lære at sætte bogstavernes lyde sammen til nye ord i den dynamiske afkodningstest, kan det derfor antages, at disse elever også vil udvise fundamentale vanskeligheder med afkodning på længere sigt.

3. Formål og fokus

Formålet med den her rapporterede del af projektet var at undersøge, i hvilket omfang man på basis af elevers testresultater i 0. og 1. klasse kan forudsige afkodningsvanskeligheder (herunder ordblindhed) i slutningen af 2. klasse – med særligt fokus på at belyse, hvorvidt dynamiske test i 0. og 1. klasse kan bidrage til at forudsige senere afkodningsvanskeligheder, når man allerede kender elevernes præstationer på traditionelle test af læseforudsætninger og læsefærdigheder.

Det *primære formål* med undersøgelsen var at bidrage til svar på følgende forskningsspørgsmål:

1. Kan en dynamisk test af fonologisk opmærksomhed i efteråret 0. klasse bidrage til at forudsige senere afkodningsvanskeligheder, også når man allerede kender elevernes præstationer på traditionelle test af læseforudsætninger i efteråret 0. klasse?
2. Kan en dynamisk afkodningstest i slutningen af 0. klasse bidrage til at forudsige senere afkodningsvanskeligheder, også når man allerede kender elevernes præstationer på traditionelle test af læsning og læseforudsætninger i slutningen af 0. klasse? Og hvis dette er tilfældet og man i slutningen af 0. klasse bruger en dynamisk afkodningstest til at forudsige senere afkodningsvanskeligheder, styrker det så forudsigelsen, at man også inkluderer traditionelle test?
3. Kan en dynamisk afkodningstest i midten af 1. klasse bidrage til at forudsige afkodningsvanskeligheder i slutningen af 2. klasse, også når man allerede kender elevernes præstationer på traditionelle test i midten af 1. klasse?

Det *sekundære formål* med projektet var at skabe datagrundlag for nogle mulige afgrænsninger med henblik på tidlig identifikation af elever i risiko for udvikling af afkodningsvanskeligheder (herunder ordblindhed). På Undervisningsministeriets opfordring redegør vi i denne rapport for forudsigelsesværdien af de vigtigste (i praksis de mest prædiktive) test på samtlige fire tidlige testtidspunkter (dvs. både efteråret 0. klasse, slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse), når målet er at forudsige afkodningsvanskeligheder i slutningen af 2. klasse. Dette indebærer altså, at forudsigelsesværdien af alle relevante test (såvel traditionelle som dynamiske test) undersøges, også i slutningen af 1. klasse, hvor eleverne kun gennemførte traditionelle test.

Af hensyn til overskueligheden af rapporten behandler vi de tre ovenfor nævnte forskningsspørgsmål vedr. forudsigelsesværdien af de dynamiske test i de afsnit, der drejer sig om resultaterne for de testtidspunkter, hvor de pågældende dynamiske test er gennemført. Det vil sige, at spørgsmål 1 vedr. forudsigelsesværdien af en dynamisk test af fonologisk opmærksomhed behandles under afsnit 5.4, der handler om forudsigelse på basis af testning i efteråret 0. klasse. Tilsvarende behandles spørgsmål 2 vedr. forudsigelsesværdien af en dynamisk afkodningstest i afsnit 5.5, der drejer sig om forudsigelse på basis af testning i slutningen af 0. klasse. Endelig behandles spørgsmål 3 i afsnit 5.6, der fokuserer på forudsigelse på basis af testning i midten af 1. klasse.

Der fokuseres altså i denne rapport på at undersøge, hvor godt afkodningsvanskeligheder hos elever kan forudsiges på basis af testresultater hos de samme elever på forskellige tidligere tidspunkter i indskolingen. Der lægges i denne rapport størst vægt på at redegøre for, hvor præcist man på basis af test i hhv. 0. og 1. klasse kan forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse (altså ved langtidsundersøgelsens sluttidspunkt), da elevernes læseudvikling på dette tidspunkt i undersøgelsen er nået længst og giver bedst grundlag for at vurdere, om der er tale om

vedvarende afkodningsvanskeligheder. Det er et eksplicit ønske fra Undervisningsministeriet, at der fokuseres på forudsigelsen af *alvorlige* afkodningsvanskeligheder (herunder ordblindhed).

Rapporten redegør kun kortfattet for, hvorvidt man på basis af test i 0. klasse kan forudsige afkodningsvanskeligheder i 1. klasse, og af hensyn til overskueligheden er disse korte redegørelser placeret i bilag 2. Læsere, der måtte være interesserede i mere udførlige redegørelser for, hvorvidt man på basis af test i 0. klasse kan forudsige afkodningsvanskeligheder i 1. klasse, henvises til artikler, hvor vi har beskrevet foreløbige resultater fra langtidsundersøgelsen frem til hhv. midten af 1. klasse (Gellert & Elbro, 2014) og slutningen af 1. klasse (Gellert & Elbro, 2015a; Gellert & Elbro, 2015b).

4. Metode

4.1. Design

Vi gennemførte en langtidsundersøgelse af elever fra efteråret 0. klasse til slutningen af 2. klasse. Eleverne blev testet på følgende tidspunkter:

- Efterår 0. klasse
- Slutningen af 0. klasse
- Midten af 1. klasse
- Slutningen af 1. klasse
- Slutningen af 2. klasse

4.2. Deltagere

4.2.3. Rekruttering og udvælgelse

I foråret 2012 kontaktede Videnscenter for Handicap, Hjælpemidler og Socialpsykiatri (Socialstyrelsen) læsekonsulenter fra hhv. hovedstadskommunen (Københavns Kommune), en forstadskommune (Herlev Kommune) og en kommune i provinsen (Næstved Kommune) og spurgte, om disse kommuner ville indgå i den planlagte langtidsundersøgelse. Læsekonsulenterne tog kontakt til skoleledere i deres kommuner, og alle tre kommuner meldte tilbage, at de gerne ville være med. Der blev udvalgt én skole fra Næstved Kommune (Sydbyskolen) og én skole fra Herlev Kommune (Kildegårds skolen). Fra Københavns Kommune blev der udvalgt fire skoler (Blågård Skole, Tingbjerg Heldags skole, Tove Ditlevsens Skole og Ålholm Skole). Flere af de københavnske skoler blev valgt, fordi de havde en høj andel af elever fra sproglige minoritetsgrupper og således gav mulighed for senere at gennemføre analyser med fokus på deltagere med dansk som andetsprog. Samtlige kommende 21 børnehaveklasser på disse i alt seks folkeskoler blev meldt til projektet i foråret 2012. Efter sommerferien blev nogle af de planlagte børnehaveklasser dog slået sammen, så der endte med at indgå i alt 19 børnehaveklasser fordelt på de seks skoler som vist nedenfor:

- Sydbyskolen (Sydby, Næstved Kommune): 4 klasser
- Kildegårds skolen (Herlev Kommune): 4 klasser
- Blågård Skole (Nørrebro, Københavns Kommune): 2 klasser
- Tingbjerg Heldagsskole (Brønshøj, Københavns Kommune): 2 klasser

- Tove Ditlevsens Skole (Vesterbro, Københavns Kommune): 4 klasser
- Ålholm Skole (Valby, Københavns Kommune): 3 klasser

Vi udarbejdede et informationsbrev om projektet til forældrene til børnehaveklasseeleverne fra de seks skoler. Informationsbrevet blev efter samråd med børnehaveklasselederne og læsevejlederne på de involverede skoler oversat til 13 forskellige sprog, således at forældrene så vidt muligt kunne få brevet på deres eget sprog. I brevet blev forældrene spurgt, om de ville give tilladelse til, at deres barn deltog i projektet. Børnehaveklasselederne på de involverede skoler deltog inden sommerferien i et informationsmøde om projektet og fik desuden tilsendt skriftlige oplysninger. Børnehaveklasselederne informerede efter sommerferien forældrene til de nystartede børnehaveklasseelever om projektet, uddelte informationsbrevene og indsamlede forældretilladelser. Omkring 90 % af børnehaveklasseeleverne på de involverede skoler fik skriftlig tilladelse fra deres forældre til at deltage i projektet.

I september 2012 gennemførte Center for Læseforskning ved hjælp af gruppetest af bogstavkendskab og fonologisk opmærksomhed en screening af de børnehaveklasseelever, der havde fået skriftlig tilladelse til at deltage i projektet. Screeningen omfattede i alt 430 elever. På basis af resultaterne af screeningen udvalgte 200 børnehaveklasseelever, som efterfølgende blev testet individuelt.

Halvdelen af de 200 elever ($N = 100$) blev valgt på basis af svage screeningsresultater. Således blev omkring 5-6 elever med de svageste resultater fra hver klasse valgt til denne gruppe. Den anden halvdel af eleverne ($N = 100$) blev udvalgt tilfældigt blandt klassekammeraterne til den omtalte gruppe af elever, der klarede sig dårligt på screeningen.

Der blev således med vilje foretaget en oversampling af deltagere med formodet svage læseforudsætninger og deltagere med dansk som andetsprog. Denne oversampling blev foretaget mhp. at kunne få mere pålidelige oplysninger om netop disse undergrupper og stadig holde omfanget af undersøgelsen inden for de givne økonomiske og praktiske rammer.

4.2.4. Frafald og tilbageværende deltagere

I slutningen af 2. klasse var 33 af de 200 elever gået ud af undersøgelsen. 27 af disse 33 elever var flyttet til en anden skole, mens 5 elever skulle gå en klasse om, og 1 elev var flyttet fra sin klasse for at blive placeret i en modtageklasse (et særligt tilbud for elever, der ikke kan dansk nok til at følge undervisning i en almindelig skoleklasse). Endvidere mangler der nogle data fra 7 elever, som pga. fx langvarig sygdom eller udlandsrejse ikke blev testet på alle fem testtidspunkter, samt fra yderligere 2 elever, der nægtede at deltage i dele af testningen.

Det er kun data fra de 158 tilbageværende elever (75 drenge og 83 piger), der indgår i de nedenfor rapporterede analyser vedr. forudsigelse af afkodningsvanskeligheder i slutningen af 2. klasse på basis af testning gennemført i efteråret og slutningen af 0. klasse samt midten og slutningen af 1. klasse. 55 af disse 158 elever havde et andet modersmål end dansk. Gennemsnitsalderen var 6 år og 2 måneder ved screeningen i september i børnehaveklassen.

Af de 158 tilbageværende deltagere var 76 fra den gruppe af de oprindeligt 100 elever, der blev udvalgt på basis af svage screeningsresultater, mens 82 kom fra den gruppe, der blev tilfældigt udvalgt.

De 158 deltagere i det fulde sample udgjorde således ikke et repræsentativt udvalg af elever, hvilket heller ikke var hensigten. Men som det senere beskrives (jf. afsnit 5.3.3), anvendte vi som grundlag for beregningerne af scoregrænser for hhv. alvorlige og moderate afkodningsvanskeligheder i slutningen af 2. klasse et reduceret sample, idet vi fra analyserne fjernede de deltagere, der i efteråret 0. klasse bidrog til oversamplingen. I dette reducerede og mere repræsentative sample indgik der kun 117 deltagere (mod 158 i det fulde sample). Elevernes gruppetestresultater med

Ordlæseprøve 2 (Juul & Møller, 2010b) i slutningen af 2. klasse indikerede, at det reducerede sample på 117 elever faktisk var nogenlunde repræsentativt med hensyn til ordlæsefærdigheder, idet det gennemsnitlige antal korrekt besvarede opgaver for denne gruppe var 53,4 (med en standardafvigelse på 22,6) og dermed lå ganske tæt på normen for foråret 2. klasse på 49,5 korrekt besvarede opgaver (jf. vejledningen til *Ordlæseprøve 1-2* i Møller & Juul, 2010, s. 38) med en standardafvigelse på 21,4 (Juul, 2015).

4.3. Testmaterialer

Deltagerne i langtidsundersøgelsen gennemførte en række test, herunder både gruppetest og individuelle test. Der henvises til bilag 1 for en fuldstændig oversigt over de anvendte test. Af hensyn til overskueligheden fokuseres der i det følgende på de test, der viste sig at være bedst til at forudsige senere afkodningsvanskeligheder og/eller på anden vis var særlig relevante for besvarelsen af forskningsspørgsmålene (jf. kapitel 3 tidligere i rapporten). Hovedparten af de test, der beskrives nærmere nedenfor og indgår i de senere analyser, blev afviklet individuelt med deltagerne på skolerne ved testmedarbejdere fra Center for Læseforskning. Den eneste undtagelse er materialet *Sprogvurdering af børn i treårsalderen, inden skolestart og børnehaveklassen*, som indeholder både gruppetest og individuelle test og som i efteråret 0. klasse blev afviklet ved de involverede klassers børnehaveklasseledere i overensstemmelse med normal praksis på skolerne.

Tabel 1 indeholder en oversigt over de udvalgte test, der beskrives nærmere i de følgende afsnit og indgår i de senere analyser.

Testtidspunkt	Test
Efterår 0. klasse	Forlydsidentifikation – statisk
	Forlydsidentifikation – dynamisk
	Bogstavkendskab
	Sprogvurdering af børn i treårsalderen, inden skolestart og børnehaveklassen
Slutningen af 0. klasse	Dynamisk afkodning A
	Afkodningspræcision – Elbros ordlister med rigtige ord
	Bogstavkendskab
	Forlydsidentifikation – statisk
	Fonologisk syntese
	Ordforråd
Midten af 1. klasse	RAN: Hurtig automatisk benævnelse af genstande
	Dynamisk afkodning B
	Afkodningspræcision – Elbros ordlister med rigtige ord og nonord
	Afkodningseffektivitet – Elbros ordlister med rigtige ord og nonord
	Bogstavkendskab
	Fonologisk syntese
Slutningen af 1. klasse	RAN: Hurtig automatisk benævnelse af cifre
	Afkodningspræcision – Elbros ordlister med rigtige ord og nonord
	Afkodningseffektivitet – Elbros ordlister med rigtige ord og nonord
Slutningen af 2. klasse	RAN: Hurtig automatisk benævnelse af cifre
	Afkodningspræcision 1 – Elbros ordlister med rigtige ord og nonord
	Afkodningspræcision 2 – lister med rigtige ord og nonord fra DVO-ordblindetest
	Afkodningseffektivitet 1 – Elbros ordlister med rigtige ord og nonord
	Afkodningseffektivitet 2 – UVM-ordblindetest
	Ordforråd – UVM-ordblindetest
Fonologisk subtraktion – DVO-ordblindetest	

Tabel 1. Oversigt over de centrale test i undersøgelsen opdelt efter testtidspunkt

4.3.1. Udvalgte testmaterialer i efteråret 0. klasse

4.3.1.1. Statisk test af fonologisk opmærksomhed (forlydsidentifikation)

En test af forlydsidentifikation med multiple choice-format blev udviklet til projektet. Testen havde samme format som i tidligere undersøgelser, hvor den har vist god prædiktionsværdi for den første læseudvikling (fx Elbro, Borstrøm & Petersen, 1998). Testtypen indgår i flere publicerede, danske testmaterialer til afdækning af fonologisk opmærksomhed i børnehaveklassen (fx Borstrøm & Petersen, 2004). I den anvendte test blev eleven i hver opgave præsenteret for et ark med seks illustrationer af hyppige ord (fx flag, bamse, dukke, snegl, taske, dreng), som testlederen benævnte (jf. figur 1 nedenfor).¹ Testlederen udtalte derefter en lyd, fx /t/, og bad eleven om at pege på billedet af det ord, som begynder med den pågældende lyd (jf. figur 1 nedenfor). Efter to øveopgaver blev eleven præsenteret for i alt 15 opgaver. Scoren var antal korrekt besvarede opgaver (max. 15). Overensstemmelsen mellem deltestens enkelte opgaver (den interne opgavehomogenitet) var 0,82 (vurderet ved Cronbachs Alfa) og dermed høj.



Figur 1. Eksempel på opgave i test af forlydsidentifikation.

”Flag, bamse, dukke, snegl, taske, dreng...
/t/...Hvad for et ord starter med /t/?”

4.3.1.2. Dynamisk test af fonologisk opmærksomhed (forlydsidentifikation)

Til projektet blev der udviklet en dynamisk test i tilknytning til den ovenfor beskrevne statiske test af fonologisk opmærksomhed. Umiddelbart efter den statiske test fik eleven i en dynamisk test igen de opgaver, som han/hun ikke besvarede korrekt i den statiske test af fonologisk opmærksomhed. I den dynamiske test indgik følgende fire niveauer af hjælp:

1. Eleven bliver bedt om selv at benævne billederne og igen forsøge at finde det ord, der starter med fokus-lyden i den pågældende opgave (fx /t/).
2. Eleven bliver bedt om at gentage fokus-lyden efter testlederen og påny prøve at finde det ord, der starter med lyden.

¹ I en umiddelbart forudgående test af stavelsesidentifikation var eleven blevet præsenteret for de samme billeder, blevet bedt om at benævne dem og om nødvendigt blevet korrigeret af testlederen. Derefter blev eleven bedt om at udpege det ord, der begynde med en bestemt stavelse (fx *bam-*). Testen af stavelsesidentifikation inddrages ikke i de analyser, der rapporteres i denne artikel, da opmærksomhed på stavelser, rim og andre større segmenter generelt ikke er så stærk en prædikator af børns læsefærdigheder som opmærksomhed på fonemer (fx Melby-Lervåg, 2012)

3. Testlederen siger fokus-lyden foran alle seks ord og beder eleven om igen at forsøge at finde det ord, der starter med lyden.
4. Testlederen giver svaret.

Hvis eleven ikke løser opgaven med hjælp på niveau 1, gås der videre til niveau 2. Hvis eleven heller ikke her kan løse opgaven med den tilhørende hjælp, fortsættes der til niveau 3 – og så fremdeles. Hvis eleven derimod løser opgaven med hjælp på niveau 1, 2 eller 3, droppes de efterfølgende niveauer, og der gås i stedet videre til den næste fejlbesvarede opgave fra den statiske test af fonologisk opmærksomhed. Opgaverne blev scoret ved hjælp af følgende pointsystem:

- 4 point = eleven besvarede opgaven korrekt allerede i den statiske test og behøvede derfor ingen hjælp
- 3 point = korrekt svar efter hjælp på niveau 1 i den dynamiske test
- 2 point = korrekt svar efter hjælp på niveau 2 i den dynamiske test
- 1 point = korrekt svar efter hjælp på niveau 3 i den dynamiske test
- 0 point = eleven svarede ikke selv korrekt

Den maksimale score var 60 point.

4.3.1.3. Bogstavkendskab

Til afdækning af elevernes bogstavkendskab anvendtes en test af bogstavbenævnelse fra materialet *Læseevaluering på begyndertrinet* (Borstrøm & Petersen, 2004). Testen har to dele – én med alle de store bogstaver og én med alle de små bogstaver. I begge dele udpeger testlederen bogstaverne ét for et i læseretningen og spørger eleven, hvad bogstavet hedder. Scorerne var antal korrekt benævnte store henholdsvis små bogstaver. Opgavehomogeniteten var i denne undersøgelse henholdsvis 0,92 og 0,91 (Cronbachs Alfa). For at styrke pålideligheden af målet for bogstavkendskab blev resultaterne på de to deltest (som indbyrdes korrelerede 0,84) slået sammen til én score, hvor scorerne på de to deltest blev vægtet ligeligt.

4.3.1.4. Sprog vurderingsmateriale

Materialet *Sprog vurdering af børn i treårsalderen, inden skolestart og i børnehaveklassen* (herefter benævnt Sprog vurderingsmaterialet) er udviklet af Center for Børnesprog ved Syddansk Universitet og udsendt af Kommunernes Landsforening, Undervisningsministeriet, Socialministeriet og Finansministeriet til brug for den lovpligtige sprog vurdering. I den aktuelle langtidsundersøgelse indgik fra dette materiale de redskaber, som er udviklet til elever i begyndelsen af børnehaveklassen og som normalt blev administreret af børnehaveklasselederne (herunder papirbaserede gruppetest og individuelle test samt it-baserede forældre rapporter).

Sprog vurderingsmaterialet giver ifølge den tilhørende vejledning (Socialstyrelsen, 2010) grundlag for vurdering af børn inden for følgende fire dimensioner af sproget:

- 1) Produktivt talesprog
- 2) Receptivt talesprog
- 3) Lydlige kompetencer
- 4) Kommunikative kompetencer

På basis af alle resultaterne beregnes der for hvert barn en samlet score samt for hver af de fire sproglige dimensioner en kombineret score, som udregnes af et it-system og præsenteres i en resultatrapport som percentilscore.

I de analyser, der senere beskrives her i denne rapport om den aktuelle langtidsundersøgelse, inddrages elevernes samlede score på Sprogvurderingsmaterialet samt elevernes kombinerede score på den tredje sproglige dimension, lydlige kompetencer, da denne dimension specielt peger frem mod senere afkodningsfærdigheder (jf. Socialstyrelsen, 2010). Elevernes kombinerede score på lydlige kompetencer er baseret på elevernes resultater på gruppetest af bogstavkendskab og rim eller – i de tilfælde, hvor eleven også er blevet testet individuelt – på resultaterne af individuelle test af bogstavkendskab, rim, skelnen af sproglyde og hurtig navngivning. For en nærmere beskrivelse af redskaberne i Sprogvurderingsmaterialet og udregning af scorer henvises der til den vejledning til materialet, som er udsendt til skoler og dagtilbud (Socialstyrelsen, 2010), den metodiske oversigt (Bleses m.fl., 2011) og et teknisk bilag (Socialstyrelsen, 2014).

I modsætning til alle de øvrige test, der indgik i vores langtidsundersøgelse, blev redskaberne i Sprogvurderingsmaterialet administreret af børnehaveklasselederne (i overensstemmelse med normal praksis i skolerne) og altså ikke af testmedarbejdere fra Center for Læseforskning. 36 % af de 158 elever, der deltog i vores langtidsundersøgelse, gennemførte både gruppetest og individuelle test fra sprogvurderingsmaterialet. De resterende 64 % gennemførte kun gruppetest.

Vi fik via Rambøll og de deltagende skoler adgang til elevernes sprogvurderingsresultater fra efteråret 0. klasse med tilladelse fra Undervisningsministeriet, Socialstyrelsen, de deltagende kommuner og elevernes forældre.

4.3.2. Testmaterialer i slutningen af 0. klasse

4.3.2.1. Dynamisk Afkodningstest A

Eleverne gennemførte en dynamisk afkodningstest, som tidligere er udsendt til voksne med dansk som andetsprog med titlen *Dynamisk ordblindetest (DOT)* (Daugaard, Elbro & Gellert, 2011a). Alle instruktionerne foregår non-verbalt ved at testlederen viser eleven, hvordan man kan udføre opgaverne og ved at testlederen bruger forskellige gestus til at opfordre og give feedback. Testen består af tre deltest, som beskrives nedenfor.

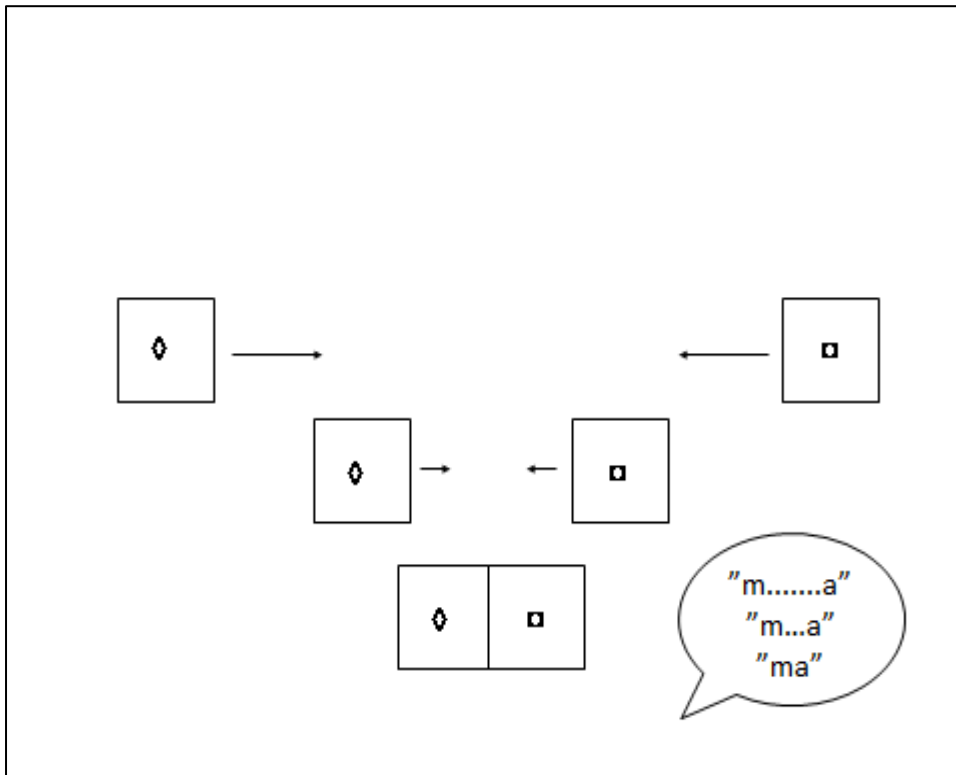
Del 1: Indlæring af et nyt alfabet. I denne deltest skal eleven tilegne sig koblingen mellem tre nye bogstaver (ϣ, Ϙ og ϙ) og deres tilhørende lyde (/s/, /m/ og /a/). Eleven bliver først præsenteret for de tre nye bogstaver og deres tilhørende lyde, idet testlederen peger på ét symbol ad gangen på et øveark og samtidig udtaler den tilhørende lyd. Derefter får eleven på de efterfølgende testark vist de tre nye bogstaver i varierende rækkefølge i op til ti runder og skal benævne bogstaverne ved de tilhørende lyde. Undervejs får eleven hjælp og feedback. Hvis eleven benævner alle de tre bogstaver korrekt i tre på hinanden følgende runder inden den tiende runde, bliver deltesten afbrudt. De resterende runder giver fuldt pointtal, idet det antages, at eleven nu har tilegnet sig koblingen mellem bogstaverne og deres lyde. Scoren var i den aktuelle undersøgelse antal korrekt benævnte bogstaver (max. 30). Uanset scoren på denne deltest går eleven videre til deltest 2. Den interne opgavehomogenitet var i denne undersøgelse 0,81 (vurderet ved Cronbachs Alfa).

Del 2: Læseindlæring med et nyt alfabet. I deltest 2 skal eleven lære at forbinde de nye bogstaver til nonord på blot to bogstaver: *sa, ma, as, am*. Den lydlige syntese bliver illustreret visuelt ved hjælp af bogstavbrikker, der gradvist bliver flyttet sammen (jf. figur 2 nedenfor), og eleven får hjælp til at huske de nye bogstavers lyde, hvis han eller hun bliver i tvivl om dem undervejs. De fire korte nonord bliver præsenteret for eleven i op til fem runder med korrigerende feedback. Scoren var i denne undersøgelse antal korrekt læste nonord (max. 20). Kun hvis eleven læser alle

fire nonord korrekt i to på hinanden følgende runder, går han eller hun videre til deltest 3. Deltestens opgavehomogenitet var i denne undersøgelse 0,96 (Cronbachs Alfa).

Del 3: Videre læsning med nyt alfabet. I den sidste deltest skal eleven bruge det lærte på egen hånd. Eleven præsenteres for forskellige kombinationer af de nye bogstaver svarende til 12 nonord af stigende kompleksitet (fx *masa* og *smams*) og skal prøve at læse disse nonord højt. Testen afbrydes efter tre på hinanden følgende fejl, dvs. tre nonord, der er læst forkert. Scoren var antal korrekt læste nonord (max. 12). Eleven fik tildelt scoren 0 på denne deltest, hvis han eller hun ikke nåede kriteriet på Del 2 og derfor ikke gik videre til Del 3. Opgavehomogeniteten for Del 3 var i denne undersøgelse 0,93 (Cronbachs Alfa).

Elevernes scorer på deltest 2 og 3 i Dynamisk Afkodningstest A blev adderet.



Figur 2. Illustration af deltest 2 i den dynamiske afkodningstest (læseindlæring med et nyt alfabet). Den lydige syntese bliver illustreret visuelt ved hjælp af bogstavbrikker, der gradvist flyttes sammen.

4.3.2.2. Bogstavkendskab

Til afdækning af elevernes bogstavkendskab anvendtes den traditionelle test med benævnelse af store og små bogstaver, som deltagerne også gennemførte i efteråret 0. klasse. For en nærmere beskrivelse af testen henvises der til afsnit 4.3.1.3.

4.3.2.3. Fonologisk opmærksomhed (forlydsidentifikation)

Som den ene af to test af fonologisk opmærksomhed i slutningen af 0. klasse indgik den statiske test af forlydsidentifikation, som deltagerne også gennemførte i efteråret 0. klasse. For en nærmere beskrivelse af testen henvises der til afsnit 4.3.1.1.

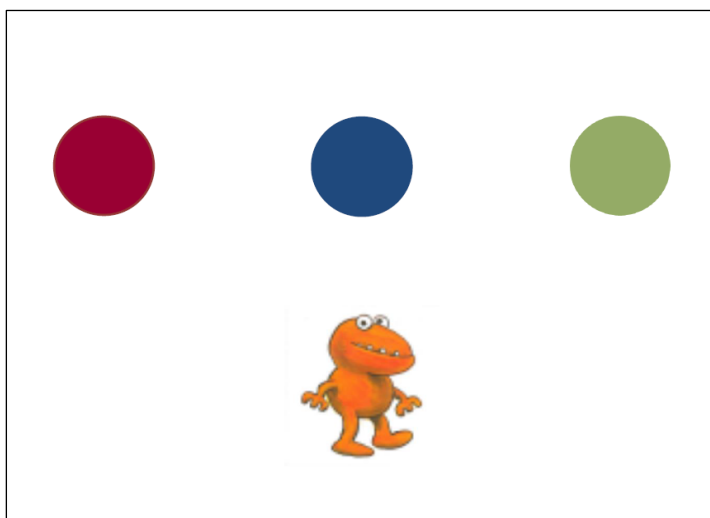
4.3.2.4. *Fonologisk opmærksomhed (fonologisk syntese)*

En test af fonologisk syntese blev udviklet til projektet. Elevernes færdigheder i fonologisk syntese blev undersøgt med to deltest, én med rigtige ord og én med nonord. I begge deltest blev eleven via høretelefoner præsenteret for to eller tre på forhånd indtalte enkeltlyde og bedt om at sætte lydene sammen. På en computerskærm så eleven to eller tre farvede cirkler (afhængigt af antallet af lyde). Når den første lyd blev sagt, blinkede den første cirkel på skærmen. Efter 1½ sekunds pause blev den næste lyd sagt, samtidig med at den næste cirkel blinkede, og så fremdeles. Efter en kort pause hørte eleven igen de to eller tre lyde med 1½ sekund imellem, og igen blinkede cirklerne i takt med lydene. Eleven skulle efter to præsentationer af lydene sige, hvad det blev til, når lydene blev sat sammen.

Fonologisk syntese – rigtige ord. I denne deltest blev eleven efter to øveopgaver præsenteret for først fem opgaver med hver to lyde (fx /s/ - /ø/), der skulle sættes sammen til et ord (*sø*). Derefter fulgte fem opgaver med hver tre lyde (fx /s/ - /o/ - /l/), der skulle sættes sammen. Scoren var antal korrekt besvarede opgaver (max. 10). Deltestens opgavehomogenitet var 0,84 (Cronbachs Alfa).

Fonologisk syntese – nonord. I denne deltest så eleven på computerskærmen under de omtalte cirkler et billede af et fantasidyr (jf. figur 3). Eleven fik at vide, at det var et mærkeligt dyr, som ikke fandtes i virkeligheden, og at dyret også havde et mærkeligt navn. Efter to øveopgaver blev eleven præsenteret for først fem opgaver med hver et nyt fantasidyr og to lyde (fx /y/ - /s/), som eleven skulle sætte sammen for at finde ud af dyrets navn (*Ys*). Derefter fulgte fem opgaver med hver et nyt fantasidyr og tre lyde (fx /s/ - /u/ - /f/), der skulle sættes sammen (*Suf*). Scoren var antal korrekt besvarede opgaver (max. 10). Deltestens opgavehomogenitet var 0,9 (Cronbachs Alfa).

For at styrke pålideligheden af målet for fonologisk syntese blev resultaterne på de to deltest (som indbyrdes korrelerede over 0,7) slået sammen til én score, hvor scorerne på de to deltest blev vægtet ligeligt.



Figur 3. Eksempel på opgave i test af fonologisk syntese (nonord)

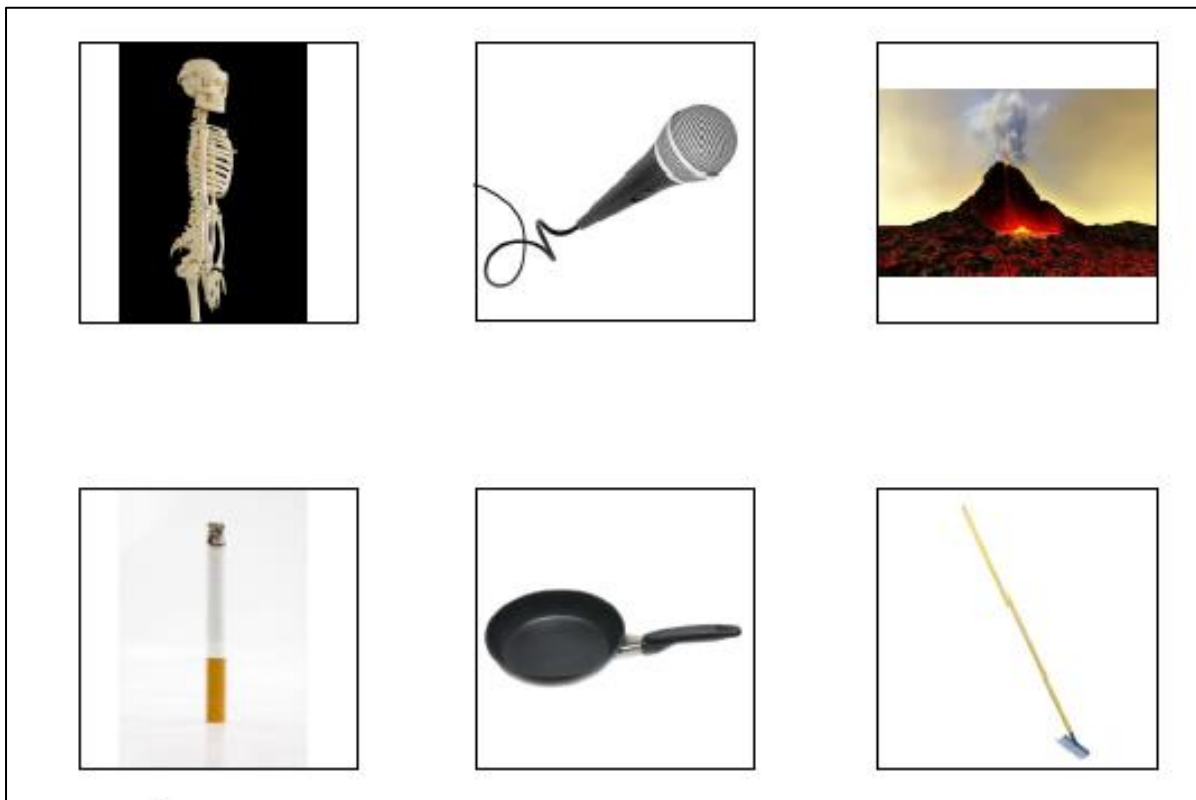
”/s/ - /u/ - /f/... /s/ - /u/ - /f/...

Hvad bliver det til, når man sætter lydene sammen?”

4.3.2.5. *Ordforråd*

Elevernes ordforråd blev undersøgt med en billedbenævnelsestest, som rapportens første forfatter havde udviklet i samarbejde med Rikke Vang Christensen med udgangspunkt i en tidligere anvendt ordforrådstest til større børn (Gellert & Elbro, 2013). Der var seks øveopgaver, hvorefter hver elev blev bedt om at benævne hvert af i alt 84 farvefoto-

grafier med ét ord, fx *jordbær*, *blæksprutte*, *hue*, *kogle*, *vulkan*, *skjold*, *akvarium*, *reol*, *kranium* og *rollator*. Fotografierne blev præsenteret på 14 ark med seks fotografier på hvert ark (jf. figur 4 nedenfor). Scoren var antal korrekt benævnte billeder. Testens interne opgavehomogenitet var 0,96 (Cronbachs Alfa). Eleverne fik også denne ordforrådstest på de øvrige fire testtidspunkter (efteråret 0. klasse, midten og slutningen af 1. klasse samt slutningen af 2. klasse), og test-retest-reliabiliteten over tid var over 0,9.



Figur 4. Eksempel på opgave i test af ordforråd.

”Hvad er det?”

4.3.2.6. RAN: Hurtig automatisk benævnelse af genstande

Denne test af hurtig automatisk benævnelse (på engelsk *rapid automatized naming*, RAN) blev udviklet til projektet. Eleven skulle benævne fire rækker med fem figurer (*sol*, *bil*, *hat*, *fisk* og *is*) så hurtigt som muligt. De samme fem figurer blev gentaget i tilfældig rækkefølge. Før den egentlige test begyndte, blev eleven præsenteret for figurerne og øvede sig på at benævne figurerne i to rækker så hurtigt som muligt. Scoren på testen var gennemførelstiden i sekunder. Eleverne gennemførte også denne test i efteråret 0. klasse, og test-retest-reliabiliteten over tid var 0,69.

4.3.2.7. Afkodningspræcision

Ord-

Ordlæsning: Denne statiske afkodningstest består i højtlesning af 30 rigtige, enkeltstående ord (fx *du*, *glad* og *kærlighed*) (jf. Elbro, 2007; tabel 3.4). Ordene, der alle er hyppige og mellem to og ni bogstaver lange, blev præsenteret for eleven i tre korte lister på hver sit ark. Disse ordlister er blevet udviklet til og anvendt i tidligere dansk forskning (Elbro, 1990; Elbro & Petersen, 2004). Inden de egentlige testord blev eleven præsenteret for et ark med fem lette øveord. Hvis eleven ikke kunne læse nogen af øveordene på egen hånd, gik eleven ikke videre til testordene. Scoren var antal korrekt læste testord. Test-retest-reliabiliteten over tid for denne test var 0,78 i såvel denne undersøgelse som i en tidligere undersøgelse, hvor elever i både 2. og 3. klasse fik den samme test (Elbro & Petersen, 2004).

4.3.3. Testmaterialer i midten af 1. klasse

4.3.3.1. Dynamisk Afkodningstest B

Dynamisk Afkodningstest B blev udviklet til projektet og konstrueret, så den i opbygning var helt parallel med *Dynamisk Afkodningstest A* (jf. afsnit 4.3.2.1), blot med tre andre lyde og symboler. Alle instruktionerne foregik også i *Dynamisk Afkodningstest B* non-verbalt. Testen bestod af tre deltest, som kun kort beskrives nedenfor. For en nærmere redegørelse for deltestenes udformning henvises til afsnit 4.3.2.1, hvor paralleltesten *Dynamisk Afkodningstest A* beskrives detaljeret.

Del 1: Indlæring af et nyt alfabet. I denne deltest skulle eleven tilegne sig koblingen mellem tre nye bogstaver (ϣ, ϕ og ϑ) og deres tilhørende lyde (/f/, /l/ og /v/). Scoren var antal korrekt benævnte bogstaver (max. 30). Uanset scoren på denne deltest gik eleven videre til deltest 2. Overensstemmelsen mellem deltestens enkelte opgaver (den interne opgavehomogenitet) var 0,82 (vurderet ved Cronbachs Alfa).

Del 2: Læseindlæring med et nyt alfabet. I deltest 2 skulle eleven lære at forbinde de nye bogstaver til fire forskellige nonord på to bogstaver. Disse fire korte nonord blev præsenteret for eleven i op til fem runder med korrigerende feedback. Scoren var i denne undersøgelse antal korrekt læste nonord (max. 20). Kun hvis eleven læste alle fire nonord korrekt to på hinanden følgende runder, gik han eller hun videre til deltest 3. Deltestens opgavehomogenitet var 0,93 (Cronbachs Alfa).

Del 3: Videre læsning med nyt alfabet. I den sidste deltest blev eleven præsenteret for forskellige kombinationer af de nye bogstaver svarende til 12 nonord af stigende kompleksitet og skulle prøve at læse disse nonord højt. Scoren var antal korrekt læste nonord (max. 12). Eleven fik tildelt scoren 0 på denne deltest, hvis han eller hun ikke nåede kriteriet på Del 2 og derfor ikke gik videre til Del 3. Opgavehomogeniteten for Del 3 var 0,9 (Cronbachs Alfa).

Elevernes scorer på deltest 2 og 3 i *Dynamisk Afkodningstest B* blev adderet.

4.3.3.2. Afkodningspræcision

Ordlæsning: Eleverne gennemførte igen Elbros test af ordlæsning, som de tidligere havde fået i slutningen af 0. klasse (jf. beskrivelsen ovenfor i afsnit 4.3.2.7).

Nonordslæsning: Denne test består i højtlesning af nonord og giver et traditionelt mål for udnyttelse af skriftens grundlæggende lydprincip. Testen er parallel med den ovenstående test af ordlæsning og indeholder 30 nonord, som matcher de 30 rigtige ord i henseende til længde og konsonant-vokal-struktur (jf. Elbro 2007; tabel 3.4). Scoren var antal korrekt læste nonord. Eleverne gennemførte også denne test i slutningen af hhv. 1. og 2. klasse, og test-retest-reliabiliteten over tid for denne test var minimum 0,76.

For at styrke pålideligheden af målet for afkodningspræcision blev resultaterne fra midten af 1. klasse på Elbros test af hhv. ordlæsning og nonordslæsning (som indbyrdes korrelerede over 0,9) slået sammen til én samlet afkodningspræcisionsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning i midten af 1. klasse blev vægtet ligeligt.

4.3.3.3. Afkodningseffektivitet

Antal korrekt læste ord pr. minut ved højtlesning af Elbros ordlister med hhv. 30 ord og 30 nonord (jf. ovenfor) blev anvendt som mål for *afkodningseffektivitet*, dvs. en kombination af præcisionen og hastigheden i elevens afkodning. Resultaterne på listerne med ordlæsning og nonordslæsning korrelerede over 0,9 og blev slået sammen til én samlet afkodningseffektivitetsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning i midten af 1. klasse blev vægtet ligeligt.

4.3.3.4. Bogstavkendskab

Til afdækning af elevernes bogstavkendskab anvendtes den traditionelle test med benævnelse af store og små bogstaver, som deltagerne også gennemførte i efteråret og slutningen af 0. klasse. For en nærmere beskrivelse af testen henvises der til afsnit 4.3.1.3.

4.3.3.5. Fonologisk syntese

Testen blev udviklet til projektet og lignede den test af fonologisk syntese, som eleverne gennemførte i slutningen af 0. klasse (jf. afsnit 4.3.2.4). I den test, som blev udviklet til og anvendt i midten af 1. klasse til afdækning af elevernes færdigheder i fonologisk syntese, indeholdt alle opgaverne dog nonord (i modsætning til testen i slutningen af 0. klasse, hvor halvdelen af opgaverne indeholdt rigtige ord, mens den anden halvdel indeholdt nonord), og der indgik i midten af 1. klasse nonord med en lidt mere kompleks lydstruktur end i slutningen af 0. klasse.

Lige som i testen af fonologisk syntese i slutningen af 0. klasse blev eleven via høretelefoner præsenteret for to eller tre på forhånd indtalte enkeltlyde og bedt om at sætte lydene sammen. På en computerskærm så eleven to eller tre farvede cirkler (afhængigt af antallet af lyde) samt under disse cirkler et billede af et fantasidyr. Eleven fik at vide, at det var et mærkeligt dyr, som ikke fandtes i virkeligheden, og at dyret også havde et mærkeligt navn. Efter to øveopgaver blev eleven præsenteret for først fem opgaver med hver et nyt fantasidyr og to lyde (fx /f/ - /o/), som eleven skulle sætte sammen for at finde ud af dyrets navn (*Fo*). Derefter fulgte otte opgaver med hver et nyt fantasidyr og tre lyde (fx /s/ - /l/ - /y/), der skulle sættes sammen (*Sly*). Når den første lyd blev sagt, blinkede den første cirkel på skærmen. Efter 1½ sekunds pause blev den næste lyd sagt, samtidig med at den næste cirkel blinkede, og så fremdeles. Efter en kort pause hørte eleven igen de to eller tre lyde med 1½ sekund imellem, og igen blinkede cirklerne i takt med lydene. Eleven skulle efter to præsentationer af lydene sige, hvad det blev til, når lydene blev sat sammen. Scoren var antal korrekt besvarede opgaver (max. 13). Deltestens opgavehomogenitet var 0,87 (Cronbachs Alfa).

4.3.3.6. RAN: Hurtig automatisk benævnelse af cifre

Denne test af hurtig automatisk benævnelse (på engelsk *rapid automatized naming, RAN*) blev udviklet til projektet. Eleven skulle benævne 5 rækker med fem cifre (7, 2, 9, 8 og 5) så hurtigt som muligt. De samme fem cifre blev gentaget i tilfældig rækkefølge. Før den egentlige test med benævnelse af i alt 35 cifre begyndte, blev eleven præsenteret for cifrene og øvede sig på at benævne cifrene så hurtigt som muligt i to rækker. Scoren på testen var gennemførelstiden i sekunder. Eleverne gennemførte også denne test i slutningen af 1. klasse, og test-retest-reliabiliteten over tid var 0,75.

4.3.4 Testmaterialer i slutningen af 1. klasse

4.3.4.1. Afkodningspræcision

Ordlæsning: Eleverne fik igen Elbros test af ordlæsning (jf. beskrivelsen ovenfor i afsnit 4.3.2.7).

Nonordslæsning: Eleverne fik igen Elbros test af nonordslæsning (jf. beskrivelsen i afsnit 4.3.3.2).

Resultaterne fra slutningen af 1. klasse på Elbros test af ordlæsning og nonordslæsning (som indbyrdes korrelerede over 0,9) blev slået sammen til én samlet afkodningspræcisionsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning blev vægtet ligeligt.

4.3.4.2. Afkodningseffektivitet

Antal korrekt læste ord pr. minut ved højtlesning af Elbros ordlister med hhv. 30 ord og 30 nonord blev lige som i midten af 1. klasse anvendt som mål for afkodningseffektivitet (jf. afsnit 4.3.3.3). Resultaterne på listerne med ordlæsning og nonordslæsning blev også i slutningen af 1. klasse slået sammen til én samlet afkodningseffektivitetsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning blev vægtet ligeligt.

4.3.4.3. RAN: Hurtig automatisk benævnelse af cifre

Eleverne gennemførte igen testen af hurtig automatisk benævnelse af cifre, som de tidligere havde fået i midten af 1. klasse (jf. beskrivelsen ovenfor i afsnit 4.3.3.5.).

4.3.5. Testmaterialer i slutningen af 2. klasse

4.3.5.1. Valg af test med henblik på vurdering af afkodningsvanskeligheder i slutningen af 2. klasse

Fra projektets start var det planlagt at anvende Dansk Videnscenter for Ordblindheds testmateriale til identifikation af elever i risiko for dysleksi (Arnbak & Borstrøm, 2007) med henblik på vurdering af, hvorvidt elever i den aktuelle undersøgelse i slutningen af 2. klasse udviste alvorlige afkodningsvanskeligheder. Endvidere inddrog vi på Undervisningsministeriets opfordring som noget ekstra (og altså ud over det oprindeligt lovede testbatteri) den it-baserede ordblindetest, som Undervisningsministeriet udsendte tidligere i 2015 til afdækning af ordblindhed blandt elever fra 3. klasse samt på senere klassetrin, ungdomsuddannelser og videregående uddannelser (jf. Ordblindetestkonsortiet, 2015). Vi valgte desuden også at anvende Elbros ordlister med højtlesning af ord og nonord, da disse ordlister har indgået i tidligere forskningsprojekter (fx Elbro, 1990; Elbro & Petersen, 2004) og er alment udbredte i praksis. Disse forskellige test står nærmere beskrevet nedenfor.

4.3.5.2. Afkodningspræcision 1

Ordlæsning A: Eleverne fik igen Elbros test med højtlesning af 30 rigtige ord (jf. beskrivelse i afsnit 4.3.2.7). Antal korrekt læste ord blev anvendt som mål for afkodningspræcision.

Nonordslæsning A: Eleverne fik igen Elbros test med højtlesning af 30 nonord (jf. beskrivelse i afsnit 4.3.3.2). Antal korrekt læste nonord blev anvendt som mål for afkodningspræcision.

Resultaterne fra slutningen af 2. klasse på Elbros test af ordlæsning og nonordslæsning (som korrelerede over 0,9) blev slået sammen til én samlet afkodningspræcisionsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning i 2. klasse blev vægtet ligeligt.

4.3.5.3. Afkodningspræcision 2

Ordlæsning B: Deltest med højtlesning af rigtige ord fra Dansk Videnscenter for Ordblindheds testmateriale (Arnbak og Borstrøm, 2007), herefter benævnt DVO-ordblindetesten. Deltesten består af lister med 4 x 10 enstavelsesord med henholdsvis 2, 3, 4 og 5 bogstaver.

Nonordslæsning B: Deltest med højtlesning af nonord fra DVO-ordblindetesten. Deltesten består af lister med 4 x 10 nonord med henholdsvis 2, 3, 4 og 5 bogstaver.

Resultaterne fra slutningen af 2. klasse på DVO-ordblindetestens deltest af ordlæsning og nonordslæsning (som korrelerede over 0,9) blev slået sammen til én samlet afkodningspræcisionsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning i 2. klasse blev vægtet ligeligt.

4.3.5.4. Afkodningseffektivitet 1

Antal korrekt læste ord pr. minut ved højtlesning af Elbros ordlister med hhv. 30 ord og 30 nonord blev igen anvendt som mål for afkodningseffektivitet (jf. tidligere beskrivelse i afsnit 4.3.3.3). Resultaterne på listerne med ordlæsning og nonordslæsning blev slået sammen til én samlet afkodningseffektivitetsscore, hvor scorerne på testene af henholdsvis ordlæsning og nonordslæsning i 2. klasse blev vægtet ligeligt.

4.3.5.5. Afkodningseffektivitet 2

Find stavemåden (Ordblindetestkonsortiet, 2015). Denne test indgår i Undervisningsministeriets it-baserede ordblindetest, herefter benævnt UVM-ordblindetesten, og består af 40 opgaver, hvor deltageren hører et nonsensord (fx ”olt”) og skal vælge den rigtige stavemåde ud af fem valgmuligheder (*otl, olp, ølt, olt* og *oft*). Deltageren har fem minutter til at løse så mange opgaver som muligt.

Find det der lyder som et ord (Ordblindetestkonsortiet, 2015). Denne test indgår også i UVM-ordblindetesten og består af 44 opgaver, hvor deltageren ser fem nonord (fx *bænne, jalle, håkke, dimme, fykte*). Deltageren skal vælge det ord, der lyder som et rigtigt ord, når man læser det (*håkke* lyder som ”hugge”, når man læser det). Deltageren har syv minutter til at løse så mange opgaver som muligt.

Den samlede score for de to ovenfor beskrevne deltest fra UVM-ordblindetesten blev anvendt som mål for afkodningseffektivitet.

4.3.5.6. Ordforråd

I UVM-ordblindetesten indgår en basal ordforrådtest, som består af 15 opgaver (Ordblindetestkonsortiet, 2015). I hver delopgave hører deltageren et ord (fx ”avis”) og skal så vælge det billede, som forestiller ordet. Der er fire valgmuligheder. Ordforrådsprøven tjekker ifølge vejledningen (Ordblindetestkonsortiet, 2015), om deltageren har et tilstrækkeligt ordforråd på dansk til at løse *Find det der lyder som et ord*-opgaverne. Resultater med den basale ordforrådsprøve på 10 eller lavere kan ifølge vejledningen muligvis tyde på, at begrænsninger i ordforrådet har haft en lille (negativ) betydning for resultatet af Ordblindetesten.

4.3.5.7. Fonologisk subtraktion

Denne deltest fra Dansk Videnscenter for Ordblindheds testmateriale (Arnbak og Borstrøm, 2007) består af 30 delopgaver, hvor eleven skal tage lyde væk forskellige steder i ord (fx ”Hvis man siger ’kost’ og så tager /k/ væk, hvad bliver der så tilbage?”). Deltesten indeholder 10 opgaver i forlyd, 10 opgaver i indlyd og 10 opgaver i udlyd.

5. Resultater

5.1. Testresultater

I tabel 2 er de udvalgte test opdelt efter testtidspunkt. Tabellen giver en oversigt over de opnåede middelværdier og standardafvigelser for såvel den samlede gruppe af elever som for hver af grupperne af elever med enten dansk som modersmål eller dansk som andetsprog.

Test-tidspunkt	Test	Max. score	Middelværdi (standardafvigelse)		
			Alle deltagere (N = 158)	Dansk som modersmål (N = 103)	Dansk som andetsprog (N = 55)
Efterår 0. klasse	Forlydsidentifikation – statisk	15	9,2 (3,2)	9,3 (3,3)	8,9 (3,1)
	Forlydsidentifikation – dynamisk	60	48,2 (7,8)	48,3 (8,4)	48,1 (6,6)
	Bogstavkendskab (gnm.)	29	14,5 (7,7)	15,4 (7,4)	12,7 (7,9) *
	Sprogvurdering – samlet score	100	40,4 (31,4)	50,3 (30,6)	21,8 (23,6) **
	Sprogvurdering – lydige kompetencer	100	38,5 (31,6)	45,1 (32,5)	26,0 (25,7) **
Slutningen af 0. klasse	Dynamisk afkodning A – del 1	30	26,2 (4,2)	26,3 (4,2)	26,2 (4,3)
	Dynamisk afkodning A – del 2-3	32	15,7 (10,1)	16,2 (10,1)	14,8 (10,0)
	Afkodningspræcision (Elbros ordlister med rigtige ord)	30	4,3 (7,2)	4,9 (7,5)	3,1 (6,3)
	Bogstavkendskab (gnm.)	29	23,2 (5,8)	23,6 (5,4)	22,5 (6,4)
	Forlydsidentifikation – statisk	15	11,7 (3,2)	12,0 (3,1)	11,3 (3,2)
	Fonologisk syntese (gnm.)	10	3,8 (2,9)	4,0 (2,9)	3,4 (2,8)
	Ordforråd	84	49,5 (14,3)	56,4 (8,8)	36,5 (13,6) **
	RAN – genstande		1,1 (0,3)	1,1 (0,3)	1,0 (0,2) *
Midten af 1. klasse	Dynamisk afkodning B – del 1	30	27,0 (4,7)	27,4 (3,7)	26,1 (6,0)
	Dynamisk afkodning B – del 2-3	32	22,0 (9,2)	22,6 (9,1)	20,8 (9,3)
	Afkodningspræcision – Elbros ordlister med rigtige ord og nonord (gnm.)	30	11,2 (8,4)	11,6 (8,5)	10,4 (8,3)
	Afkodningseffektivitet – Elbros ordlister med rigtige ord og nonord (gnm.)		5,1 (6,6)	5,4 (7,0)	4,7 (6,0)
	Bogstavkendskab (gnm.)	29	26,6 (3,2)	26,8 (2,8)	26,3 (3,9)
	Fonologisk syntese	13	7,9 (4,2)	8,1 (4,1)	7,5 (4,3)
	RAN – cifre		1,3 (0,3)	1,3 (0,3)	1,4 (0,3)
Slutningen af 1. klasse	Afkodningspræcision – Elbros ordlister med rigtige ord og nonord (gnm.)	30	18,4 (8,5)	19,0 (8,4)	17,3 (8,7)
	Afkodningseffektivitet – Elbros ordlister med rigtige ord og nonord (gnm.)		15,1 (14,3)	15,9 (15,5)	13,6 (11,8)
	RAN – cifre		1,5 (0,3)	1,5 (0,3)	1,6 (0,4)
Slutningen af 2. klasse	Afkodningspræcision 1 – Elbros ordlister med rigtige ord og nonord (gnm.)	30	23,1 (7,1)	23,9 (6,7)	21,7 (7,8)
	Afkodningspræcision 2 – lister med rigtige ord og nonord fra DVO-ordblindetest (gnm.)	40	32,2 (8,7)	33,1 (7,9)	30,6 (9,3)
	Afkodningseffektivitet 1 – Elbros ordlister med rigtige ord og nonord (gnm.)		26,6 (17,6)	28,0 (18,5)	23,9 (15,5)
	Afkodningseffektivitet 2 – samlet testscore fra UVM-ordblindetest		71,5 (10,5)	72,4 (10,4)	69,7 (10,5)
	Ordforråd – UVM-ordblindetest	15	12,0 (1,7)	12,7 (1,2)	10,7 (1,6) **
	Fonologisk subtraktion – DVO-ordblindetest	30	23,1 (7,2)	24,0 (6,3)	21,5 (8,6)

Tabel 2. Middelværdier og standardafvigelser for den samlede gruppe af deltagere og for hver af grupperne med enten dansk som modersmål eller dansk som andetsprog.

* Deltagerne med dansk som andetsprog scorede gennemsnitligt lavere på denne test end deltagerne med dansk som modersmål. Forskellen er signifikant med $p < 0,05$.

**Deltagerne med dansk som andetsprog gennemsnitligt scorede signifikant lavere på denne test end deltagerne med dansk som modersmål. Forskellen er signifikant med $p < 0,01$.

Der var en loftseffekt på Del 1 i begge de dynamiske afkodningstest (A og B, som blev gennemført i hhv. slutningen af 0. klasse og midten af 1. klasse). Resultaterne på Del 1, hvor eleverne skulle lære forbindelserne mellem de nye symboler og deres lyde, korrelerede kun svagt med de to andre deltest i de dynamiske afkodningstest og heller ikke højt med nogen af de øvrige test i undersøgelsen. Endvidere bidrog Del 1 ikke med signifikant prædiktionsværdi i nogen af de analyser, der rapporteres i det følgende. Af hensyn til overskueligheden i fremstillingen udelades Del 1 derfor i rapporteringen af de senere analyser, således at vi fremover kun inddrager de kombinerede scorer for Del 2 og Del 3 i de dynamiske afkodningstest A og B.

Der sås desuden en vis loftseffekt på følgende mål: Bogstavkendskab i slutningen af 0. klasse og midten af 1. klasse, forlydsidentifikation i slutningen af 0. klasse, Dynamisk Afkodningstest B (Del 2-3) i midten af 1. klasse samt målene for afkodningspræcision og fonologisk subtraktion i slutningen af 2. klasse. Endvidere var der en betydelig gulveffekt på målet for afkodningspræcision i slutningen af 0. klasse. Spredningen i elevernes scorer kan antages at være reduceret pga. de nævnte lofts- og gulveffekter.

T-test viste, at gruppen af deltagere med dansk som andetsprog gennemsnitligt scorede signifikant lavere på nogle af de anvendte test end gruppen af deltagere med dansk som modersmål (jf. markeringerne med stjerner i tabel 2 ovenfor). Dette gjaldt primært for sproglige mål (herunder test af ordforråd). Derimod var der ingen signifikante forskelle mellem gruppernes gennemsnitlige præstationer på hverken de traditionelle afkodningstest eller de dynamiske afkodningstest.

5.2. Indbyrdes sammenhænge mellem testresultater

Tabel 3 nedenfor viser korrelationerne (Spearman's *rho*) mellem elevernes resultater på de udvalgte test i den aktuelle langtidsundersøgelse.

Det bemærkes, at der i slutningen af 2. klasse ses meget stærke korrelationer mellem de forskellige afkodningsmål fra hhv. UVM-ordblindetesten, DVO-ordblindetesten og Elbros ordlister (Spearman's *rho* = 0,76-0,91).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1. Forlydsidentifikation – statisk E 0. kl.																										
2. Forlydsidentifikation – dynamisk E 0. kl.	,93																									
3. Bogstavkendskab efterår 0. kl.	,53	,53																								
4. Sprogvrurdering – samlet score E 0. kl.	,32	,33	,50																							
5. Sprogvrurdering – lydl. kompetencer E 0. kl.	,43	,45	,63	,85																						
6. Dynamisk afkodning A slut 0. kl.	,42	,50	,44	,38	,50																					
7. Afkodningspræcision slut 0. kl.	,48	,52	,57	,48	,60	,76																				
8. Bogstavkendskab slut 0. kl.	,41	,47	,75	,44	,59	,65	,71																			
9. Forlydsidentifikation – statisk slut 0. kl.	,67	,68	,59	,49	,61	,71	,72	,66																		
10. Fonologisk syntese slut 0. kl.	,55	,59	,48	,43	,55	,80	,78	,65	,78																	
11. Ordforråd slut 0. kl.	,26	,30	,33	,62	,46	,29	,34	,27	,36	,34																
12. RAN – genstande slut 0. kl.	,15	,18	,26	,28	,30	,28	,36	,37	,31	,32	,24															
13. Dynamisk afkodning B midt 1. kl.	,44	,50	,36	,41	,46	,76	,64	,54	,65	,72	,35	,35														
14. Afkodningspræcision midt 1. kl.	,40	,45	,43	,40	,50	,80	,76	,67	,67	,80	,28	,41	,84													
15. Afkodningseffektivitet midt 1. kl.	,39	,43	,45	,40	,52	,77	,77	,69	,67	,77	,23	,46	,80	,95												
16. Bogstavkendskab midt 1. kl.	,29	,34	,57	,39	,47	,48	,56	,74	,53	,46	,20	,44	,50	,63	,63											
17. Fonologisk syntese midt 1. kl.	,29	,34	,33	,40	,48	,71	,66	,56	,59	,71	,25	,32	,73	,83	,83	,52										
18. RAN – cifre midt 1. kl.	,17	,21	,28	,20	,27	,33	,42	,41	,31	,35	,02	,69	,38	,49	,59	,55	,42									
19. Afkodningspræcision slut 1. kl.	,31	,37	,36	,45	,51	,70	,66	,59	,61	,70	,30	,42	,75	,84	,83	,63	,76	,47								
20. Afkodningseffektivitet slut 1. kl.	,31	,37	,38	,41	,51	,67	,68	,61	,62	,66	,21	,48	,72	,82	,87	,63	,74	,60	,91							
21. RAN – cifre slut 1. kl.	,14	,20	,19	,16	,25	,33	,37	,38	,34	,35	,03	,60	,35	,47	,55	,49	,40	,76	,55	,67						
22. Afkodningspræcision 1 (Elbro) slut 2. kl.	,29	,37	,38	,47	,53	,65	,60	,61	,61	,62	,35	,44	,69	,74	,73	,59	,67	,48	,81	,80	,55					
23. Afkodningspræcision 2 (DVO) slut 2. kl.	,21	,26	,32	,39	,46	,61	,58	,57	,51	,57	,32	,35	,60	,70	,68	,54	,67	,42	,74	,73	,49	,86				
24. Afkodningseffektivitet 1 (Elbro) slut 2. kl.	,21	,28	,26	,41	,48	,56	,51	,51	,54	,52	,25	,47	,61	,64	,69	,54	,59	,54	,77	,84	,67	,86	,76			
25. Afkodningseffektivitet 2 (UVM) slut 2. kl.	,23	,30	,35	,42	,49	,60	,57	,56	,58	,60	,29	,44	,65	,71	,73	,57	,63	,51	,77	,83	,60	,88	,82	,91		
26. Ordforråd (UVM) slut 2. kl.	,01	,05	,20	,51	,34	,16	,17	,14	,17	,14	,74	,14	,17	,15	,13	,09	,23	,01	,12	,07	,01	,18	,21	,10	,14	
27. Fonologisk subtraktion (DVO) slut 2. kl.	,23	,25	,27	,41	,46	,60	,56	,44	,51	,54	,29	,30	,61	,64	,62	,40	,62	,29	,70	,64	,32	,71	,63	,63	,69	,15

Tabel 3. Korrelationer mellem elevernes resultater på udvalgte test i hhv. efteråret 0. kl., slutningen af 0. kl., midten af 1. kl., slutningen af 1. kl. og slutningen af 2. kl.

Korrelationer over 0,16 er signifikante med $p < 0,05$; korrelationer over 0,21 er signifikante med $p < 0,01$.

5.3. Kategorisering af elever med afkodningsvanskeligheder i slutningen af 2. klasse

5.3.1. Problemstilling

Da sigtet med projektet var at undersøge, hvorvidt man på basis af testresultater i 0. og 1. klasse kan forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, var det nødvendigt at udpege elever med sådanne vanskeligheder ved undersøgelsens sluttidspunkt. Der findes imidlertid ikke noget objektivt eller alment accepteret kriterium for, hvor dårligt elever i slutningen af 2. klasse skal læse for at kunne siges at have alvorlige afkodningsvanskeligheder. Men deltagerne i den aktuelle undersøgelse gennemførte på dette tidspunkt som beskrevet i afsnit 4.3.5 følgende testmaterialer, der i mangel af bedre kunne anvendes som grundlag for vurdering af eventuelle afkodningsvanskeligheder:

- DVO-ordblindetesten
- UVM-ordblindetesten
- Elbros lister med 30 rigtige ord og 30 nonord

Nedenfor redegøres der for hvert af de tre testmaterialer for de kategoriseringer, vi har anvendt i de følgende analyser.

5.3.2. Kategorisering på basis af DVO-ordblindetesten

Dansk Videnscenter for Ordblindheds testmateriale til identifikation af elever i risiko for dysleksi blev afprøvet for omkring 10 år siden blandt elever i begyndelsen og slutningen af 3. klasse (Arnbak & Borstrøm, 2007). De vejledende afgrænsninger for begyndelsen af 3. klasse benyttes i den aktuelle undersøgelse, selv om deltagerne kun gik i slutningen af 2. klasse. For det første er læsefærdighedsniveauet i grundskolens yngre klasser steget i perioden siden afprøvningen af dette materiale (Mejding & Rønberg, 2012, s. 32), og det kan derfor med rimelighed antages, at læsefærdighedsniveauet blandt elever i slutningen af 2. klasse i 2015 ikke er lavere end det tilsvarende niveau i begyndelsen af 3. klasse for omkring 10 år siden. For det andet benyttes DVO-ordblindetesten i praksis ofte allerede fra slutningen af 2. klasse.

Ifølge retningslinjerne i vejledningen til DVO-ordblindetesten placeres en elev i en af ordblindekategorierne ("svære dyslektiske vanskeligheder" eller "dyslektiske vanskeligheder"), hvis eleven scorer under bestemte grænser på mindst 2 ud af 3 test (*højtlesning af rigtige ord*, *højtlesning af nonsensord* og/eller *fonologisk subtraktion*). I de følgende analyser kategoriseres de deltagere, der i slutningen af 2. klasse falder i en af ordblindekategorierne på DVO-ordblindetesten, som elever med alvorlige afkodningsvanskeligheder. Dette svarer til 16,5 % af deltagerne i den aktuelle undersøgelse. Denne relativt høje andel skal ses i lyset af, at der i den aktuelle undersøgelse som tidligere beskrevet (jf. afsnit 4.2.3) var foretaget en oversampling af elever med svage læseforudsætninger.

5.3.3. Kategorisering på basis af UVM-ordblindetesten

Undervisningsministeriets it-baserede ordblindetest retter sig mod elever fra foråret 3. klasse samt senere klassetrin, ungdomsuddannelser og videregående uddannelser (jf. Ordblindetestkonsortiet, 2015).

På basis af resultaterne af afprøvningen af UVM-ordblindetesten blandt elever fra foråret 3. klasse og opefter har ministeriet fastlagt øvre grænseværdier for hhv. ordblindhed og usikker fonologisk kodning svarende til hhv. 8-percentilen og 20-percentilen i uselekerede grupper. Deltagere, der på UVM-ordblindetesten placerer sig svarende til eller under 8-percentilen, kategoriseres således som ordblinde (den såkaldte røde gruppe), mens deltagere, der placerer sig mellem 8- og 20-percentilen, kategoriseres som havende usikker fonologisk kodning (den såkaldte gule gruppe) (jf. Ordblindetestkonsortiet, 2014; Ordblindetestkonsortiet, 2015).

UVM-ordblindetesten er altså ikke tidligere afprøvet helt ned til slutningen af 2. klasse, og den tilhørende vejledning indeholder derfor ikke for dette klassetrin grænser for, hvornår elever kan siges at være ordblinde eller have usikker fonologisk kodning. Da afprøvningen af UVM-ordblindetesten blandt elever i 3. klasse først fandt sted i foråret, må det antages, at elever i slutningen af 2. klasse vil præstere væsentligt lavere på UVM-ordblindetesten, end hvis de først var blevet testet omkring trekvart år senere i foråret 3. klasse. Derfor ville det være problematisk at anvende UVM-ordblindetestens konkrete scoregrænser for 3. klasse med elever, der som i vores undersøgelse kun går i slutningen af 2. klasse.

I stedet valgte vi at beregne estimater af, hvor hhv. 8- og 20-percentilen ligger i et nogenlunde repræsentativt sample i slutningen af 2. klasse og anvende disse estimater som grundlag for kategorisering af elever med afkodningsvanskeligheder på dette tidspunkt. Som nævnt blev der i udvælgelsen af deltagerne i den aktuelle undersøgelse foretaget en oversampling af elever med formodet svage forudsætninger for læsning, og derfor er dette sample ikke repræsentativt. For at opnå et så vidt muligt repræsentativt sample som grundlag for beregningerne af 8- og 20-percentilen fjernede vi fra analyserne de deltagere, der i efteråret 0. klasse bidrog til oversamplingen. I dette reducerede og mere repræsentative sample indgik der kun 117 deltagere (mod 158 i det fulde sample).² Derefter beregnede vi den gennemsnitlige samlede testscore (afkodningseffektivitetsscoren) på UVM-ordblindetesten samt hhv. 8-percentilen og 20-percentilen for dette formodet repræsentative sample. De beregnede talværdier fremgår af den nedenstående tabel 4.

Middelværdi (standardafvigelse)	8-percentil	20-percentil
72,3 (10,6)	57	63

Tabel 4. Middelværdi, standardafvigelse, 8-percentil og 20-percentil for afkodningseffektivitetsscoren fra UVM-ordblindetesten for den formodet repræsentative gruppe af deltagere i slutningen af 2. klasse (N = 117)

² Som tidligere nævnt (jf. afsnit 4.2.4) indikerede elevernes gruppetestresultater med *Ordlæseprøve 2* (Juul & Møller, 2010b) i slutningen af 2. klasse, at det reducerede sample på 117 elever faktisk var nogenlunde repræsentativt med hensyn til ordlæsefærdigheder.

Som det fremgår af tabel 4 ovenfor, svarede 8-percentilen i den formodet repræsentative deltagergruppe i slutningen af 2. klasse ($N = 117$) til en samlet testscore på 57 på UVM-ordblindetesten, mens 20-percentilen svarede til en score på 63. Disse to scoregrænser på hhv. 57 og 63 ligger som forventet et stykke lavere end de tilsvarende scoregrænser svarende til 8- og 20-percentilen i 3. klasse, som er hhv. 62 og 69 (Ordblindetestkonsortiet, 2014; Ordblindetestkonsortiet, 2015). Vi anvendte de to scoregrænser svarende til 8- og 20-percentilen i vores formodet repræsentative sample ($N = 117$) i slutningen af 2. klasse som grundlag for kategorisering af deltagerne i det fulde sample ($N = 158$). Således blev alle de deltagere, der på UVM-ordblindetesten i slutningen af 2. klasse opnåede en samlet testscore på 57 eller derunder kategoriseret som havende alvorlige afkodningsvanskeligheder (svarende til den røde gruppe i vejledningen til UVM-ordblindetesten). Denne gruppe udgjorde 11 % af det fulde sample på 158 deltagere i den aktuelle undersøgelse. Ligeledes blev de deltagere, der på UVM-ordblindetesten opnåede en samlet testscore mellem 58 og 63 kategoriseret som havende moderate afkodningsvanskeligheder (svarende til den gule gruppe i vejledningen til UVM-ordblindetesten). Denne gruppe udgjorde 13 % af det fulde sample på 158 deltagere. Tilsammen udgjorde gruppen med moderate til alvorlige afkodningsvanskeligheder målt med UVM-ordblindetesten altså 24 % af det fulde sample.

Som nævnt er det et ønske fra Undervisningsministeriet, at der i denne rapport fokuseres på forudsigelsen af *alvorlige* afkodningsvanskeligheder (herunder ordblindhed). Det ønske opfylder analyserne nedenfor, når de fokuserer på eleverne under ordblindegrænsen (dvs. under 8-percentilen for klassetrinnet). Imidlertid udgør afkodningsfærdigheder et kontinuum, hvor en del elever har mere moderate afkodningsvanskeligheder, hvilket der også tages højde for i vejledningen til UVM-ordblindetesten. Den deltagergruppe, der som beskrevet ovenfor er blevet udpeget som havende moderate afkodningsvanskeligheder, vil derfor også blive inddraget i de senere analyser mhp. forslag til mulige afgrænsninger af elever i risikogruppe (jf. afsnit 6.2).

5.3.4. Kategorisering på basis af Elbros lister med højtlesning af ord og nonord

Elbros lister med højtlesning af 30 ord og nonord har som nævnt indgået i tidligere forskningsprojekter og er alment udbredte i praksis. Der findes imidlertid ikke nyere normer for slutningen af 2. klasse for disse lister. Med henblik på at kategorisere deltagerne i slutningen af 2. klasse på basis af deres afkodningseffektivitetsscorer på Elbros ordlister anvendte vi derfor en metode, der var helt parallel med den, vi brugte til at kategorisere deltagerne på basis af UVM-ordblindetesten (jf. afsnit 5.3.3 ovenfor).

Vi tog således udgangspunkt i, hvor hhv. 8- og 20-percentilen lå for afkodningseffektivitetsscoren fra Elbros ordlister i det ovenfor beskrevne formodede repræsentativt sample på 117 deltagere i slutningen af 2. klasse (jf. afsnit 5.3.3). Disse percentiler fremgår af den nedenstående tabel 5.

Middelværdi (standardafvigelse)	8-percentil	20-percentil
28,1 (17,4)	6,4	9,8

Tabel 5. Middelværdi, standardafvigelse, 8-percentil og 20-percentil for afkodningseffektivitetsscoren fra Elbros ordlister for den formodet repræsentative gruppe af deltagere i slutningen af 2. klasse ($N = 117$)

Som det fremgår af tabel 5 ovenfor, svarede 8-percentilen i den formodet repræsentative deltagergruppe i slutningen af 2. klasse ($N = 117$) til en afkodningseffektivitetsscore på 6,4 på Elbros ordlister, mens 20-percentilen svarede til en afkodningseffektivitetsscore på 9,8. Disse to scoregrænser (6,4 og 9,8) anvendte vi som grundlag for kategorisering af deltagerne i det fulde sample ($N = 158$). Således blev de deltagere, der på Elbros ordlister opnåede en afkodningseffektivitetsscore på 6,4 eller derunder kategoriseret som havende alvorlige afkodningsvanskeligheder. Denne gruppe udgjorde 11 % af det fulde sample på 158 deltagere. Ligeledes blev de deltagere, der på Elbros ordlister opnåede en afkodningseffektivitetsscore i området 6,5-9,8 kategoriseret som havende moderate afkodningsvanskeligheder. Denne gruppe udgjorde 12 % af det fulde sample på 158 deltagere. Tilsammen udgjorde gruppen med moderate til alvorlige afkodningsvanskeligheder målt med Elbros ordlister altså 23 % af det fulde sample.

5.4. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra efteråret 0. klasse

I dette afsnit redegøres der for forudsigelse af afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra efteråret 0. klasse. Der fokuseres på at besvare følgende spørgsmål: *Kan en dynamisk test af fonologisk opmærksomhed i efteråret 0. klasse bidrage til at forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, også når man allerede kender elevernes præstationer på traditionelle test af læseforudsætninger i efteråret 0. klasse?* De test, der inddrages fra efteråret 0. klasse, står beskrevet i afsnit 4.3.1.

Vi gennemførte en række logistiske regressionsanalyser³ med de 158 deltagere kategoriseret efter afkodningsstatus i slutningen af 2. klasse (alvorlige afkodningsvanskeligheder eller ej) som afhængig variabel. Vi anvendte som beskrevet i afsnit 5.3 elevernes resultater på både DVO-ordblindetesten, UVM-ordblindetesten og Elbros lister med højt læsning af ord og nonord som grundlag for kategoriseringen. De tidligere beskrevne kategoriseringer af deltagere med alvorlige afkodningsvanskeligheder på basis af disse tre testmaterialer blev i de følgende regressionsanalyser på skift anvendt som afhængige variable.

I de tre regressionsanalyser, som rapporteres i dette afsnit, indsatte vi i første trin det kombinerede mål for lydige kompetencer fra Sprog vurderings materialet, da denne score i efteråret 0. klasse var den stærkeste enkelte prædikator af elevernes afkodningsfærdigheder i slutningen af 2. klasse målt som enten afkodningspræcision (vha. DVO-ordblindetesten) eller afkodningseffektivitet (vha. UVM-ordblindetesten og Elbros

³ En regressionsanalyse undersøger sammenhængen mellem en afhængig variabel og andre specificerede uafhængige variable. I en *logistisk* regressionsanalyse er den afhængige variabel dikotom (dvs. at variabelen kun har to mulige værdier). For en nærmere beskrivelse af regressionsanalyser henvises der til indføringer i statistik (fx Elbro & Poulsen, 2015; Howitt & Cramer, 2011).

ordlister). Som det fremgår af korrelationsmatricen i tabel 3, korrelerede elevernes afkodningsfærdigheder således i slutningen af 2. klasse i størrelsesordenen 0,46-0,53 med målet for lydlig kompetencer i efteråret 0. klasse og dermed højere med dette mål end med den samlede score på Sprogvurderingsmaterialet (Spearman's $\rho = 0,39-0,47$) og de separate mål for hhv. bogstavkendskab (0,26-0,37) og fonologisk opmærksomhed (0,21-0,37) i efteråret 0. klasse. Efter at have indsat det kombinerede mål for lydlig kompetencer fra Sprogvurderingsmaterialet i regressionsanalysens første trin inkluderede vi i analysens efterfølgende trin dernæst de øvrige traditionelle test af læseforudsætninger (herunder den statiske test af forlydsidentifikation) og til sidst den dynamiske test af forlydsidentifikation for at undersøge, om nogen af disse test i efteråret 0. klasse kunne bidrage yderligere til forudsigelsen af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse.

På basis af elevernes scorer på kombinationsmålet for lydlig kompetencer fra Sprogvurderingsmaterialet i efteråret 0. klasse kunne man forudsige 12-17 % af variationen i elevernes senere afkodningsstatus (Nagelkerke $R^2 = 0,12-0,17$, jf. trin 1 i tabel 6-8 nedenfor). Efter der var taget højde for elevernes scorer på kombinationsmålet for lydlig kompetencer, kunne den statiske test af forlydsidentifikation bidrage med yderligere 5-9 % til forudsigelsen af elevernes afkodningsstatus, når denne afkodningsstatus i slutningen af 2. klasse var baseret på enten UVM-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister (jf. trin 2 i tabel 7 og 8 nedenfor). Når elevernes afkodningsstatus i slutningen af 2. klasse derimod var baseret på DVO-ordblindetesten, kunne hverken den statiske test af forlydsidentifikation eller nogen af de øvrige prædiktorer fra efteråret 0. klasse bidrage signifikant med yderligere forudsigelse efter kontrol for elevernes lydlig kompetencer. Efter der i analyserne var taget højde for elevernes scorer på de traditionelle test af læseforudsætninger, som kunne bidrage signifikant til forudsigelsen (dvs. kombinationsmålet for lydlig kompetencer og evt. den statiske test af forlydsidentifikation), kunne den dynamiske test af forlydsidentifikation *ikke* bidrage yderligere til forudsigelsen af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse.

Som tidligere beskrevet i afsnit 2.2. angiver AUC-værdierne forholdet mellem sensitivitet og specificitet, og det anbefales internationalt, at den samlede AUC-værdi for testbatterier med henblik på tidlig identifikation af fx elever i risiko for udvikling af læsevanskeligheder skal være på 0,80 eller derover. Som det fremgår af tabel 6-8 nedenfor, ligger AUC-værdierne for den kombinerede score på lydlig kompetencer fra Sprogvurderingsmaterialet i størrelsesordenen 0,72-0,74 og dermed under den anbefalede grænse. Selv i de to analyser, hvor den statiske test af forlydsidentifikation kunne bidrage selvstændigt til forudsigelse af senere alvorlige afkodningsvanskeligheder, var den samlede AUC kun på 0,78-0,79 (jf. trin 2 i tabel 7 og 8) og altså stadig under den anbefalede grænse på 0,80.

Trin	Prædiktor (efterår 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Lydlige kompetencer (Sprogvurdering)	,17	,17	16,8**	11,6	,74

** $p < ,01$.

Tabel 6. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af DVO-ordblindetesten (afkodningspræcision + fonologisk opmærksomhed) og prædiktorer fra efterår 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (efterår 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Lydlige kompetencer (Sprogvurdering)	,13	,13	8,9**	3,4	,73
2	Statisk test af forlydsidentifikation	,18	,05	4,3*	4,0	,78

* $p < ,05$; ** $p < ,01$.

Tabel 7. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af UVM-ordblindetesten (afkodningseffektivitet) og prædiktorer fra efterår 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (efterår 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Lydlige kompetencer (Sprogvurdering)	,12	,12	7,8**	2,5	,72
2	Statisk test af forlydsidentifikation	,21	,09	6,9**	6,0	,79

* $p < ,05$; ** $p < ,01$.

Tabel 8. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af Elbros ordlister (afkodningseffektivitet) og prædiktorer fra efterår 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Vi gennemførte endvidere analyser udelukkende med deltagerne med dansk som andetsprog ($N = 55$) i stedet for med alle deltagerne ($N = 158$) som ovenfor rapporteret. Resultaterne fra de analyser, der udelukkende inddrog deltagere med dansk som andetsprog, må tages med forbehold pga. den beskedne gruppestørrelse. Disse analyser viste imidlertid i lighed med de ovenstående analyser, at kombinationsscoren på lydlige kompetencer fra Sprogvurderingsmaterialet i efteråret 0. klasse var den bedste enkelte prædiktor af afkodningsvanskeligheder i slutningen af 2. klasse.

Sammenfattende viser de ovenfor rapporterede resultater i dette afsnit (5.4), at elevernes kombinationsscore på *lydlige kompetencer* fra Sprogvurderingsmaterialet i efteråret 0. klasse var den bedste enkelte prædiktor af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. En *traditionel, statisk test af fonologisk opmærksomhed (forlydsidentifikation)* kunne dog i hovedparten af analyserne bidrage til en mere præcis

forudsigelse. Når der først var taget højde for elevernes præstationer på traditionelle test af læseforudsætninger (herunder kombinationsmålet for lydige kompetencer og den statistiske test af forlydsidentifikation) kunne en dynamisk test af fonologisk opmærksomhed i efteråret 0. klasse ikke bidrage yderligere til at forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. De rapporterede mål for forholdet mellem sensitivitet og specificitet (AUC-værdierne) ligger under den internationalt anbefalede grænse på 0,80 og indikerer, at forudsigelsen er forbundet med så høj en grad af usikkerhed, at den i praksis har relativt begrænset værdi.

5.5. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra slutningen af 0. klasse

I dette afsnit redegøres der for forudsigelse af senere afkodningsvanskeligheder på basis af testresultater fra slutningen af 0. klasse. Der fokuseres på at besvare følgende spørgsmål: *Kan en dynamisk afkodningstest i slutningen af 0. klasse bidrage til at forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, også når man allerede kender elevernes præstationer på traditionelle test i slutningen af 0. klasse?* De test, der inddrages fra slutningen af 0. klasse, står beskrevet i afsnit 4.3.2.

Som det fremgår af korrelationsmatricen i tabel 3, var Dynamisk Afkodningstest A i slutningen af 0. klasse den stærkeste enkelte prædikator af elevernes afkodningsfærdigheder i slutningen af 2. klasse målt som enten afkodningseffektivitet (vha. UVM-ordblindetesten og Elbros ordlister) eller afkodningspræcision (vha. Elbros ordlister eller DVO-ordblindetesten (Spearman's $\rho = 0,56-0,65$) og altså stærkere end de traditionelt bedste prædiktorer af senere afkodningsfærdigheder, herunder bogstavkendskab, forlydsidentifikation og afkodningspræcision, som alle korrelerede i størrelsesordenen 0,51-0,61 med elevernes afkodningsfærdigheder i slutningen af 2. klasse.

Vi gennemførte logistiske regressionsanalyser med alle eleverne kategoriseret efter afkodningsstatus i slutningen af 2. klasse (alvorlige afkodningsvanskeligheder eller ej) som afhængig variabel. De tidligere beskrevne kategoriseringer ud fra elevernes resultater på hhv. DVO-ordblindetesten (afkodningspræcision og fonologisk opmærksomhed), UVM-ordblindetesten (afkodningseffektivitet) og Elbros ordlister (afkodningseffektivitet) indgik i de følgende regressionsanalyser på skift som afhængige variable.

Blandt de uafhængige variable i regressionsanalyserne, som rapporteres nedenfor, indgik i de første trin de traditionelle test af læseforudsætninger og læsefærdigheder, som i slutningen af 0. klasse hver for sig bidrog selvstændigt til forudsigelse af, hvorvidt eleverne udviste alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse. Derefter inkluderede vi i et efterfølgende trin Dynamisk Afkodningstest A i analyserne for at undersøge, om denne test i slutningen af 0. klasse efter kontrol for de traditionelle prædiktorer yderligere kunne bidrage til forudsigelsen af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder i slutningen af 2. klasse.

I alle de tre regressionsanalyser, som rapporteres i dette afsnit, indsatte vi i første trin elevernes resultater på testen af forlydsidentifikation, som i slutningen af 0. klasse var den stærkeste eller den næststærkeste *traditionelle* prædiktor af deltagernes afkodningsfærdigheder i slutningen af 2. klasse (jf. tabel 3). Samtidig er testen af forlydsidentifikation en gængs testtype i Danmark i modsætning til den anden test af fonologisk opmærksomhed i slutningen af 0. klasse, fonologisk syntese, som også var en stærk prædiktor af senere afkodningsfærdigheder, men som blev konstrueret specielt til projektet og tog længere tid at gennemføre end testen af forlydsidentifikation. På basis af elevernes resultater alene på den traditionelle test af forlydsidentifikation i slutningen af 0. klasse kunne man forudsige 22-39 % af variationen i deres afkodningsstatus i slutningen af 2. klasse (Nagelkerke $R^2 = 0,22-0,39$, jf. trin 1 i tabellerne 9-11 nedenfor). Når der først var taget højde for elevernes resultater på testen af forlydsidentifikation, var bogstavkendskab i alle analyserne derefter den stærkeste traditionelle prædiktor af deltagernes afkodningsstatus i 2. klasse og bidrog med yderligere 6-8 % til forudsigelsen (ændring i Nagelkerke $R^2 = 0,06-0,08$, jf. trin 2 i tabellerne 9-11). Efter kontrol for både forlydsidentifikation og bogstavkendskab i slutningen af 0. klasse bidrog ingen af de øvrige traditionelle prædiktorer (hverken afkodningspræcision, fonologisk syntese, RAN eller ordforråd) yderligere til forudsigelsen af elevernes afkodningsstatus i slutningen af 2. klasse, når den afhængige variabel var enten DVO-ordblindetesten eller UVM-ordblindetesten. Med afkodningseffektivitetsscoren på Elbros ordlister som afhængig variabel bidrog afkodningspræcision imidlertid yderligere med 5 % til forudsigelsen efter kontrol for forlydsidentifikation og bogstavkendskab (jf. tabel 11).

Dynamisk Afkodningstest A kunne bidrage med yderligere 7 % til forudsigelsen af, hvorvidt eleverne i 2. klasse udviste alvorlige afkodningsvanskeligheder eller ej selv efter skrap kontrol med alle de traditionelle prædiktorer i slutningen af 0. klasse, når den afhængige variabel var enten DVO-ordblindetesten eller UVM-ordblindetesten (jf. trin 3 i tabel 9-10). Derimod var det ekstra bidrag fra den dynamiske afkodningstest ikke signifikant med afkodningseffektivitetsscoren på Elbros ordlister som afhængig variabel (jf. tabel 11).

Trin	Prædiktor (slut 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Forlydsidentifikation	,22	,22	21,6**	0,3	,79
2	Bogstavkendskab	,28	,06	6,8**	4,3	,82
3	Dyn. afkodningstest A	,35	,07	8,2**	7,2	,84

** $p < ,01$.

Tabel 9. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af DVO-ordblindetesten (afkodningspræcision + fonologisk opmærksomhed) og prædiktorer fra slutningen af 0. klasse. Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 0. kl.)	Nagel-kerke R^2	Ændring i Nagel-kerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Forlydsidentifikation	,39	,39	33,6**	1,2	,86
2	Bogstavkendskab	,45	,06	6,1*	9,4	,91
3	Dyn. afkodningstest A	,52	,07	7,3**	5,5	,93

* $p < ,05$; ** $p < ,01$.

Tabel 10. Forudsigtelse af afkodningsstatus i 2. klasse med kategorisering på basis af UVM-ordblindetesten (afkodningseffektivitet) og prædiktorer fra slutningen af 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 0. kl.)	Nagel-kerke R^2	Ændring i Nagel-kerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Forlydsidentifikation	,30	,30	26,5**	1,8	,85
2	Bogstavkendskab	,38	,08	7,7**	4,1	,88
3	Afkodningspræcision	,43	,05	4,8*	0,0	,89
4	Dyn. afkodningstest A	,45	,02	2,4	2,1	,90

* $p < ,05$; ** $p < ,01$.

Tabel 11. Forudsigtelse af afkodningsstatus i 2. klasse med kategorisering på basis af Elbros ordlister (afkodningseffektivitet) og prædiktorer fra slutningen af 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Som nævnt var Dynamisk Afkodningstest A i slutningen af 0. klasse den stærkeste enkelte prædiktor af elevernes senere afkodningsfærdigheder i slutningen af 2. klasse (jf. korrelationsmatricen i tabel 3). Endvidere viste hovedparten af analyserne i det foregående afsnit, at denne dynamiske afkodningstest kunne bidrage selvstændigt til forudsigtelsen af, hvorvidt eleverne i 2. klasse ville udvise afkodningsvanskeligheder eller ej selv efter skrap kontrol med alle de traditionelle prædiktorer i slutningen af 0. klasse (jf. tabel 9-11 ovenfor). På denne baggrund er det relevant at stille det følgende spørgsmål: *Kan den dynamiske afkodningstest helt erstatte de traditionelle test i slutningen af 0. klasse – eller styrker det forudsigtelsen af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, hvis man foruden den dynamiske afkodningstest inkluderer traditionelle test i slutningen af 0. klasse?*

For at besvare dette spørgsmål gennemførte vi logistiske regressionsanalyser svarende til de ovenstående, men nu med Dynamisk Afkodningstest A i første trin og derefter de traditionelle prædiktorer fra 0. klasse (bogstavkendskab, forlydsidentifikation, fonologisk syntese, ordforråd, RAN og afkodningspræcision) i de følgende trin. Lige som i de foregående analyser blev eleverne kategoriseret efter afkodningsstatus i slutnin-

gen af 2. klasse (alvorlige afkodningsvanskeligheder eller ej). De tidligere beskrevne kategoriseringer ud fra elevernes resultater på hhv. DVO-ordblindetesten (afkodningspræcision og fonologisk opmærksomhed), UVM-ordblindetesten (afkodningseffektivitet) og Elbros ordlister (afkodningseffektivitet) indgik også i de følgende regressionsanalyser på skift som afhængige variable.

Når Dynamisk Afkodningstest A blev indsat på det første trin i analyserne, kunne man på grundlag af denne test alene forudsige 29-36 % af variationen i deltagernes afkodningsstatus i slutningen af 2. klasse (jf. trin 1 i tabel 12-14 nedenfor). Efter kontrol for den dynamiske afkodningstest var bogstavkendskab den stærkeste traditionelle prædikator af deltagernes afkodningsstatus i 2. klasse og bidrog med yderligere 6-15 % til forudsigelsen (jf. trin 2 i tabel 12-14). Efter kontrol for både den dynamiske afkodningstest og bogstavkendskab bidrog ingen af de øvrige traditionelle prædiktorer yderligere til forudsigelsen af elevernes afkodningsstatus i 2. klasse, når denne afkodningsstatus var baseret på enten DVO-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister. Når UVM-ordblindetesten var den afhængige variabel, bidrog RAN (hurtig automatisk benævnelse af genstande) imidlertid med yderligere 4 % til forudsigelsen (jf. tabel 13, trin 3). Ingen af de øvrige traditionelle prædiktorer i slutningen af 0. klasse (hverken afkodningspræcision, fonologisk syntese eller ordforråd) kunne bidrage yderligere til forudsigelsen af deltagernes afkodningsstatus, når der først var taget højde for deres resultater på den dynamiske afkodningstest, bogstavkendskab og RAN.

Som tidligere beskrevet i afsnit 2.2 angiver AUC-værdierne forholdet mellem sensitivitet og specificitet, og det anbefales internationalt, at den samlede AUC-værdi for testbatterier mhp. tidlig identifikation af fx elever i risiko for udvikling af læsevanskeligheder skal være på mindst 0,80. Som det fremgår af tabel 12-14 nedenfor, ligger AUC-værdierne i alle analyserne for den dynamiske afkodningstest i kombination med bogstavkendskab på 0,84-0,92 og altså dermed et stykke over den anbefalede minimumsgrænse på 0,80, uanset hvilket mål der bruges for elevernes afkodningsstatus i slutningen af 2. klasse.

Trin	Prædikator (slut 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Dyn. afkodningstest A	,29	,29	30,1**	10,8	,82
2	Bogstavkendskab	,35	,06	6,2*	6,0	,84

* $p < ,05$; ** $p < ,01$.

Tabel 12. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af DVO-ordblindetesten (afkodningspræcision + fonologisk opmærksomhed) og prædiktorer fra slutningen af 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Dyn. afkodningstest A	,36	,36	30,9**	8,3	,87
2	Bogstavkendskab	,51	,15	14,8**	11,4	,92
3	RAN (genstande)	,55	,04	4,1**	4,1	,94

** $p < ,01$.

Tabel 13. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af UVM-ordblindetesten (afkodningseffektivitet) og prædiktorer fra slutningen af 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 0. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Dyn. afkodningstest A	,29	,29	25,5**	7,3	,84
2	Bogstavkendskab	,40	,11	10,0**	9,2	,88

** $p < ,01$.

Tabel 14. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af Elbros ordlister (afkodningseffektivitet) og prædiktorer fra slutningen af 0. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Vi gennemførte endvidere analyser udelukkende med deltagerne med dansk som andetsprog ($N = 55$) i stedet for med alle deltagerne ($N = 158$) som ovenfor rapporteret. Resultaterne fra de analyser, der udelukkende inddrog deltagere med dansk som andetsprog, må tages med forbehold pga. den beskedne gruppestørrelse. Disse analyser viste imidlertid i lighed med de ovenstående analyser, at den dynamiske afkodningstest i slutningen af 0. klasse var den bedste enkelte prædiktor af afkodningsvanskeligheder i slutningen af 2. klasse.

Sammenfattende viser analyserne rapporteret i dette afsnit (5.5), at den *dynamiske afkodningstest* i slutningen af 0. klasse var den stærkeste enkelte prædiktor af afkodningsfærdigheder i slutningen af 2. klasse. Selv efter skrap kontrol med alle traditionelle prædiktorer bidrog denne dynamiske test selvstændigt til at forudsige afkodningsvanskeligheder i slutningen af 2. klasse i hovedparten af analyserne. Endvidere viser analyserne, at selv om den dynamiske afkodningstest i slutningen af 0. klasse var den stærkeste enkelte prædiktor af senere afkodningsfærdigheder, kunne denne dynamiske test *ikke* helt erstatte de traditionelle test. Især testen af *bogstavkendskab* bidrog yderligere til forudsigelse af senere afkodningsvanskeligheder. De rapporterede mål for forholdet mellem sensitivitet og specificitet (AUC-værdierne) indikerer, at der på basis af den dynamiske afkodningstest i kombination med en traditionel test af bogstavkendskab i slutningen af 0. klasse kan opnås en præcision i forudsigelsen, der ligger et godt stykke over den internationalt anbefalede grænse for testbatterier mhp. tidlig identifikation.

5.6. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra midten af 1. klasse

I dette afsnit redegøres der for forudsigelse af senere afkodningsvanskeligheder på basis af testresultater fra midten af 1. klasse. Der fokuseres på at besvare følgende spørgsmål: *Kan en dynamisk afkodningstest i midten af 1. klasse bidrage til at forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, også når man allerede kender elevernes præstationer på traditionelle test i midten af 1. klasse?* De test, der inddrages fra midten af 1. klasse, står beskrevet i afsnit 4.3.3.

Som det fremgår af korrelationsmatricen i tabel 3, var der stærke sammenhænge mellem på den ene side afkodningspræcision og afkodningseffektivitet i midten af 1. klasse og på den anden side afkodningsfærdigheder i 2. klasse målt som enten afkodningspræcision (vha. DVO-ordblindetesten) eller afkodningseffektivitet (vha. UVM-ordblindetesten og Elbros ordlister) (Spearman's $\rho = 0,64-0,74$). Som det også fremgår af tabel 3, var sammenhængene mellem elevernes resultater på Dynamisk Afkodningstest B i midten af 1. klasse og deres afkodningsfærdigheder i slutningen af 2. klasse lidt lavere, men stadig relativt stærke (Spearman's $\rho = 0,60-0,69$).

For at undersøge, om Dynamisk Afkodningstest B i midten af 1. klasse kunne bidrage til at forudsige senere afkodningsvanskeligheder, også når man allerede kender elevernes præstationer på traditionelle test, gennemførte vi logistiske regressionsanalyser med alle eleverne kategoriseret efter afkodningsstatus i slutningen af 2. klasse (alvorlige afkodningsvanskeligheder eller ej) som afhængig variabel. De tidligere beskrevne kategoriseringer ud fra elevernes resultater på hhv. DVO-ordblindetesten, UVM-ordblindetesten og Elbros ordlister indgik i de følgende regressionsanalyser på skift som afhængige variable.

Blandt de uafhængige variable i regressionsanalyserne, som rapporteres nedenfor, indgik i de første trin de traditionelle test, som i midten af 1. klasse hver for sig bidrog selvstændigt til forudsigelse af, hvorvidt eleverne udviste afkodningsvanskeligheder eller ej i slutningen af 2. klasse. Derefter inkluderede vi i et efterfølgende trin Dynamisk Afkodningstest B i analyserne for at undersøge, om denne test i midten af 1. klasse yderligere kunne bidrage til forudsigelsen af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder i slutningen af 2. klasse.

I alle de tre regressionsanalyser, som rapporteres i dette afsnit, indsatte vi i første trin elevernes afkodningspræcision målt med højtlesning af Elbros ordlister med rigtige ord og nonord, da det i praksis er lettere at måle afkodningspræcision end afkodningseffektivitet, og da elevernes afkodningspræcision i midten af 1. klasse var den stærkeste *traditionelle* prädiktor af deltagernes afkodningsfærdigheder i slutningen af 2. klasse målt som afkodningspræcision på basis af ordlister fra DVO-ordblindetesten eller Elbros ordlister (jf. tabel 3). Alene på basis af elevernes afkodningspræcision i midten af 1. klasse kunne man forudsige 35-61 % af variationen i deres afkodningsstatus i slutningen af 2. klasse (Nagelkerke $R^2 = 0,35-0,61$, jf. trin 1 i tabel 15-17 nedenfor). Når der først var taget højde for elevernes afkodningspræcision i midten af 1. klasse, bidrog hverken afkodningseffektivitet, fonologisk syntese eller RAN yderligere til forudsigelsen af elevernes afkod-

ningsstatus i slutningen af 2. klasse i nogen af analyserne. Imidlertid bidrog bogstavkendskab med yderligere 5-9 % til forudsigelsen af deltagernes afkodningsstatus (jf. trin 2 i tabel 15-17).

Den dynamiske afkodningstest kunne efter kontrol for de traditionelle mål i midten af 1. klasse *ikke* bidrage selvstændigt til forudsigelsen af elevernes senere afkodningsstatus, uanset om denne afkodningsstatus var baseret på DVO-ordblindetesten, UVM-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister.

Trin	Prædiktor (midt 1. kl.)	Nagel-kerke R^2	Ændring i Nagel-kerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,35	,35	36,5**	9,2	,85
2	Bogstavkendskab	,40	,05	5,9*	5,2	,86

* $p < ,05$; ** $p < ,01$.

Tabel 15. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af DVO-ordblindetesten (afkodningspræcision og fonologisk opmærksomhed) og prædiktorer fra midten af 1. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (midt 1. kl.)	Nagel-kerke R^2	Ændring i Nagel-kerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,61	,61	56,9**	4,7	,95
2	Bogstavkendskab	,68	,07	7,5**	6,0	,96

** $p < ,01$.

Tabel 16. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af UVM-ordblindetesten (afkodningseffektivitet) og prædiktorer fra midten af 1. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (midt 1. kl.)	Nagel-kerke R^2	Ændring i Nagel-kerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,54	,54	50,4**	6,0	,94
2	Bogstavkendskab	,63	,09	10,2**	7,8	,95

** $p < ,01$.

Tabel 17. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af Elbros ordlister (afkodningseffektivitet) og prædiktorer fra midten af 1. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Sammenfattende viser analyserne rapporteret i dette afsnit (5.6), at man ud fra elevernes *afkodningspræcision* i midten af 1. klasse kunne forudsige en betydelig del af variationen i elevernes afkodningsstatus i slutningen af 2. klasse. Blot på basis af målet for afkodningspræcision i midten af 1. klasse lå AUC-værdierne i området 0,85-0,94 og dermed et godt stykke over den anbefalede minimumsgrænse på 0,80. Når

der først var taget højde for elevernes afkodningspræcision i midten af 1. klasse, blev præcisionen i forudsigelsen af deres afkodningsstatus i 2. klasse *ikke* signifikant bedre, hvis man inddrog mål for deres afkodningseffektivitet i midten af 1. klasse. Derimod kunne der påvises et signifikant ekstra bidrag til forudsigelsen af afkodningsvanskeligheder i slutningen af 2. klasse på basis af et traditionelt mål for bogstavkendskab i midten af 1. klasse. I midten af 1. klasse, hvor elevernes afkodningspræcision i sig selv var en meget stærk prædikator af deres senere afkodningsstatus, kunne en dynamisk afkodningstest *ikke* bidrage til en mere præcis forudsigelse.

5.7. Forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testresultater fra slutningen af 1. klasse

Som det fremgår af korrelationsmatricen i tabel 3, var der meget stærke sammenhænge mellem på den ene side elevernes afkodningspræcision og afkodningseffektivitet i slutningen af 1. klasse og på den anden side deres senere afkodningsfærdigheder i 2. klasse målt som enten afkodningspræcision (vha. DVO-ordblindetesten) eller afkodningseffektivitet (vha. UVM-ordblindetesten og Elbros ordlister) (Spearman's $\rho = 0,73-0,84$). Som tidligere nævnt er det i praksis lettere at måle afkodningspræcision end afkodningseffektivitet, og det var derfor værd at undersøge, om man kan nøjes med at måle elevens afkodningspræcision i slutningen af 1. klasse mhp. udpegning af elever, der risikerer at udvikle alvorlige afkodningsvanskeligheder, eller om afkodningseffektivitet og evt. andre prædiktorer derudover kan bidrage væsentligt til forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse.

For at undersøge dette gennemførte vi logistiske regressionsanalyser med alle eleverne kategoriseret efter afkodningsstatus i slutningen af 2. klasse (alvorlige afkodningsvanskeligheder eller ej) som afhængig variabel. I første trin indsatte vi elevernes afkodningspræcision målt med højtlesning af Elbros ordlister med ord og nonord i slutningen af 1. klasse. Alene på basis af elevernes afkodningspræcision på dette tidspunkt kunne man forudsige 52-64 % af variationen i deres afkodningsstatus i slutningen af 2. klasse (Nagelkerke $R^2 = 0,52-0,64$, jf. trin 1 i tabel 18-20 nedenfor). Efter at der var taget højde for elevernes afkodningspræcision i slutningen af 1. klasse, bidrog afkodningseffektivitet med yderligere 3-5 % til forudsigelsen af elevernes afkodningsstatus i slutningen af 2. klasse, når denne afkodningsstatus var baseret på elevernes resultater på UVM-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister (jf. trin 2 i tabel 19 og 20). Afkodningseffektivitet i slutningen af 1. klasse bidrog imidlertid ikke signifikant til forudsigelsen af afkodningsvanskeligheder i slutningen af 2. klasse i analysen med DVO-ordblindetesten som den afhængige variabel, efter der var taget højde for elevernes afkodningspræcision i slutningen af 1. klasse.⁴ Ingen af de øvrige

⁴ I DVO-ordblindetesten er vurderingen af elevernes afkodningsfærdigheder udelukkende baseret på præcision, hvorimod også hastigheden spiller en rolle for vurderingen i både UVM-ordblindetesten og på det anvendte afkodningseffektivitetsmål for Elbros ordlister. Det er derfor forståeligt, at elevernes afkodningseffektivitet i slutningen af 1. klasse ikke kunne bidrage ekstra til forudsigelsen af deres afkodningsstatus i slutningen af 2. klasse, når denne afkodningsstatus var baseret på DVO-ordblindetesten.

prædiktorer i slutningen af 1. klasse bidrog yderligere til forudsigelsen i nogen af analyserne, når der først var taget højde for elevernes afkodningsfærdigheder på dette tidspunkt.

Trin	Prædiktor (slut 1. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,52	,52	58,0**	31,7	,91

** $p < ,01$.

Tabel 18. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af DVO- ordblindetesten (afkodningspræcision og fonologisk opmærksomhed) og prædiktorer fra slutningen af 1. klasse. Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 1. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,64	,64	60,1**	3,6	,96
2	Afkodningseffektivitet	,67	,03	4,0*	2,8	,97

* $p < ,05$; ** $p < ,01$.

Tabel 19. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af UVM-ordblindetesten (afkodningseffektivitet) og prædiktorer fra slutningen af 1. klasse. Beregningerne er baseret på det fulde sample ($N = 158$).

Trin	Prædiktor (slut 1. kl.)	Nagelkerke R^2	Ændring i Nagelkerke R^2	Ændring i χ^2	Wald (i sidste trin)	AUC
1	Afkodningspræcision	,63	,63	60,8**	2,7	,96
2	Afkodningseffektivitet	,68	,05	5,4*	3,8	,97

* $p < ,05$; ** $p < ,01$.

Tabel 20. Forudsigelse af afkodningsstatus i 2. klasse med kategorisering på basis af Elbros ordlister (afkodningseffektivitet) og prædiktorer fra slutningen af 1. klasse. Beregningerne er baseret på det fulde sample ($N = 158$).

Endvidere gennemførte vi tilsvarende regressionsanalyser, hvor vi i det første trin indsatte afkodningseffektivitet som prædiktor i stedet for som ovenfor afkodningspræcision. I to ud af tre analyser blev forudsigelsen af elevernes senere afkodningsstatus *bedre*, når man indsatte afkodningseffektivitet i stedet for afkod-

ningspræcision i det første trin. Dette var tilfældet, når den afhængige variabel var enten UVM-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister i slutningen af 2.klasse. Kun når DVO-ordblindetesten var den afhængige variabel, var afkodningseffektivitet i slutningen af 1. klasse i sig selv en ringere prædiktor end afkodningspræcision. Dette stemmer overens med resultaterne vist i tabel 19 og 20 ovenfor, hvor det netop ses, at afkodningseffektivitet efter kontrol for afkodningspræcision bidrager selvstændigt til forudsigelse af senere afkodningsvanskeligheder vurderet på basis af UVM-ordblindetesten eller afkodningseffektivitetsscoren på Elbros ordlister i slutningen af 2. klasse.

Sammenfattende viser hovedparten af analyserne rapporteret i dette afsnit (5.7), at selv efter der var taget højde for det stærke forudsigelsesbidrag fra elevernes afkodningspræcision i slutningen af 1. klasse, bidrog deres *afkodningseffektivitet* i slutningen af 1. klasse yderligere og gjorde forudsigelsen af elevernes afkodningsstatus i slutningen af 2. klasse endnu mere præcis.

5.8. Sammenfatning og diskussion vedr. forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse på basis af testning i 0. og 1. klasse

Det nedenstående skema (tabel 21) giver en oversigt over de test, der som rapporteret i de foregående afsnit (5.4-5.7) på de fire forskellige testtidspunkter i løbet af 0. og 1. klasse kan udpeges som de vigtigste prædiktorer af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. For hvert af prædiktionstidspunkterne angives de tilhørende AUC-værdier for de prædiktorer, der i skemaet angives som de vigtigste.

Prædiktionstidspunkt	Vigtigste prædiktorer	AUC
Efterår 0. kl.	Lydlige kompetencer (Sprogvurdering) + Statisk test af forlydsidentifikation	,72-,74 ,74-,79
Slutningen af 0. kl.	Dynamisk Afkodningstest A + Bogstavkendskab	,82-,87 ,84-,92
Midten af 1. kl.	Afkodningspræcision + Bogstavkendskab	,85-,95 ,87-,96
Slutningen af 1. kl.	Afkodningseffektivitet	,91-,97

Tabel 21. Oversigt over vigtigste prædiktorer af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse fra hhv. efteråret og slutningen af 0. klasse samt midten og slutningen af 1. klasse med tilhørende AUC-værdier.

Som det fremgår af den opsummerende tabel 21 ovenfor samt tabel 6-20 i afsnit 5.4-5.7, er der for hvert af prædiktionstidspunkterne lidt variation i den samlede præcision i forudsigelsen (jf. spændvidden i AUC-værdierne, der som tidligere nævnt udgør et samlet mål for forholdet mellem sensitivitet og specificitet).

Ligeledes viste analyserne i de foregående afsnit, at forudsigelsesbidragene fra de uafhængige variable (prædiktorerne) varierede lidt i forhold til, hvilken afhængig variabel der anvendtes som grundlag for kategoriseringen af eleverne i slutningen af 2. klasse (DVO-ordblindetesten, UVM-ordblindetesten eller Elbros ordlister). Der sås en tendens til, at elevernes afkodningsstatus i slutningen af 2. klasse i lidt mindre omfang kunne forudsiges på basis af testene på de tidligere testtidspunkter, når den afhængige variabel var baseret på

DVO-ordblindetesten end på UVM-ordblindetesten og Elbros ordlister (jf. afsnit 5.4-5.7). Denne tendens hænger muligvis sammen med, at DVO-ordblindetesten tester lidt andre færdigheder end de to andre anvendte testmaterialer. I DVO-ordblindetesten er kategoriseringen af eleverne således baseret på deres afkodningspræcision og ikke på deres afkodningseffektivitet som i UVM-ordblindetesten og Elbros ordlister. Derudover er kategoriseringen af eleverne i DVO-ordblindetesten i modsætning til i de to øvrige testmaterialer ikke blot baseret på deres præstationer på afkodningstest, men også på deres præstationer på en test af fonologisk opmærksomhed (fonologisk subtraktion). Desuden er der den forskel, at mens kun 11 % af deltagerne i den aktuelle undersøgelse blev kategoriseret som havende alvorlige afkodningsvanskeligheder på basis af UVM-ordblindetesten og Elbros ordlister, blev 16,5 % og altså en større andel af deltagerne kategoriseret som havende alvorlige afkodningsvanskeligheder på basis af DVO-ordblindetesten (jf. afsnit 5.3). Det må dog understreges, at der i den aktuelle undersøgelse generelt var meget stærke sammenhænge mellem elevernes resultater på afkodningsmålene fra hhv. DVO-ordblindetesten, UVM-ordblindetesten og Elbros ordlister i slutningen af 2. klasse (jf. tabel 3), og at der for de tre afhængige variable (DVO-ordblindetesten, UVM-ordblindetesten og Elbros ordlister) var relativt små forskelle mellem forudsigelsesbidragene og AUC-værdierne for prædiktorerne på de forskellige prædiktions tidspunkter.

Som tabel 21 viser – og som det mere grundigt er rapporteret i forbindelse med analyserne i afsnit 5.4-5.7 – gav blot en enkelt prædiktor i slutningen af 1. klasse den bedst mulige forudsigelse af alvorlige afkodningsvanskeligheder (afkodningseffektivitet). I modsætning hertil kunne der i både efteråret 0. klasse, slutningen af 0. klasse og midten af 1. klasse opnås en mere præcis forudsigelse, hvis flere prædiktorer blev kombineret. Derfor står der i tabel 21 ud for prædiktions tidspunktet efterår 0. klasse to prædiktorer. Den stærkeste prædiktor på dette tidspunkt, lydlig kompetencer fra Sprog vurderings materialet, og de tilhørende AUC-værdier står øverst. Derefter står den statiske test af forlydsidentifikation, som i efteråret 0. klasse kunne bidrage yderligere til forudsigelsen efter kontrol for elevernes scorer på kombinationsmålet for lydlig kompetencer, samt de AUC-værdier, der opnåedes på basis af disse to prædiktorer tilsammen. Tilsvarende er der i tabel 21 ud for prædiktions tidspunkterne i hhv. slutningen af 0. klasse og midten af 1. klasse indsat to prædiktorer. Således står der ud for slutningen af 0. klasse først Dynamisk Afkodningstest A og de tilhørende AUC-værdier. Derefter står bogstavkendskab, som i slutningen af 0. klasse kunne bidrage yderligere til forudsigelsen efter kontrol for elevernes scorer på den dynamiske afkodningstest, samt de AUC-værdier, der opnåedes på basis af disse to prædiktorer tilsammen. Ligeledes er der ud for midten af 1. klasse anført både afkodningspræcision og bogstavkendskab samt de tilhørende AUC-værdier.

Som det fremgår af de rapporterede mål for forholdet mellem sensitivitet og specificitet (AUC-værdierne) i tabel 21 samt af de øvrige rapporterede analyser i de foregående afsnit (5.4-5.7), øges den samlede præcision i forudsigelsen af alvorlige afkodningsvanskeligheder gradvist fra efteråret 0. klasse frem til slutningen af 1. klasse. Dette svarer overordnet til, hvad andre undersøgelser (fx Catts m.fl., 2009; Nielsen & Poulsen, 2012, Poulsen m.fl., 2015) har vist: jo tidligere man forsøger at udpege elever med risiko for udvik-

ling af læsevanskeligheder, jo mere usikkert er grundlaget for vurderingen af, hvorvidt eleverne vil udvikle vanskeligheder eller ej.

Som tidligere nævnt (jf. afsnit 5.4) var elevernes scorer på kombinationsmålet for lydlige kompetencer fra Sprogvurderingsmaterialet i efteråret 0. klasse den stærkeste enkelte prædikator af, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. Imidlertid lå AUC-værdierne for den kombinerede score på lydlige kompetencer fra Sprogvurderingsmaterialet kun i størrelsesordenen 0,72-0,74 (jf. tabel 21 ovenfor) og dermed under den grænse ($AUC \geq 0,80$), der internationalt anbefales for testbatterier med henblik på tidlig identifikation af elever i risiko for udvikling af læsevanskeligheder.

Den aktuelle undersøgelse viste, at en traditionel, statisk test af forlydsidentifikation i efteråret 0. klasse kunne bidrage yderligere til forudsigelsen og således gøre den samlede forudsigelse af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse mere præcis. I efteråret 0. klasse kunne ingen af de øvrige traditionelle test eller den dynamiske test af forlydsidentifikation derudover bidrage signifikant til forudsigelsen. Når man inddrog den statiske test af forlydsidentifikation som supplement til kombinationsmålet for lydlige kompetencer fra Sprogvurderingsmaterialet, var den samlede AUC-værdi kun på 0,78-0,79 og dermed stadig under den anbefalede grænse på 0,80. I modsætning hertil lå den forudsigelse, der i den aktuelle undersøgelse kunne opnås i både slutningen af 0. klasse og i midten og slutningen af 1. klasse med AUC-værdier i størrelsesordenen 0,84-0,97 et godt stykke over den anbefalede grænse. Det er ikke overraskende, at forudsigelsen af senere vanskeligheder på basis af testning i efteråret 0. klasse var væsentligt mindre præcis end på basis af senere testning. Den relativt upræcise forudsigelse på basis af testning i efteråret 0. klasse ligger på linje med tidligere undersøgelser, som har vist, at det er forbundet med betydelig usikkerhed at forsøge at forudsige læsevanskeligheder på individniveau på basis af testning, der gennemføres lang tid før påbegyndelsen af den formelle læseundervisning, herunder fra begyndelsen af børnehaveklassen i Danmark (fx Elbro, Borstrøm & Petersen, 1998; Juul, 2007). Under alle omstændigheder må det bemærkes, at den aktuelle langtidsundersøgelses resultater vedr. forudsigelsesværdien af Sprogvurderingsmaterialet (eller dele heraf) snart primært vil have historisk interesse, da Sprogvurderingsmaterialet forventes at blive udsendt i en revideret version (jf. Center for Børnesprog, 2014).

I slutningen af 0. klasse var elevernes præstationer på en traditionel afkodningstest (højtlesning af ordlister) ikke nogen god prædikator af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. Dette var forudsigeligt, da den formelle læseundervisning på dette tidspunkt endnu ikke var påbegyndt eller kun havde stået på i relativt kort tid. Som tidligere nævnt er elever, der allerede på dette tidspunkt kan læse mange enkeltord korrekt, næppe i risiko for senere at udvikle afkodningsvanskeligheder. Men en meget stor del af eleverne kan før og omkring påbegyndelsen af den formelle læseundervisning af gode grunde slet ikke læse nogen ord, og langt fra alle ikke-læserne er i risiko for at udvikle afkodningsvanskeligheder (jf. afsnit 2.2). Da traditionelle læsetest altså er for svære til at kunne bruges til udpegning af egentlige risikoelever før og omkring påbegyndelse af læseundervisningen, forsøger man i stedet traditionelt at vurdere elevernes forud-

sætninger for udvikling af læsefærdigheder, bl.a. ved hjælp af test af elevernes bogstavkendskab og fonologiske opmærksomhed. Som både tidligere undersøgelser (fx Catts m.fl., 2009; Poulsen m.fl., 2015) og den aktuelle undersøgelse har påvist, giver sådanne test ikke på individniveau nogen særlig præcis udpegning af netop de elever, der faktisk senere udvikler afkodningsvanskeligheder. Den aktuelle undersøgelse dokumenterer, at det som formodet kan betale sig at inddrage en dynamisk test af afkodning i slutningen af 0. klasse med henblik på identifikation af elever i risikogruppen. I slutningen af 0. klasse var en dynamisk afkodningstest således den bedste af de anvendte test til at forudsige elevernes afkodningsfærdigheder i slutningen af 2. klasse. Dette gjaldt ikke blot for elever med dansk som modersmål, men også for elever med dansk som andetsprog, hvor det ellers ofte er ekstra vanskeligt at vurdere, om de er i risiko for udvikling af afkodningsvanskeligheder. Selv når man i forvejen kendte elevernes resultater på den slags test, som man traditionelt bruger til at udpege elever i risiko for udvikling af afkodningsvanskeligheder, bidrog den dynamiske afkodningstest med ekstra forudsigelsesværdi. Dette er i både dansk og international sammenhæng et væsentligt resultat med betydelig praktisk relevans. Samtidig viste undersøgelsen imidlertid også, at den dynamiske afkodningstest ikke gjorde alle de traditionelle test overflødige. Forudsigelsen af, hvorvidt eleverne i slutningen af 2. klasse ville udvise alvorlige afkodningsvanskeligheder eller ej, blev således mere præcis, hvis man ud over den dynamiske afkodningstest i slutningen af 0. klasse også inddrog en traditionel test af bogstavkendskab.

På dette grundlag anbefales det, at der i den kommende testvejledning med henblik på identifikation af risikoelever i slutningen af 0. klasse tages udgangspunkt i elevernes resultater på en test af bogstavkendskab (bogstavbenævnelse) og den dynamiske afkodningstest. Selv om en traditionel afkodningstest med højtlesning af ord som nævnt ikke i sig selv på dette tidspunkt er en god prædikator af senere afkodningsvanskeligheder, anbefales det endvidere, at en individuel testsession med henblik på identifikation af risikoelever i slutningen af 0. klasse i praksis indledes med at lade eleven forsøge at læse enkeltord højt. Formålet med denne sidste anbefaling er at give testvejlederne hjælp til at frasortere elever, som allerede kan læse adskillige ord og som derfor ikke er i risiko for at udvikle afkodningsvanskeligheder. Derved kan det undgås, at alle elever i slutningen af 0. klasse pr. automatik testes individuelt med den dynamiske afkodningstest og bogstavkendskabstesten, hvilket ville være unødigt tids- og ressourcekrævende. Endvidere anbefales det ligeledes af praktiske grunde, at de elever, der kun kan læse ganske få ord korrekt, efterfølgende gennemfører bogstavkendskabstesten, som normalt kun varer et par minutter. Kun for de elever, der scorer under et vist niveau på bogstavkendskabstesten, anbefales det at gennemføre den mere tidskrævende dynamiske afkodningstest, som kan vare omkring et kvarter.

En tidligere dansk undersøgelse har vist, at en afdækning af elevernes afkodningsfærdigheder ved hjælp af højtlesning af lister med rigtige ord og nonord fra omkring midten af 1. klasse kunne danne grundlag for en vurdering af, hvorvidt en elev var i risiko for at udvikle afkodningsvanskeligheder (Poulsen m.fl., 2015). Den aktuelle undersøgelse bekræfter dette. I både midten og slutningen af 1. klasse kunne man således alene

på basis af traditionelle afkodningstest med højtlesning af Elbros lister med ord og nonord med en relativt høj grad af præcision forudsige, hvorvidt eleverne ville udvise alvorlige afkodningsvanskeligheder eller ej i slutningen af 2. klasse (jf. afsnit 5.6-5.7). Derudover kunne en traditionel bogstavkendskabstest i midten af 1. klasse bidrage til en endnu mere præcis forudsigtelse af alvorlige afkodningsvanskeligheder, og det kan derfor anbefales også at inddrage bogstavkendskab som grundlag for udpegning af risikoelever på dette tidspunkt. Efter kontrol for elevernes præstationer på de traditionelle afkodningstest samt bogstavkendskabstest i midten af 1. klasse kunne en dynamisk afkodningstest ikke bidrage yderligere til forudsigtelsen.

Som tidligere beskrevet var det i hhv. midten og slutningen af 1. klasse forskellige afkodningsmål, der i den aktuelle undersøgelse var de bedste prædiktorer af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. I midten af 1. klasse var elevernes afkodnings*præcision* således den bedste enkelte prædiktor af alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, hvorimod elevernes afkodnings*effektivitet* i slutningen af 1. klasse var den bedste enkelte prædiktor (jf. afsnit 5.6-5.7 og tabel 21 ovenfor). På dette grundlag anbefales det, at der i den kommende testvejledning med henblik på identifikation af risikoelever i midten af 1. klasse tages udgangspunkt i elevernes afkodningspræcision, mens det i slutningen af 1. klasse anbefales at tage udgangspunkt i deres afkodningseffektivitet.⁵

Som tidligere nævnt (jf. afsnit 2.1) er elever med dansk som andetsprog som udgangspunkt ikke i særlig risiko for at få alvorlige afkodningsvanskeligheder. I overensstemmelse hermed viste den aktuelle undersøgelse, at de gennemsnitlige præstationer på de traditionelle og de dynamiske afkodningstest *ikke* var signifikant forskellige for hhv. deltagere med dansk som modersmål og deltagerne med dansk som andetsprog (jf. afsnit 5.1).

Imidlertid scorede deltagerne med dansk som andetsprog gennemsnitligt signifikant lavere end deltagerne med dansk som modersmål på sproglige mål som test af ordforråd. Ordforråd var i den aktuelle undersøgelse ikke nogen stærk prædiktor af elevernes afkodningsfærdigheder. Dette er i overensstemmelse med andre undersøgelser, som har vist, at skoleelevers ordforråd kun i moderat grad er relateret til deres afkodningsfærdigheder, hvorimod ordforråd korrelerer relativt højt med deres (senere) tekstforståelse (fx Ricketts, Nation & Bishop, 2007; Verhoeven, van Leeuwe & Vermeer, 2011). Selv om elever med begrænset ordforråd altså ikke er i særlig risiko for at udvikle alvorlige afkodningsvanskeligheder, er de i særlig risiko for at udvikle alvorlige vanskeligheder med *tekstforståelse* (jf. Droop & Verhoeven, 2003; Nielsen, 1998; Rønberg, 2014). Således kan en del af de elever, som i vores undersøgelse ikke udviste afkodningsvanskeligheder, meget vel tænkes senere at få alvorlige problemer med at forstå tekster som følge af et begrænset ordforråd. Deltagerne i den aktuelle undersøgelse var imidlertid ikke nået langt nok i deres læseudvikling til, at det var muligt at opnå pålidelig information om sammenhængen mellem deres ordforråd og tekstforståelse, da elevers tekstforståelse på dette relativt tidlige stadium i læseudviklingen (slutningen af 2. klasse) stadig i høj

⁵ Som tidligere nævnt er det i praksis lettere at måle afkodningspræcision end afkodningseffektivitet. Vi har derfor drøftet med projektets følgegruppe, om det er hensigtsmæssigt at anbefale testvejledere at bruge mål for afkodningseffektivitet i slutningen af 1. klasse. Følgegruppen støttede denne anbefaling.

grad afhænger af deres afkodningsfærdigheder. Derfor fastholder vi undersøgelsens fokus på forudsigelse af afkodningsvanskeligheder.

I det følgende vil vi redegøre for, hvordan man i praksis på basis af de vigtigste prædiktorer (jf. tabel 21) kan afgrænse elever i risiko for udvikling af alvorlige afkodningsvanskeligheder på hvert af de tre tidlige testtidspunkter, hvor resultaterne af langtidsundersøgelsen indikerer, at man kan opnå en præcision i forudsigelsen, der ligger over den internationalt anbefalede grænse (dvs. i slutningen af 0. klasse samt i midten og slutningen af 1. klasse).

6. Mulige afgrænsninger og deres implikationer

6.1. Præcision i tidlig identifikation af elever i risiko for alvorlige afkodningsvanskeligheder ved forskellige niveauer af sensitivitet

Vi har af hensyn til overskueligheden valgt at fokusere på UVM-ordblindetesten som den afhængige variabel i det følgende, da denne test må betragtes som den nye ”gold standard” med hensyn til afdækning af afkodningsvanskeligheder i Danmark.

Hvis man på basis af testresultater i 0. og 1. klasse så vidt muligt vil forsøge at udpege de elever, der er i risiko for senere at udvikle alvorlige afkodningsvanskeligheder, står man i det dilemma, at man samtidig fejlagtigt kan komme til at udpege en stor gruppe elever som værende i risikogruppe (jf. afsnit 2.2 i indledningen). Dette dilemma fremgår af den nedenstående tabel 22, som med UVM-ordblindetesten som afhængig variabel for hvert af de tre tidlige testtidspunkter (hhv. slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse) ved forskellige niveauer af sensitivitet (hhv. 60, 70, 80 og 90 %) viser, hvor stor en andel af falsk positive (falske alarmer) der følger med. Fx kan man se i tabel 22, at hvis man i slutningen af 0. klasse korrekt vil udpege 60 % af de elever, der i slutningen af 2. klasse viser sig at udvise alvorlige afkodningsvanskeligheder vurderet med UVM-ordblindetesten, indebærer det en rate af falsk positive på 5 %. Hvis man i stedet for 60 % i slutningen af 0. klasse korrekt vil udpege fx 80 % af de elever, der i slutningen af 2. klasse viser sig at have alvorlige afkodningsvanskeligheder vurderet med UVM-ordblindetesten, får man en væsentligt højere rate af falsk positive (11 %). Og hvis man i slutningen af 0. klasse vil udpege så mange som 90 % af de elever, der i slutningen af 2. klasse viser sig at have alvorlige afkodningsvanskeligheder, får man en endnu højere andel af falsk positive (35 %). Som det fremgår af resten af tabel 22 med UVM-ordblindetesten som afhængig variabel er der tale om et generelt billede: Jo flere egentlige risikoelever (sandt positive) man vil udpege, jo højere en andel af falske alarmer (falsk positive) følger der med. Vi har i bilag 3 vedlagt tilsvarende tabeller med hhv. Elbros ordlister og DVO-ordblindetesten som afhængige variable, og her ses det samme billede.

Forudsigtelse af: UVM-ordblindetest (afkodningseffektivitet slut 2. kl.)			Sensitivitet (raten af sandt positive)			
			,60	,70	,80	,90
Prædiktions-tidspunkt	Prædiktorer	AUC	Andel falsk positive	Andel falsk positive	Andel falsk positive	Andel falsk positive
Slut 0. kl.	Dynamisk afkodning + bogstavkendskab	,92	,05	,06	,11	,35
Midt 1. kl.	Afkodningspræcision + bogstavkendskab	,96	,02	,06	,07	,10
Slut 1. kl.	Afkodningseffektivitet	,97	,03	,06	,08	,13

Tabel 22. Præcision i tidlig identifikation af elever i risiko for alvorlige afkodningsvanskeligheder ved forskellige niveauer af sensitivitet og kategorisering på basis af UVM-ordblindetest (afkodningseffektivitet) i slutningen af 2. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

De praktiske konsekvenser af valget af sensitivitetsniveau og den tilhørende andel af falsk positive kan illustreres med det følgende eksempel, hvor vi inddrager en fiktiv skoles årgang med 100 børnehaveklasse-elever. Hvis vi som tidligere tager udgangspunkt i Undervisningsministeriets afgrænsning ved 8-percentilen på UVM-ordblindetesten, kan det antages, at omkring 8 af de fiktive 100 elever i 2.-3. klasse vil udvise så alvorlige afkodningsvanskeligheder, at de på UVM-ordblindetesten vil placere sig i den røde ordblindegruppe. De resterende 92 af de 100 elever kan antages at placere sig svarende til den gule eller den grønne gruppe på UVM-ordblindetesten i 2.-3. klasse.

Med en sensitivitet på 70 % i slutningen af 0. klasse vil man som risikoelever i slutningen af 0. klasse korrekt få udpeget 70 % (svarende til i runde tal 6 elever) af de 8 elever, der i det tænkte eksempel ender med at placere sig i den røde ordblindegruppe på UVM-ordblindetesten i 2.-3. klasse og kan betegnes som sandt positive. Med en sensitivitet på 70 % vil man samtidig blandt de 92 elever, der i det tænkte eksempel i 2.-3. klasse placerer sig i den gule eller den grønne gruppe på UVM-ordblindetesten, i slutningen af 0. klasse fejlagtigt få udpeget 6 % (svarende til i runde tal 6 elever) som risikoelever (jf. andelen af falsk positive på 6 % i tabel 22 ved et sensitivitetsniveau på 70). Med en sensitivitet på 70 % vil man i det tænkte eksempel med 100 elever i slutningen af 0. klasse få udpeget i runde tal i alt omkring 11-12 elever (i runde tal 6 + 6) som risikoelever, der bør tildeles særlig støtte i forbindelse med læseundervisningen. Samtidig vil man overse 30 % af de elever, der faktisk i 2.-3. klasse placerer sig i den røde ordblindegruppe. I det tænkte eksempel vil rundt regnet 2 ud af de 8 elever, der faktisk senere udviser alvorlige afkodningsvanskeligheder og placerer sig i den røde ordblindegruppe, altså fejlagtigt blive erklæret uden for risiko i slutningen af 0. klasse.

Hvis man i slutningen af 0. klasse i stedet vælger et sensitivitetsniveau svarende til 90 %, vil man korrekt få udpeget 90 % (svarende til i runde tal 7 elever) ud af de 8 elever, der i det tænkte eksempel ender med at placere sig i den røde ordblindegruppe på UVM-ordblindetesten i 2.-3. klasse. Man vil således kun overse 10 % (svarende til i runde tal 1 elev) ud af de 8 elever, der senere faktisk placerer sig i den røde ordblindegruppe. Med en sensitivitet på 90 % vil man imidlertid samtidig som risikoelever fejlagtigt få udpeget 35 %

af de 92 elever, der i det tænkte eksempel i 2.-3. klasse placerer sig i den gule eller den grønne gruppe på UVM-ordblindetesten (svarende til i runde tal 32 elever) (jf. andelen af falsk positive på 35 % i tabel 22 ved et sensitivitetsniveau på 90). Med en sensitivitet på 90 % vil man i det tænkte eksempel med 100 elever således i slutningen af 0. klasse få udpeget i alt omkring 39 elever (7 + 32) som risikoelever, der bør tildeles særlig støtte i forbindelse med læseundervisningen. Af disse 39 elever er de 32 elever falske alarmer (falsk positive), idet de trods den tidlige udpegning som risikoelever faktisk *ikke* udviser alvorlige afkodningsvanskeligheder senere. Når andelen af falsk positive er høj (som ved sensitivitetsniveauet på 90 i slutningen af 0. klasse i dette eksempel), kan det have store omkostninger. Dels bliver der mange elever og forældre, som bliver unødigt bekymrede; dels bliver det vanskeligt at tildele ret mange ekstra ressourcer til netop de elever, der har størst behov.

I den kategorisering, der danner basis for tabel 22 ovenfor, er eleverne i slutningen af 2. klasse i lighed med i de tidligere rapporterede regressionsanalyser (jf. kapitel 5) kun opdelt efter, hvorvidt de havde alvorlige afkodningsvanskeligheder (svarende til den røde gruppe i UVM-ordblindetesten) eller ej. Imidlertid udgør afkodningsfærdigheder et kontinuum, hvor en del elever har mere moderate afkodningsvanskeligheder, og det er relevant at vide, hvorvidt nogle af de elever, der umiddelbart falder uden for kategorien med alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, faktisk har en moderat grad af afkodningsvanskeligheder (svarende til den gule gruppe i UVM-ordblindetesten). Dette spørgsmål belyser vi i de følgende analyser, hvor vi opererer med to grupper af elever med afkodningsvanskeligheder svarende til opdelingen på UVM-ordblindetesten, hvor elever, der placerer sig svarende til eller under 8-percentilen, kategoriseres som ordblinde (svarende til den røde gruppe), mens elever, der placerer sig mellem 8- og 20-percentilen, kategoriseres som havende usikker fonologisk kodning (svarende til den gule gruppe) (jf. Ordblindetestkonsortiet, 2014; Ordblindetestkonsortiet, 2015). Som tidligere beskrevet blev 8- og 20-percentilerne overført til vores sample, således at vi i analyserne kunne anvende scoregrænser, der passede med, at deltagerne i den aktuelle undersøgelse kun gik i slutningen af 2. klasse (jf. afsnit 5.3.3). Vi beregnede således, at 8- og 20-percentilen i en formodet repræsentativ gruppe i slutningen af 2. klasse svarede til samlede testscorer på UVM-ordblindetesten på hhv. 57 og 63. Som tidligere beskrevet i afsnit 5.3.3 blev de deltagere, der på UVM-ordblindetesten opnåede en afkodningseffektivitetsscore på 57 eller derunder, følgelig kategoriseret som havende alvorlige afkodningsvanskeligheder (svarende til den røde gruppe i vejledningen til UVM-ordblindetesten). Ligeledes blev de deltagere, der på UVM-ordblindetesten opnåede en samlet score mellem 58 og 63, kategoriseret som havende moderate afkodningsvanskeligheder (svarende til den gule gruppe).

6.2. Mulige øvre grænseværdier mhp. udpegnings af elever i risiko for alvorlige afkodningsvanskeligheder

Nedenfor har vi for hhv. slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse ”oversat” forskellige niveauer af sensitivitet til konkrete scoregrænser på de anbefalede test med henblik på udpegnings af elever i risiko for alvorlige afkodningsvanskeligheder for hvert af disse tre testtidspunkter (jf. tabel 23-25). Af tabellerne fremgår for hver af de angivne mulige scoregrænser for det første den procentvise samlede andel af elever, som ved den pågældende scoregrænse ville blive udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder og som i praksis på dette grundlag skulle tildeles særlig opmærksomhed. Denne andel udgøres af både korrekt og fejlagtigt udpegede risikoelever (jf. de praktiske konsekvenser af valget af sensitivitetsniveau, som blev illustreret i eksemplet med de 100 fiktive elever i afsnit 6.1 ovenfor). For det andet vises den procentvise andel af elever med alvorlige afkodningsvanskeligheder i 2. klasse, der ved den pågældende scoregrænse på et tidligere tidspunkt i indskoling ville blive udpeget *korrekt* som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder, idet de i slutningen af 2. klasse faktisk placerede sig svarende til den røde ordblindegruppe og altså kunne kategoriseres som sandt positive. For det tredje viser tabellerne for hver af grænseværdierne for gruppen af elever uden alvorlige vanskeligheder i slutningen af 2. klasse den procentvise andel, der ved den pågældende scoregrænse ville blive *fejlagtigt* udpeget til risikogruppen, idet de i slutningen af 2. klasse faktisk ikke udviste *alvorlige* afkodningsvanskeligheder og altså umiddelbart kunne betegnes som falsk positive. Denne kategori dækker som nævnt imidlertid ikke blot elever, der placerer sig svarende til den grønne normalgruppe på UVM-ordblindetesten, men også elever, der placerer sig svarende til den gule gruppe med usikker fonologisk kodning. Vi har derfor i tabel 23-25 ud over den procentvise andel af elever, der ikke faldt i den røde ordblindegruppe, specifikt angivet den procentvise andel af elever, der faldt i den gule gruppe og altså udviste *moderate* afkodningsvanskeligheder.

6.2.1. Mulige øvre grænseværdier i slutningen af 0. klasse

Som nævnt i afsnit 5.8 kan det på baggrund af analyserne rapporteret i afsnit 5.5 anbefales, at der med henblik på identifikation af risikoelever i slutningen af 0. klasse tages udgangspunkt i elevernes resultater på såvel Dynamisk Afkodningstest A som test af bogstavkendskab (bogstavbenævnelse). En logistisk regressionsanalyse med UVM-ordblindetesten i slutningen af 2. klasse som afhængig variabel og hhv. den dynamiske afkodningstest og bogstavkendskabstesten i slutningen af 0. klasse som uafhængige variable viste, at de to prædiktorer havde præcis samme vægt i forudsigelsen. B-værdien var således 0,184 for både den dynamiske afkodningstest og for bogstavkendskabstesten. På denne baggrund er det mest oplagt simpelt hen at lægge scorerne fra hver af de to test sammen til én kombineret score, som så kan danne grundlag for vurderingen af, hvorvidt elever i slutningen af 0. klasse er i risiko for at udvikle alvorlige afkodningsvanskeligheder eller ej. Da den maksimale samlede score på Del 2 og 3 i den dynamiske afkodningstest er 32 og den maksimale

gennemsnitlige score på bogstavkendskabstesten er 29, bliver den maksimale kombinerede score på de to test 51. Den kombinerede score for den dynamiske afkodningstest og bogstavkendskabstesten er anvendt i de analyser, der rapporteres i dette afsnit.

I tabel 23 vises der for slutningen af 0. klasse nogle mulige øvre grænseværdier svarende til de tidligere nævnte niveauer af sensitivitet (60 %, 70 %, 80 % og 90 %) på basis af elevernes kombinerede scorer på den dynamiske afkodningstest og bogstavkendskabstesten. For hver af grænseværdierne er der i den anden kolonne angivet den procentvise samlede andel af elever, som ved den pågældende scoregrænse ville blive udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder. Denne andel rummer som nævnt ovenfor både de korrekt og fejlagtigt udpegede risikoelever. I den tredje kolonne står den procentvise andel af elever med alvorlige afkodningsvanskeligheder i 2. klasse, der ved den pågældende scoregrænse i slutningen af 0. klasse ville blive udpeget *korrekt* som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder, idet de i slutningen af 2. klasse faktisk placerede sig svarende til den røde ordblindegruppe og altså kunne kategoriseres som sandt positive. I den fjerde kolonne angives for gruppen af elever uden alvorlige vanskeligheder i slutningen af 2. klasse den procentvise andel, der ved den pågældende scoregrænse ville blive *fejlagtigt* udpeget til risikogruppen, idet de i slutningen af 2. klasse faktisk ikke udviste *alvorlige* afkodningsvanskeligheder og altså umiddelbart kunne betegnes som falsk positive. Denne kategori indeholder som nævnt imidlertid ikke blot elever, der placerede sig svarende til den grønne normalgruppe på UVM-ordblindetesten i slutningen af 2. klasse, men også elever, der placerede sig svarende til den gule gruppe med usikker fonologisk kodning. Vi har derfor i den fjerde kolonne udover den procentvise andel af elever, der ikke faldt i den røde ordblindegruppe, specifikt angivet den procentvise andel af elever, der faldt i den gule gruppe og altså udviste *moderate* afkodningsvanskeligheder i slutningen af 2. klasse.

Som nævnt ovenfor (afsnit 6.1) kategoriseres elever, der i slutningen af 2. klasse på UVM-ordblindetesten placerer sig svarende til eller under 8-percentilen for en formodet repræsentativ gruppe, som ordblinde (svarende til den røde gruppe), mens elever, der placerer sig mellem 8- og 20-percentilen for en formodet repræsentativ gruppe, kategoriseres som havende usikker fonologisk kodning (svarende til den gule gruppe). Som tidligere beskrevet blev afgrænsningerne svarende til 8- og 20-percentilerne for et formodet repræsentativt sample overført til vores fulde sample (jf. afsnit 5.3.3), og disse afgrænsninger danner også grundlag for beregningerne vist nedenfor.

Øvre grænseværdier i slutningen af 0. kl.	Kategorisering på basis af UVM-ordblindetest i slutningen af 2. kl.		
	<i>Elever udpeget til risikogruppe i slutningen af 0. klasse (både korrekt og fejlagtigt udpegede risikoelever)</i>	<i>Korrekt udpegede risikoelever (Ordblinde = rød gruppe i 2. kl.) (Begrundet bekymring)</i>	<i>Fejlagtigt udpegede risikoelever (Ikke ordblinde = grøn og gul gruppe i 2. kl.) (Overflødig bekymring)</i>
18	9 %	60 %	5 % (heraf de 2 % i gul gruppe)
21	11 %	70 %	6 % (heraf de 3 % i gul gruppe)
25	17 %	80 %	11 % (heraf de 4 % i gul gruppe)
34	39 %	90 %	35 % (heraf de 10 % i gul gruppe)

Tabel 23. Eksempler på mulige øvre grænseværdier for risikogrupper i slutningen af 0. kl. vurderet med en kombination af Dynamisk Afkodningstest A og bogstavkendskabstest.

Beregningerne er baseret på det fulde sample ($N = 158$).

Som vist i tabel 23 ovenfor ville en øvre scoregrænse på 18 i slutningen af 0. klasse indebære, at elever med en kombineret score mellem 0 og 18 på basis af den dynamiske afkodningstest og bogstavkendskabstesten ville blive kategoriseret som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder. Som det fremgår af tabel 23 ovenfor, ville kun 60 % af de elever, der i slutningen af 2. klasse placerede sig svarende til den røde ordblindegruppe på UVM-ordblindetesten, med en øvre scoregrænse på 18 blive korrekt udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder i slutningen af 0. klasse. Samtidig ville en øvre scoregrænse på 18 i slutningen af 0. klasse indebære, at i alt 5 % ville blive fejlagtigt udpeget som risikoelever. Dog ville 2 ud af de i alt 5 % af de fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder.

Hvis der vælges en højere øvre scoregrænse på 21 i slutningen af 0. klasse, ville 70 % af de elever, der i slutningen af 2. klasse placerede sig svarende til den røde ordblindegruppe på UVM-ordblindetesten, blive korrekt udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder i slutningen af 0. klasse (jf. tabel 23 ovenfor). Samtidig ville i alt 6 % blive fejlagtigt udpeget som risikoelever. Dog ville halvdelen (3 % ud af de i alt 6 %) af de fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder.

Hvis den øvre scoregrænse i stedet sættes op til 25, ville det medføre, at 80 % af de elever, der i slutningen af 2. klasse placerede sig svarende til ordblindekategorien (den røde gruppe) på UVM-ordblindetesten, ville blive korrekt udpeget som risikoelever i slutningen af 0. klasse (jf. tabel 23). Samtidig med, at man altså med denne højere scoregrænse ville opnå at få udpeget en højere andel af risikoelever korrekt, ville man imidlertid også få en noget højere andel af falsk positive, idet 11 % nu ville blive fejlagtigt udpeget som risikoelever. Dog ville 4 % ud af de i alt 11 % fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder i slutningen af 2. klasse.

Hvis den øvre scoregrænse sættes yderligere op til 34, ville man korrekt få udpeget 90 % af de elever, der i slutningen af 2. klasse placerede sig svarende til ordblindekategorien (den røde gruppe) på UVM-ordblindetesten, som risikoelever i slutningen af 0. klasse (jf. tabel 23). Samtidig ville andelen af de fejlagtigt udpegede risikoelever (de falsk positive) imidlertid stige kraftigt til i alt 35 % (hvoraf de 10 % dog placerede sig svarende til den gule gruppe på UVM-ordblindetesten og altså udviste en moderat grad af afkodningsvanskeligheder i slutningen af 2. klasse).

Det kan også være relevant at se på, hvor stor en risiko de enkelte elever, der ved en given afgrænsning på de anbefalede test placeres i risikogruppen i slutningen af 0. klasse, faktisk har for i slutningen af 2. klasse at udvise afkodningsvanskeligheder. Dette kan udledes på basis af tallene i tabel 23. Ved fx en øvre scoregrænse på 25 (svarende til 80 % sensitivitet) udpeges i alt 17 % til risikogruppen i slutningen af 0. klasse. Ud af disse 17 % placerer de 6 % (altså ca. 35 %) sig i slutningen af 2. klasse i den røde gruppe på UVM-ordblindetesten, mens 4 % ud af de 17 % udpegede risikoelever (altså ca. 24 %) placerer sig i den gule gruppe på UVM-ordblindetesten i slutningen af 2. klasse. Det vil sige, at ved en øvre scoregrænse på 25 ville omkring 41 % af de elever, der udpeges til risikogruppen i 0. klasse, faktisk alligevel udvikle deres afkodningsfærdigheder så godt, at de i slutningen af 2. klasse ville placere sig svarende til den grønne gruppe på UVM-ordblindetesten.

6.2.2. Mulige øvre grænseværdier i midten af 1. klasse

Som nævnt i afsnit 5.8 kan det på baggrund af analyserne rapporteret i afsnit 5.6 anbefales, at der med henblik på identifikation af risikoelever i midten af 1. klasse tages udgangspunkt i elevernes afkodningspræcision målt med højt-læsning af Elbros lister med rigtige ord og nonord i kombination med test af bogstavkendskab (bogstavbenævnelse). En logistisk regressionsanalyse med UVM-ordblindetesten i slutningen af 2. klasse som afhængig variabel og hhv. afkodningspræcision og bogstavkendskab i midten af 1. klasse som uafhængige variable viste, at vægten (B-værdien) af afkodningspræcision baseret på gennemsnittet af elevens scorer på Elbros lister med rigtige ord og nonord ($B = 6,2$) var godt og vel 30 gange så høj som for bogstavkendskab ($B = 0,2$). I praksis vil det være mest hensigtsmæssigt, at testvejlederen opgør og indtaster elevens afkodningspræcisionsscorer på hhv. ordlister og nonordslister hver for sig i et særligt regneark, og at regnearket blot lægger de to scorer sammen *uden* at beregne gennemsnittet. Dette giver nemlig mulighed for, at det på regnearket markeres for testvejlederen, hvis eleven allerede på listerne med rigtige ord scorer så højt, at eleven placerer sig over risikogrænsen, således at man ikke behøver at gå videre med testningen. På denne baggrund anbefales det at anvende en kombineret score, hvor hvert korrekt læst ord tæller 15 gange så tungt som det gennemsnitlige antal korrekt benævnte store og små bogstaver, som grundlag for vurderingen af, hvorvidt elever i midten af 1. klasse er i risiko for at udvikle alvorlige afkodningsvanskeligheder eller ej. Formlen for den kombinerede score for afkodningspræcision og bogstavkendskab i midten af 1. klasse bliver således: $(15 \times \text{antal korrekt læste rigtige ord og nonord}) + \text{gennemsnitligt antal korrekt benævnte store og små bogstaver}$

små bogstaver. Med denne vægtning er den maksimale afkodningspræcisionsscore på basis af højtlesning af hhv. 30 rigtige ord og 30 nonord i alt 900 (15 x (30 + 30)), mens den maksimale gennemsnitlige score på bogstavkendskabstesten med hhv. store og små bogstaver er 29. Den maksimale kombinerede score på de to test er således 929. Den kombinerede score for afkodningspræcision og bogstavkendskab er anvendt i de analyser, der rapporteres i dette afsnit.

I tabel 24 nedenfor vises der for midten af 1. klasse nogle mulige grænseværdier svarende til en sensitivitet på hhv. 60, 70, 80 og 90 på basis af elevernes afkodningspræcisionsscorer ved højtlesning af Elbros lister med ord og nonord i kombination med bogstavkendskab. For hver af grænseværdierne er der som i den foregående tabel i den anden kolonne angivet den procentvise samlede andel af elever, der ved den pågældende scoregrænse i midten af 1. klasse ville blive udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder (altså den andel, som udgøres af de korrekt og de fejlagtigt udpegede risikoelever tilsammen). Den tredje kolonne viser den andel af elever med alvorlige afkodningsvanskeligheder i 2. klasse, der ved den pågældende scoregrænse i midten af 1. klasse ville blive udpeget korrekt som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder (de sandt positive). Vi har i den fjerde kolonne igen angivet den procentvise andel af elever uden senere alvorlige vanskeligheder, som ved den givne grænse ville blive udpeget som risikoelever, men som i 2. klasse ikke faldt i den røde ordblindegruppe og altså umiddelbart kan betegnes som falsk positive, samt i parentes specifikt angivet den procentvise andel af elever, der faldt i den gule gruppe og altså udviste *moderate* afkodningsvanskeligheder.

Øvre grænseværdier i midten af 1. kl.	Kategorisering på basis af UVM-ordblindetest i slutningen af 2. kl.		
	<i>Elever udpeget til risikogrupper i midten af 1. klasse (både korrekt og fejlagtigt udpegede risikoelever)</i>	<i>Korrekt udpegede risikoelever (Ordblinde = rød gruppe i 2. kl.) (Begrundet bekymring)</i>	<i>Fejlagtigt udpegede risikoelever (Ikke ordblinde = grøn og gul gruppe i 2. kl.) (Overflødig bekymring)</i>
26	7 %	60 %	2 % (heraf de 1 % i gul gruppe)
36	11 %	70 %	6 % (heraf de 2 % i gul gruppe)
43	13 %	80 %	7 % (heraf de 2 % i gul gruppe)
55	16 %	90 %	10 % (heraf de 3 % i gul gruppe)

Tabel 24. Eksempler på mulige øvre grænseværdier for risikogrupper i midten af 1. kl. vurderet med afkodningspræcisionsscorer på basis af højtlesning af Elbros lister med ord og nonord i kombination med bogstavkendskab. Beregningerne er baseret på det fulde sample ($N = 158$).

Som det fremgår af tabel 24 ovenfor, ville en øvre scoregrænse på fx 36 indebære, at 70 % af de elever, der i slutningen af 2. klasse placerede sig svarende til den røde ordblindegruppe på UVM-ordblindetesten, ville blive korrekt udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder i midten af 1. klasse. Samtidig ville i alt 6 % blive fejlagtigt udpeget som risikoelever. Dog ville en tredjedel (2 % ud af de i alt 6 %) af de fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder.

Hvis der i stedet vælges en lidt højere øvre grænse på 43, ville det indebære, at 80 % af de elever, der i slutningen af 2. klasse placerede sig svarende til ordblindekategorien (den røde gruppe) på UVM-ordblindetesten, ville blive korrekt udpeget som risikoelever i midten af 1. klasse (jf. tabel 24). Samtidig ville i alt 7 % blive fejlagtigt udpeget som risikoelever. Dog ville 2 % ud af de i alt 7 % fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise moderate afkodningsvanskeligheder i slutningen af 2. klasse.

Hvis den øvre scoregrænse i stedet sættes op til 55, ville det medføre, at hele 90 % af de elever, der i slutningen af 2. klasse placerede sig svarende til ordblindekategorien (den røde gruppe) på UVM-ordblindetesten, ville blive korrekt udpeget som risikoelever i midten af 1. klasse (jf. tabel 24). Samtidig ville andelen af de fejlagtigt udpegede risikoelever stige til i alt 10 %. Dog ville 3 % ud af de i alt 10 % fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder i slutningen af 2. klasse.

Man kan også spørge, hvor stor en risiko de enkelte elever, der ved en given afgrænsning på de anbefalede test placeres i risikogruppen i midten af 1. klasse, faktisk har for i slutningen af 2. klasse at udvise afkodningsvanskeligheder. Dette kan udledes på basis af tallene i tabel 24. Ved fx en øvre scoregrænse på 55 (svarende til 90 % sensitivitet) udpeges i alt 16 % til risikogruppen i midten af 1. klasse. Ud af disse 16 % placerer de 6 % (altså ca. 38 %) sig i slutningen af 2. klasse i den røde gruppe på UVM-ordblindetesten, mens 3 % ud af de 16 % udpegede risikoelever (altså ca. 19 %) placerer sig i den gule gruppe på UVM-ordblindetesten i slutningen af 2. klasse. Det vil sige, at ved en øvre scoregrænse på 55 ville omkring 43 % af de elever, der udpeges til risikogruppen i midten af 1. klasse, faktisk alligevel udvikle deres afkodningsfærdigheder så godt, at de i slutningen af 2. klasse ville placere sig svarende til den grønne gruppe på UVM-ordblindetesten.

6.2.3. Mulige øvre grænseværdier i slutningen af 1. klasse

Som nævnt i afsnit 5.8 kan det på baggrund af analyserne rapporteret i afsnit 5.7 anbefales, at der med henblik på identifikation af risikoelever i slutningen af 1. klasse tages udgangspunkt i elevernes afkodningseffektivitet målt med højt-læsning af Elbros lister med rigtige ord og nonord. I praksis vil det være mest hensigtsmæssigt, at testvejlederen opgør og indtaster elevens afkodningseffektivitetsscorer på hhv. ordlister og nonordslister hver for sig i et særligt regneark, og at regnearket blot lægger de to scorere sammen (antal korrekt læste rigtige ord pr. minut + antal korrekt læste nonord pr. minut) *uden* at beregne gennemsnittet. Dette giver nemlig mulighed for, at det på regnearket markeres for testvejlederen, hvis eleven allerede på listerne med rigtige ord scorer så højt, at eleven placerer sig over risikogrænsen, således at man ikke behøver at gå videre med testningen.

I tabel 25 nedenfor vises der for slutningen af 1. klasse nogle mulige grænseværdier svarende til de tidligere nævnte niveauer af sensitivitet (60 %, 70 %, 80 % og 90 %) på basis af elevernes afkodningseffektivitetsscorer ved højtlesning af Elbros lister med ord og nonord (antal korrekt læste rigtige ord pr. minut + antal korrekt læste nonord pr. minut).

Øvre grænseværdier i slutningen af 1. kl.	Kategorisering på basis af UVM-ordblindetest i slutningen af 2. kl.		
	<i>Elever udpeget til risikogruppe i slutningen af 1. klasse (både korrekt og fejlagtigt udpegede risikoelever)</i>	<i>Korrekt udpegede risikoelever (Ordblinde = rød gruppe i 2. kl.) (Begrundet bekymring)</i>	<i>Fejlagtigt udpegede risikoelever (Ikke ordblinde = grøn og gul gruppe i 2. kl.) (Overflødig bekymring)</i>
3	8 %	60 %	3 % (heraf de 1 % i gul gruppe)
4	11 %	70 %	6 % (heraf de 2 % i gul gruppe)
6	14 %	80 %	8 % (heraf de 3 % i gul gruppe)
8	19 %	90 %	13 % (heraf de 4 % i gul gruppe)

Tabel 25. Eksempler på mulige øvre grænseværdier for risikogrupper i slutningen af 1. kl. vurderet med afkodningseffektivitetsscorer på basis af højtlesning af Elbros lister med ord og nonord.

Beregningerne er baseret på det fulde sample ($N = 158$).

Som det fremgår af tabel 25 ovenfor, ville en øvre scoregrænse på fx 6 korrekt læste ord pr. minut på Elbros lister med rigtige ord og nonord indebære, at 80 % af de elever, der i slutningen af 2. klasse placerede sig svarende til den røde ordblindegruppe på UVM-ordblindetesten, ville blive korrekt udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder i slutningen af 1. klasse. Samtidig ville 8 % blive fejlagtigt udpeget som risikoelever, hvoraf over en tredjedel (3 % af de i alt 8 %) ville placere sig svarende til den gule gruppe på UVM-ordblindetesten og altså udvise en moderat grad af afkodningsvanskeligheder.

Hvis den øvre scoregrænse i stedet sættes op til gennemsnitligt 8 korrekt læste ord pr. minut, ville det medføre, at 90 % af de elever, der i slutningen af 2. klasse placerede sig svarende til ordblindekategorien (den røde gruppe) på UVM-ordblindetesten, ville blive korrekt udpeget som risikoelever i slutningen af 1. klasse (jf. tabel 25). Samtidig ville andelen af de fejlagtigt udpegede risikoelever (de falsk positive) stige til i alt 13 %. Dog ville 4 % ud af de i alt 13 % fejlagtigt udpegede risikoelever placere sig svarende til den gule gruppe på UVM-ordblindetesten og således udvise moderate afkodningsvanskeligheder i slutningen af 2. klasse.

Man kan endvidere se på, hvor stor en risiko de enkelte elever, der ved en given afgrænsning på de anbefalede test placeres i risikogruppen i slutningen af 1. klasse, faktisk har for i slutningen af 2. klasse at udvise afkodningsvanskeligheder. Dette kan udledes på basis af tallene i tabel 25 ovenfor. Ved fx en øvre scoregrænse på 8 (svarende til 90 % sensitivitet) udpeges i alt 19 % til risikogruppen i slutningen af 1. klasse. Ud af disse 19 % placerer de 6 % (altså ca. 32 %) sig i slutningen af 2. klasse i den røde gruppe på UVM-ordblindetesten, mens 4 % ud af de 19 % udpegede risikoelever (altså ca. 21 %) placerer sig i den gule gruppe på UVM-ordblindetesten i slutningen af 2. klasse. Ved en øvre scoregrænse på 8 ville omkring 47 % af de elever, der udpeges til risikogruppen i slutningen af 1. klasse, således faktisk alligevel udvikle deres afkod-

ningsfærdigheder så godt, at de i slutningen af 2. klasse ville placere sig svarende til den grønne gruppe på UVM-ordblindetesten.

6.3. Det videre arbejde med afgrænsning af risikoelever

Når man på basis af testresultater i 0. og 1. klasse vil forsøge at udpege de elever, der er i særlig risiko for senere at udvikle alvorlige afkodningsvanskeligheder, står man som tidligere nævnt i det dilemma, at man samtidig fejlagtigt nemt kommer til at udpege en del elever, som faktisk ikke får vanskeligheder senere. Hvis man sætter den vejledende øvre scoregrænse på en test mhp. tidlig identifikation af risikoelever højt, får man naturligvis udpeget flere elever som risikoelever og er dermed mere sikker på, at man ikke overser ret mange af de elever, der faktisk har mest brug for ekstra støtte og opmærksomhed. Men det vil automatisk have den omkostning, at flere elever samtidig fejlagtigt udpeges til risikogruppen, således at der bliver flere såkaldte falsk positive, som ikke får vanskeligheder. Som det ses af tabel 23-25 i afsnit 6.2 ovenfor, var der dog en del af de falsk positive (altså de elever, der fejlagtigt blev udpeget som værende i risiko for udvikling af alvorlige afkodningsvanskeligheder), der i slutningen af 2. klasse placerede sig svarende til den gule gruppe med usikker fonologisk kodning på UVM-ordblindetesten. Gennemsnitligt udviste således omkring en tredjedel af de falsk positive *moderate* afkodningsvanskeligheder i slutningen af 2. klasse og var altså relativt svage læsere, selv om de ikke placerede sig svarende til den røde ordblindegruppe.

Som det fremgår af de ovenstående tabeller for de tre prædiktions-tidspunkter, er omfanget af problemet med høje andele af falsk positive størst i slutningen af 0. klasse. Dette er på linje med de analyseresultater, der blev rapporteret i kapitel 5 og svarer overordnet til, hvad også andre undersøgelser (fx Catts m.fl., 2009; Nielsen & Poulsen, 2012, Poulsen m.fl., 2015) har vist: jo tidligere man forsøger at udpege elever i risiko for udvikling af læsevanskeligheder, jo mere usikkert er grundlaget for vurderingen af, hvorvidt eleverne vil udvikle vanskeligheder eller ej.

Det er naturligvis muligt at fastlægge vejledende øvre scoregrænser mhp. udpegning af risikoelever både højere og lavere end i de ovenfor nævnte eksempler for såvel slutningen af 0. klasse som for midten og slutningen af 1. klasse. Hvis Undervisningsministeriet ønsker det, kan vi beregne implikationerne af andre øvre scoregrænser end dem, der er vist i tabellerne ovenfor. Det må imidlertid understreges, at det *ikke* er en forskningsmæssig opgave at fastlægge afgrænsninger i forbindelse med udpegning af risikoelever. Disse afgrænsninger kan få betydelige praktiske og økonomiske konsekvenser, og beslutningen vedrørende disse afgrænsninger må derfor træffes af Undervisningsministeriet. Det er ligeledes op til Undervisningsministeriet at beslutte, om der med henblik på den tidlige udpegning af risikoelever skal skelnes mellem to eller tre forskellige grupper.

Det er således muligt kun at trække én vejledende grænse på hvert af de tre tidlige testtidspunkter (slutningen af 0. klasse samt midten og slutningen af 1. klasse) og altså nøjes med at skelne mellem elever, der antages at være i risiko (evt. rød gruppe) og elever, der ikke antages at være i risiko (evt. grøn gruppe). Evt.

kan en sådan todeling suppleres med en mulighed for at lade testvejlederen eller skolen inddrage supplerende oplysninger i beslutningen om, hvorvidt eleven skal udpeges til risikogruppe. Man kan fx forestille sig, at en elev, der placerer sig i den grønne kategori, men kun ligger lidt over grænsen og samtidig kan antages at være arveligt disponeret for ordblindhed og/eller har udvist påfaldende vanskeligheder i forbindelse med læseforberedende aktiviteter eller læseundervisning, også kan udpeges til risikogruppen.

Alternativt kan der trækkes to vejledende grænser mhp. udpegning af tre forskellige grupper på hvert af de tidlige testtidspunkter (slutningen af 0. klasse samt midten og slutningen af 1. klasse), fx som skitseret i tabel 26 nedenfor med en opdeling i en hhv. grøn, gul og rød gruppe. Evt. kan der også her lægges op til, at testvejlederen eller skolen kan inddrage supplerende oplysninger om ordblindhed i familien og/eller undervisningsobservationer som grundlag for udpegning af risikoelever.

Øvre score-grænse	Gruppe	Vurdering
X-max.	A (grøn)	Eleven er sandsynligvis <i>ikke</i> i risiko for at udvikle alvorlige afkodningsvanskeligheder
Y-Z	B (gul)	Eleven er muligvis i risiko for at udvikle alvorlige afkodningsvanskeligheder
0-W	C (rød)	Eleven er i stor risiko for at udvikle alvorlige afkodningsvanskeligheder

Tabel 26. Skitse af muligt skema med oversigt over placering af elever i tre forskellige grupper på et af de tidlige testtidspunkter.

De konkrete øvre scoregrænser kan for hvert af de tre tidlige testtidspunkter indsættes i skemaet, når de er fastlagt af Undervisningsministeriet.

Som nævnt i afsnit 5.8 anbefales det specielt for *slutningen af 0. klasse*, at den individuelle testprocedure opdeles i de følgende tre trin:

1. Afkodningstest (højt læsning af Elbros ordlister)
2. Bogstavkendskabstest
3. Dynamisk Afkodningstest A

Formålet med denne procedure er at undgå, at alle elever i slutningen af 0. klasse pr. automatik testes individuelt med bogstavkendskabstesten og den dynamiske afkodningstest, hvilket ville være unødigt tids- og resourcekrævende. Med henblik på at begrænse tids- og ressourceforbruget anbefales det som tidligere nævnt, at testlederen stopper testproceduren allerede efter trin 1 for de elever, der i den indledende afkodningstest læser mere end nogle få ord korrekt, da disse elever må anses for at være uden for risiko for udvikling af alvorlige afkodningsvanskeligheder. De elever, der i slutningen af 0. klasse kun kan læse ganske få eller ingen ord korrekt, går derimod videre til trin 2 og gennemfører bogstavkendskabstesten, som normalt kun varer et par minutter. Kun de elever, der scorer under et vist niveau på bogstavkendskabstesten, går derefter

videre til trin 3 og gennemfører den mere tidskrævende dynamiske afkodningstest, som kan vare omkring et kvarter.

Der kan først gives anbefalinger angående de specifikke vejledende scoregrænser for, hvornår testvejlederen skal gå videre i den anbefalede individuelle testprocedure for slutningen af 0. klasse, når Undervisningsministeriet har fastlagt den samlede øvre vejledende grænse for den kombinerede score på bogstavkendskabstesten og den dynamiske afkodningstest i slutningen af 0. klasse (jf. tabel 23 ovenfor med eksempler på øvre scoregrænser og implikationerne af disse scoregrænser).

Specielt for *midten af 1. klasse* anbefales det, at den individuelle testprocedure opdeles i de følgende to trin:

1. Afkodningstest (højtlesning af Elbros lister med ord og nonord)
2. Bogstavkendskabstest

Med henblik på at begrænse tids- og ressourceforbruget kan det anbefales, at testlederen stopper testproceduren allerede efter trin 1 for de elever, der i den indledende afkodningstest læser relativt mange ord korrekt, da disse elever må anses for at være uden for risiko for udvikling af alvorlige afkodningsvanskeligheder. De øvrige elever går derimod videre til trin 2 og gennemfører bogstavkendskabstesten. Også for midten af 1. klasse gælder det, at der først kan gives anbefalinger angående den specifikke vejledende scoregrænse for, hvornår testvejlederen skal gå videre til trin 2 i den anbefalede individuelle testprocedure, når Undervisningsministeriet har fastlagt den samlede øvre vejledende grænse for den kombinerede score på afkodningspræcision og bogstavkendskabstesten i midten af 1. klasse (jf. tabel 24 ovenfor med eksempler på øvre scoregrænser og implikationerne af disse scoregrænser).

For *slutningen af 1. klasse* er den anbefalede individuelle testprocedure med henblik på identifikation af risikoelever mere enkel, idet den kun består af ét trin. Dette ene trin består af afkodningseffektivitet i slutningen af 1. klasse (jf. tabel 25 ovenfor med eksempler på øvre scoregrænser og implikationerne af disse scoregrænser).

På alle tre testtidspunkter kan skolerne inden den anbefalede individuelle testning eventuelt gennemføre en indledende screening med relevante gruppetest for at begrænse antallet af elever, der skal gennemføre de anbefalede individuelle test. Det bemærkes dog, at gruppetestning generelt giver mindre pålidelige resultater end individuel testning. Det kan derfor være problematisk at udelukke elever fra individuel testning alene på grundlag af upåfaldende gruppetestresultater, da elever, som faktisk er i alvorlig risiko for vedvarende afkodningsvanskeligheder, derved kan blive overset. Hvis ikke alle elever skal testes individuelt, må det derfor anbefales, at udvælgelsen ikke blot finder sted på baggrund af gruppetestresultater, men at også observationer fra undervisningen inddrages som grundlag for denne udvælgelse. Det anbefales således, at der gennemføres individuel testning af alle elever, der har vist påfaldende vanskeligheder i forbindelse med den læseforberedende undervisning og/eller selve læseundervisningen og altså ikke blot af elever, der klarer sig dårligt på relevante indledende gruppetest.

7. Undervisningsministeriets beslutninger

I den foregående del af denne projektrapport har vi givet en del af et *empirisk grundlag* for at træffe beslutninger vedr. afgrænsninger af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder. Det er som tidligere nævnt *ikke* en forskningsmæssig opgave at fastlægge de konkrete afgrænsninger til brug ved udpegning af risikoelever, men derimod en politisk beslutning. Beslutningen vedrørende disse afgrænsninger blev derfor truffet af Undervisningsministeriet. Efter aftale med ministeriet resumeres disse beslutninger nedenfor.

Ministeriet besluttede kun at trække én vejledende grænse på hvert af de tre tidlige testtidspunkter (slutningen af 0. klasse samt midten og slutningen af 1. klasse) og således nøjes med at skelne mellem elever, der kan antages at være i risiko for udvikling af alvorlige afkodningsvanskeligheder (rød gruppe) og elever, der ikke kan antages at være i risiko (grøn gruppe). Denne todeling suppleres med en mulighed for at lade testvejlederen eller skolen inddrage andre oplysninger i beslutningen om, hvorvidt eleven skal udpeges til risikogruppen. En elev, der placerer sig i den grønne kategori, men kun ligger lidt over grænsen og samtidig kan antages at være arveligt disponeret for ordblindhed og/eller har udvist påfaldende vanskeligheder i forbindelse med læseforberedende aktiviteter eller læseundervisning, kan således også udpeges til risikogruppen.

Ministeriet fastlagde på grundlag af de rapporterede resultater fra langtidsundersøgelsen nogle vejledende afgrænsninger til brug for udpegning af risikoelever. Disse afgrænsninger indebærer, at man kan forvente, at man i slutningen af 0. klasse korrekt kan udpege omkring 80 % af de elever, der ved slutningen af 2. klasse vil udvise alvorlige afkodningsvanskeligheder. Til sammenligning kan man i midten og slutningen af 1. klasse forvente, at man korrekt kan udpege 90 % af de elever, der senere vil udvise alvorlige afkodningsvanskeligheder. For nærmere oplysninger om implikationerne af de valgte afgrænsninger henvises der til kapitel 6.

De ministerielle beslutninger danner grundlag for anbefalingerne i *Vejledning til tidlig identifikation af elever i risiko for udvikling af alvorlige afkodningsvanskeligheder (herunder ordblindhed)*.

Litteraturhenvisninger

- Arnbak, E. & Borstrøm, I. (2007). Udvikling og afprøvning af procedure til identifikation af elever i risiko for dysleksi. I: B. D. Jandorf (red.). *Værd at vide om ordblindhed*. Virum: Dansk Videnscenter for Ordblindhed. http://www.nota.nu/sites/default/files/Vejledning_til_afgraensning_af_dysleksi.pdf
- Bleses, D., Lum, J., Højen, A., Jørgensen, R. N., Iachine, I., Andersen, M. K., Andersen, C., Jensen, K. Ø. & Vach, W. (2011). Sprogvurderingsmateriale til 3-årige, inden skolestart og i Børnehaveklassen. Metodisk oversigt. Working paper, *Center for Child Language e-prints, No 13*. (http://www.cfb.sdu.dk/DK/forsk/doc/E-print_13_2011.pdf)
- Borstrøm, I. & Petersen, D. K. (2004). *Læseevaluering på begyndertrinnet*. København: Alinea.
- Bridges, M. S. & Catts, H. W. (2011). The use of a dynamic screening of phonological awareness to predict risk for reading disabilities in kindergarten children. *Journal of Learning Disabilities, 44*(4), 330-338.
- Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S. & Bontempo, D. E. (2015). Early Identification of Reading Disabilities Within an RTI Framework. *Journal of Learning Disabilities, 48*(3), 281-297.
- Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S. & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities, 42*(2), 163-176.
- Center for Børnesprog (2014). Revideret version af Sprogvurdering af børn i treårsalderen, inden skolestart og i børnehaveklassen. *Center for Børnesprogs Nyhedsbrev 32*. (http://www.sdu.dk/Om_SDU/Institutter_centre/C_Boernesprog/Nye+Nyhedsbreve/Nyhedsbrev_32_okt2014)
- Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A. & Crouch, R. C. (2010). Selecting At-Risk First-Grade Readers for Early Intervention: Eliminating False Positives and Exploring the Promise of a Two-Stage Gated Screening Process. *Journal of Educational Psychology, 102*(2), 327-340.
- Coventry, W. L., Byrne, B., Olson, R. K., Corley, R. & Samuelsson, S. (2011). Dynamic and Static Assessment of Phonological Awareness in Preschool: A Behavior-Genetic Study. *Journal of Learning Disabilities, 44*(4), 322-329.
- Dansk Videnscenter for Ordblindhed (2006). *Rapport om afprøvning af visitationstest til brug i ordblindeundervisning for voksne*. København: Undervisningsministeriet.
- Daugaard, H. T., Elbro, C. & Gellert, A. S. (2011a). *Dynamisk ordblindetest*. København: Undervisningsministeriet.
- Daugaard, H. T., Elbro, C. & Gellert, A. S. (2011b). *Rapport om udvikling og afprøvning af ordblindetest til voksne med dansk som andetsprog*. København: Undervisningsministeriet. http://laes.hum.ku.dk/centerets_forskning/Projektrapport_udv_test_ordblinde_DSA.pdf

- Droop, M. & Verhoeven, L. (1998). Background knowledge, linguistic complexity, and second-language reading comprehension. *Journal of Literacy Research*, 30(2), 253-271.
- Droop, M. & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38(1), 78-103.
- Elbro, C. (1990). *Differences in Dyslexia. A study of reading strategies and deficits in a linguistic perspective*. København: Munksgaard.
- Elbro, C. (2007). *Læsevanskeligheder*. København: Gyldendal.
- Elbro, C. (2012). Den helt entydige ordblindetest. Ordblindhed i teori og praksis. I: L. Pøhler (red.). *Dysleksi – en fælles nordisk udfordring* (s. 11-21). København: Landsforeningen af Læsepædagoger.
- Elbro, C. (2014). *Læsning og læseundervisning* (3. udgave). København: Hans Reitzels Forlag.
- Elbro, C., Borstrøm, I. & Petersen, D. K. (1998). Predicting Dyslexia from Kindergarten: The Importance of Distinctness of Phonological Representations of Lexical Items. *Reading Research Quarterly*, 33, 36-60.
- Elbro, C., Daugaard, H. T. & Gellert, A. S. (2012). Dyslexia in a second language? A dynamic test of reading acquisition may provide a fair answer. *Annals of Dyslexia*, 62(3), 172-185.
- Elbro, C., Nielsen, I. & Petersen, D. K. (1994). Dyslexia in adults – evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, 44, 205-226.
- Elbro, C. & Petersen, D. K. (2004). Long-Term Effects of Phoneme Awareness and Letter Sound Training: An Intervention Study with Children at Risk for Dyslexia. *Journal of Educational Psychology*, 96(4), 660-670.
- Elbro, C. & Poulsen, M. (2015). *Hold i virkeligheden. Statistik og evidens i uddannelse*. København: Hans Reitzels Forlag.
- Elbro, C. & Scarborough, H. S. (2004). Early identification. I: T. Nunes & P. Bryant (red.). *Handbook of children's literacy* (s. 339-359). Dordrecht: Kluwer Law International.
- Everatt, J., Ocampo, D., Veii, K., Nenopoulou, S., Smythe, I., al Mannai, H. & Elbeheri, G. (2010). Dyslexia in biscriptal readers. I: N. Brunswick, S. McDougall & P. Davies (red.). *Reading and dyslexia in different orthographies* (s. 221-245). New York: Psychology Press.
- Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson, N. J. & Chaparro, E. (2015). An Examination of the Efficacy of a Multitiered Intervention on Early Reading Outcomes for First Grade Students at Risk for Reading Difficulties. *Journal of Learning Disabilities*, 48(6), 602-621.
- Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B. & Caffrey, E. (2011). The Construct and Predictive Validity of a Dynamic Assessment of Young Children Learning to Read: Implications for RTI Frameworks. *Journal of Learning Disabilities*, 44(4), 339-347.

- Gellert, A. S. (2009). *Identifikation og undervisning af voksne ordblinde med dansk som andetsprog – en faglig og pædagogisk udfordring*. København: Undervisningsministeriet og Dansk Videnscenter for Ordblindhed. http://laes.hum.ku.dk/centerets_forskning/UVM-DVO-Rapport_om_ordblinde_DSA.pdf
- Gellert, A. S. & Elbro, C. (2013). Do experimental measures of word learning predict vocabulary development over time? *Learning and Individual Differences*, 26, 1-8.
- Gellert, A. S. & Elbro, C. (2014). Kan en dynamisk afkodningstest i børnehaveklassen bidrage til at forudsige læsevanskeligheder i 1. klasse? Foreløbige resultater fra en langtidsundersøgelse. *Nydanske Sprogstudier*, 47, 9-38.
- Gellert, A. S. & Elbro, C. (2015a). Does a dynamic test of phonological awareness predict early reading difficulties? A longitudinal study from kindergarten through grade 1. *Journal of Learning Disabilities* (online before print). DOI: 10.1177/0022219415609185.
- Gellert, A. S. & Elbro, C. (2015b). Try a little bit of teaching. A dynamic assessment of word decoding as a kindergarten predictor of end of Grade 1 word reading. (Artikelmanuskript under review).
- Gough, P. B., Hoover, W. A. & Peterson, C. L. (1996). Some observations on a simple view of reading. I: C. Cornoldi & J. Oakhill (red.). *Reading comprehension difficulties. Processes and intervention* (s. 1-13). Mahwah, NJ: Erlbaum.
- Grigorenko, E. L. (2009). Dynamic Assessment and Response to Intervention: Two Sides of One Coin. *Journal of Learning Disabilities*, 42(2), 111-132.
- Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic Testing. *Psychological Bulletin*, 124(1), 75-111.
- Hansen, M., Kreiner, S. & Hansen, C. R. (1995). *CHIPS – Children's Problem Solving* (3. udgave). Virum: Hogrefe Psykologisk Forlag.
- Harn, B. A., Linan-Thompson, S. & Roberts, G. (2008). Intensifying Instruction: Does Additional instructional Time Make a Difference for the Most At-Risk First Graders? *Journal of Learning Disabilities*, 41, 115-125.
- Howitt, D. & Cramer, D. (2011). *Introduction to Statistics in Psychology. Fifth Edition*. Harlow: Pearson Education Limited.
- Johnson, E. S., Jenkins, J. R., Petscher, Y. & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities & Practice*, 24(4), 174-185.
- Juul, H. (2007). *Sproglige færdigheder og læseudvikling. En undersøgelse af danske børn fra starten af børnehaveklassen til 2. klasse*. København: Undervisningsministeriet og Center for Læseforskning. <http://holgerjuul.com/status-delrapport2-2007.pdf>
- Juul, H. (2010). *Staveprøve 1*. Virum: Hogrefe Psykologisk Forlag.
- Juul, H. & Møller, L. (2010a). *Ordlæseprøve 1*. Virum: Hogrefe Psykologisk Forlag.
- Juul, H. & Møller, L. (2010b). *Ordlæseprøve 2*. Virum: Hogrefe Psykologisk Forlag.
- Juul, H. (2015). Personlig kommunikation. (E-mail 3.12.15).

- Lyon, G. R., Shaywitz, S. E. & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53, 1-14.
- Mejding, J. & Rønberg, L. F.-J. (2012). *PIRLS 2011: En international undersøgelse om læsekompetence i 4. klasse*. København: Institut for Uddannelse og Pædagogik, Aarhus Universitet.
- Melby-Lervag, M. (2012). The Relative Predictive Contribution and Causal Role of Phoneme Awareness, Rhyme Awareness and Verbal Short-Term Memory in Reading Skills: A Review. *Scandinavian Journal of Educational research*, 56(4), 363-380.
- Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8, 283-298.
- Møller, L. (2012). *Tekstlæseprøve 4*. Virum: Hogrefe Psykologisk Forlag.
- Møller, L. & Juul, H. (2010). *Vejledning til Ordlæseprøve 1-2*. Virum: Hogrefe Psykologisk Forlag.
- National Early Reading Panel (2008). *Developing Early Literacy. A Scientific Synthesis of Early Literacy Development and Implications for Intervention*. Washington, DC: National Institute for Literacy. (<http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>)
- Nelson, J. M. (2008). Beyond correlational analysis of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS): A classification validity study. *School Psychology Quarterly*, 23, 542-552.
- Nielsen, A-M. V. & Poulsen, M. (2012). Hvad kan man bruge læsescreening til – og hvornår? *Læsepædagogen*, 60(4), 4-8.
- Nielsen, H. (2000). *DLB-kendskab. Dansk lyd- og bogstavkendskab i børnehaveklassen*. Herning: Specialpædagogisk Forlag.
- Nielsen, J. C. (1998). *Tosprogede eleveres sprog og læsning på deres første- og andetprog – en undersøgelse af tyrkisksprogede elever*. København: Danmarks Pædagogiske Institut.
- Ordblindetestkonsortiet (2014). *Teknisk rapport om Ordblindetesten*. København: Center for Læseforskning, Københavns Universitet & Skoleforskningsprogrammet, IUP, Aarhus Universitet. <https://ordblindetest.nu/vejleder/pdf/TekniskRapportOrdblindetesten.pdf>.
- Ordblindetestkonsortiet (2015). *Vejledning til Ordblindetesten (version 3)*. København: Ministeriet for Børn, Undervisning og Ligestilling. <https://ordblindetest.nu/vejleder/pdf/VejledningOrdblindetestenVersion3.pdf>.
- Petersen, D. B., Allen, M. M. & Spencer, T. D. (2016). Predicting Reading Difficulty in First Grade Using Dynamic Assessment of Decoding in Early Kindergarten: A Large-Scale Longitudinal Study. *Journal of Learning Disabilities* 49(2), 200-215.
- Petersen, D. B. & Gillam, R. B. (2015). Predicting Reading Ability for Bilingual Latino Children Using Dynamic Assessment. *Journal of Learning Disabilities*, 48(1), 3-21.
- Poskiparta, E., Niemi, P., Lepola, J., Ahtola, A. & Laine, P. (2003). Motivational-emotional vulnerability and difficulties in learning to read and spell. *British Journal of Educational Psychology*, 73, 187-206.

- Poulsen, M. & Nielsen, A-M. V., Juul, H. & Elbro, C. (2015). *Identification of future reading disability before and after the onset of formal instruction*. (Artikelmanuskript under review).
- Ricketts, J., Nation, K. & Bishop, D. V. M. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, 11(3), 235-257.
- Rønberg, L. F. (2014). Hvad er problemet? Elever med dansk som andetsprog kan jo sagtens læse! *Pædagogisk Psykologisk Tidsskrift*, 51(1), 3-18.
- Simmons, D., Coyne, M., Kwok, O., McDonagh, S, Harn, B. & Kame'enui, E. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, 41, 143-158.
- Socialstyrelsen (2010). *Vejledning > Dagtilbud og skolen. Sprogvurdering af børn i treårsalderen, inden skolestart og i børnehaveklassen*. København: Kommunernes Landsforening, Undervisningsministeriet, Socialministeriet & Finansministeriet. (<http://www.kl.dk/Aktuelle-temaer/kvalitetsportalen/Faglige-kvalitetsoplysninger-om-dagtilbud/Redskaber-pa-dagtilbudsomradet/Sprogvurdering-af-3-arige-inden-skolestart-og-i-bornehaveklassen/>)
- Socialstyrelsen (2014). *Udregning af score – teknisk bilag. Støtte- og inspirationsmateriale > Dagtilbud og skolen. Sprogvurdering af børn i treårsalderen, inden skolestart og i børnehaveklassen*. København: Kommunernes Landsforening, Undervisningsministeriet, Socialministeriet & Finansministeriet. http://www.kl.dk/ImageVaultFiles/id_71173/cf_202/Sprogvurdering_-_Udregning_af_score.PDF
- Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology*, 84, 353-363.
- Sternberg, R. J. & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: University of Cambridge.
- Swalander, L. (2012). Selvbillede, motivation og dysleksi. I: S. Samuelsson m.fl. (red.). *Dysleksi og andre vanskeligheder med skriftsproget* (s. 178-191). København: Dansk Psykologisk Forlag.
- Verhoeven, L., van Leeuwe, J. & Vermeer, A. (2011). Vocabulary Growth and Reading Development across the Elementary School Years. *Scientific Studies of Reading*, 15(1), 8-25.

Bilag 1: Testoversigt

Nedenfor gives der en oversigt over de test, som deltagerne har fået i forbindelse med projektet på følgende tidspunkter:

- efteråret 0. klasse 2012
- slutningen af 0. klasse 2013
- midten af 1. klasse 2013
- slutningen af 1. klasse 2014
- slutningen af 2. klasse 2015

Oversigten er delt op efter disse fem testtidspunkter. Der indgår både publicerede testmaterialer og upublicerede materialer, der er udviklet til dette projekt eller med tilladelse lånt fra andre forskningsprojekter. De publicerede testmaterialer er i oversigten nedenunder angivet med reference til forfatter(e) og årstal og anført i litteraturlisten i denne rapport. Ved de ikke-publicerede testmaterialer er der blot anført forfatternavn(e).

Testning i efteråret 0. klasse 2012

Testning	Testmateriale	Fokusområde
Screening/ gruppetest- ning ved Center for Læse- forskning	”Store bogstaver” fra <i>Læseevaluering på begyndertrinnet</i> (Borstrøm & Klint Petersen, 2004)	Bogstavkendskab
	”Forlydsopgaver” fra <i>DLB-kendskab</i> (Nielsen, 2000)	Fonologisk opmærksomhed
	”Find billedet” fra <i>Læseevaluering på begyndertrinnet</i> (Borstrøm & Klint Petersen, 2004)	Ordforråd
Sprogvurdering ved børnehaveklasseledere	<i>Sprogvurderingsmateriale til 3-årige, inden skolestart og i børnehaveklassen</i> (Socialstyrelsen, 2010)	Produktivt talesprog Receptivt talesprog Lydlige kompetencer Kommunikative kompetencer

Individuel testning ved Center for Læseforskning	”Bogstavbenævnelse” fra <i>Læseevaluering på begyndertrinet</i> (Borstrøm & Klint Petersen, 2004)	Bogstavkendskab
	Statisk test af forlydsidentifikation (Gellert)	Fonologisk opmærksomhed
	Dynamisk test af forlydsidentifikation (Gellert)	
	Bogstav-lydforbindelse (Klint Petersen)	Bogstavkendskab
	Lyd-bogstavforbindelse (Klint Petersen)	
	Produktivt ordforråd (Gellert & Vang Christensen)	Ordforråd
	RAN: Hurtig automatisk benævnelse af genstande (Gellert)	Aut. benævneshastighed
Tilegnelse af forbindelser mellem symboler og lyde (Gellert)	Verbal indlæring	
Tilegnelse af nye ord (Gellert)	Verbal indlæring	

Testning i slutningen af 0. klasse 2013

Testning	Testmateriale	Fokusområde
Gruppetestning ved Center for Læseforskning	”Alle bogstaver” fra <i>Læseevaluering på begyndertrinet</i> (Borstrøm & Klint Petersen, 2004)	Bogstavkendskab
	”Forlydsopgaver” fra <i>DLB-kendskab</i> (Nielsen, 2000)	Fonologisk opmærksomhed
Individuel testning ved Center for Læseforskning	”Bogstavbenævnelse” fra <i>Læseevaluering på begyndertrinet</i> (Borstrøm & Klint Petersen, 2004)	Bogstavkendskab
	”Ordlæsning – 30 ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	<i>Dynamisk Afkodningstest A</i> (tidl.: Dynamisk Ordblindetest), (Daugaard, Elbro & Gellert, 2011a)	Afkodning
	Forlydsidentifikation (Gellert)	Fonologisk opmærksomhed
	Fonologisk syntese (Gellert)	Fonologisk opmærksomhed
	Bogstav-lyd-forbindelse (Klint Petersen)	Bogstavkendskab
	Lyd-bogstav-forbindelse (Klint Petersen)	
	Produktivt ordforråd (Gellert & Vang Christensen)	Ordforråd
	RAN: Hurtig automatisk benævnelse af genstande (Gellert)	Aut. benævneshastighed
Nonordsgentagelse (Vang Christensen)	Fonologisk hukommelse	

Testning i midten af 1. klasse 2013

Testning	Testmateriale	Fokusområde
Gruppe- testning ved Center for Læse- forskning	”Ordlæseprøve 1” fra <i>Skriftsproglig udvikling</i> (Juil & Møller, 2010a)	Afkodning
	”Find billedet” fra <i>Læseevaluering på begyndertrinet</i> (Borstrøm & Klint Petersen, 2004)	Ordforråd
Individuel testning ved Center for Læse- forskning	”Ordlæsning – 30 ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	”Ordlæsning – 30 nye ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	Dynamisk Afkodningstest B (Gellert)	Afkodning
	Fonologisk syntese (Gellert)	Fonologisk opmærksomhed
	”Bogstavbenævnelse” (Små bogstaver + store bogstaver) fra <i>Læse- evaluering på begyndertrinet</i> (Borstrøm & Klint Petersen, 2004)	Bogstavkendskab
	Benævnelse af store bogstaver på tid (Klint Petersen)	Bogstavkendskab
	Produktivt ordforråd (Gellert & Vang Christensen)	Ordforråd
	RAN: Hurtig automatisk benævnelse af cifre (Gellert)	Aut. benævneshastighed
	RAN: Hurtig automatisk benævnelse af genstande (Gellert)	Aut. benævneshastighed
<i>CHIPS – Children’s Problem Solving</i> (Hansen, Kreiner & Hansen, 1995)	Kognitiv udvikling	

Testning i slutningen af 1. klasse 2014

Testning	Testmateriale	Fokusområde
Gruppe- testning ved Center for Læse- forskning	”Ordlæseprøve 1” fra <i>Skriftsproglig udvikling</i> (Juul & Møller, 2010a)	Afkodning
	”Staveprøve 1” fra <i>Skriftsproglig udvikling</i> (Juul, 2013)	Stavning
Individuel testning ved Center for Læse- forskning	”Ordlæsning – 30 ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	”Ordlæsning – 30 nye ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	Gentagen højtlesning af tre ordlister med hjælp til stavelsesdeling (Gellert)	Afkodning
	”Dinodiktat” fra <i>Læseevaluering på begyndertrinnet</i> (Borstrøm & Klint Petersen, 2004)	Stavning
	Benævnelse af store og små bogstaver på tid (Gellert)	Bogstavkendskab
	Produktivt ordforråd (Gellert & Vang Christensen)	Ordforråd
	RAN: Hurtig automatisk benævnelse af cifre (Gellert)	Aut. benævneshastighed
RAN: Hurtig automatisk benævnelse af genstande (Gellert)	Aut. benævneshastighed	

Testning i slutningen af 2. klasse 2015

Testning	Testmateriale	Fokusområde
Gruppe- testning ved Center for Læse- forskning	”Ordlæseprøve 2” fra <i>Skriftsproglig udvikling</i> (Juul & Møller, 2010b)	Afkodning
	”Nonordsstavning” fra <i>Udvikling og afprøvning af procedure til identifikation af elever i risiko for dysleksi</i> (Arnbak & Borstrøm, 2007)	Stavning
	”Tekstlæseprøve 4” fra <i>Skriftsproglig udvikling</i> (Møller, 2012)	Tekstforståelse
Individuel testning ved Center for Læse- forskning	<i>Ordblindetesten</i> (Undervisningsministeriet, 2015) <ul style="list-style-type: none"> • Spørgeskema • ”Find stavemåden” • ”Find det der lyder som et ord” • ”Basal ordforrådsprøve” 	Baggrundoplysninger Afkodning Afkodning Ordforråd
	<i>Udvikling og afprøvning af procedure til identifikation af elever i risiko for dysleksi</i> (Arnbak & Borstrøm, 2007) <ul style="list-style-type: none"> • Fonologisk subtraktion • Højtlesning af ord • Højtlesning af nonord 	Fonologisk opmærksomhed Afkodning Afkodning
	”Ordlæsning – 30 ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	”Ordlæsning – 30 nye ord” fra <i>Læsevanskeligheder</i> (Elbro, 2007)	Afkodning
	Højtlesning af tre ordlister (Gellert)	Afkodning
	Produktivt ordforråd (Gellert & Vang Christensen)	Ordforråd

Bilag 2: Forudsigelse af afkodningsvanskeligheder i 1. klasse på basis af testresultater i 0. klasse

Forudsigelse af afkodningsvanskeligheder i 1. klasse på basis af testresultater i efteråret 0. klasse

Vi har i den øvrige del af rapporten fokuseret på at undersøge, hvorvidt man kunne forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse, da elevernes læseudvikling på dette tidspunkt er nået længst og giver det bedste grundlag for at vurdere, om eleverne udviser tegn på vedvarende afkodningsvanskeligheder. Vi har imidlertid også undersøgt, om en dynamisk test af fonologisk opmærksomhed i efteråret 0. klasse kunne bidrage til at forudsige *tidlige* afkodningsvanskeligheder, dvs. afkodningsvanskeligheder i midten og slutningen af 1. klasse – i stedet for (som rapporteret i afsnit 5.4) afkodningsvanskeligheder i slutningen af 2. klasse. I både midten og slutningen af 1. klasse blev eleverne kategoriseret ud fra deres afkodningspræcision ved højtlesning af Elbros ordlister med ord og nonord. Vi fandt, at den dynamiske test af fonologisk opmærksomhed i efteråret 0. klasse faktisk godt kunne bidrage selvstændigt til at forudsige, hvilke elever der i *midten af 1. klasse* (altså godt et år senere) udviste tidlige tegn på afkodningsvanskeligheder målt med Elbros ordlister med højtlesning af ord og nonord, selv efter at der var taget højde for elevernes resultater på traditionelle test af bogstavkendskab og fonologisk opmærksomhed i efteråret 0. klasse. Imidlertid kunne den dynamiske test af fonologisk opmærksomhed *ikke* bidrage signifikant til at forudsige afkodningsvanskeligheder i *slutningen af 1. klasse*, når vi først havde taget højde for elevernes resultater på traditionelle test i efteråret 0. klasse. For en nærmere redegørelse for disse resultater henvises der til Gellert & Elbro (2015a).

Sammenfattende viste resultaterne fra vores langtidsundersøgelse, at den dynamiske test af fonologisk opmærksomhed *på kort sigt* godt kunne bidrage til at forudsige elevens vanskeligheder i forbindelse med den tidlige læseundervisning (frem til midten af 1. klasse i vores undersøgelse), men at denne dynamiske test ikke overbevisende kunne bidrage selvstændigt til *på længere sigt* at forudsige afkodningsvanskeligheder (i slutningen af hhv. 1. og 2. klasse).

Disse resultater af den aktuelle langtidsundersøgelse er i overensstemmelse med udenlandske undersøgelser, som har vist, at inddragelse af dynamiske test af fonologisk opmærksomhed i begyndelsen af børnehaveklassen kan give bedre muligheder for at identificere de børn, der omkring et halvt år senere viser tegn på vanskeligheder i den tidlige fase i læseudviklingen (fx Bridges & Catts, 2011; Spector, 1992). Samtidig er resultaterne af vores undersøgelse også i overensstemmelse med en ny amerikansk langtidsundersøgelse, der har fulgt elevernes læseudvikling længere frem og vist, at en dynamisk test af fonologisk opmærksomhed i begyndelsen af børnehaveklassen *ikke* kunne bidrage selvstændigt til at forudsige, hvilke elever der i slutningen af 1. klasse ville udvise tegn på afkodningsvanskeligheder, når man først havde taget højde for elevernes

tidlige resultater på traditionelle test af bl.a. fonologisk opmærksomhed og bogstavkendskab (Catts m.fl., 2015).

Forudsigelse af afkodningsvanskeligheder i 1. klasse på basis af testresultater i slutningen af 0. klasse

Vi har i den øvrige del af rapporten fokuseret på at forudsige alvorlige afkodningsvanskeligheder i slutningen af 2. klasse. Vi har imidlertid også undersøgt, om Dynamisk Afkodningstest A i slutningen af 0. klasse kunne bidrage til at forudsige *tidlige* afkodningsvanskeligheder, dvs. afkodningsvanskeligheder *i midten og slutningen af 1. klasse* – i stedet for (som rapporteret i afsnit 5.5) afkodningsvanskeligheder i slutningen af 2. klasse. I både midten og slutningen af 1. klasse blev eleverne kategoriseret ud fra deres afkodningspræcision ved højtlesning af Elbros ordlister med ord og nonord. Vi fandt, at den dynamiske afkodningstest i slutningen af 0. klasse var den stærkeste enkelte prædiktor af elevernes afkodningsfærdigheder i midten og slutningen af 1. klasse (lige som den som beskrevet i afsnit 5.5 var den stærkeste prædiktor af elevernes afkodningsfærdigheder i slutningen af 2. klasse). Ligeledes kunne denne dynamiske afkodningstest også bidrage selvstændigt til at forudsige, hvilke elever der i både midten og slutningen af 1. klasse udviste tegn på afkodningsvanskeligheder, selv efter at der var taget højde for elevernes resultater på traditionelle test i slutningen af 0. klasse. (For en nærmere redegørelse for disse resultater henvises der til Gellert & Elbro, 2014; Gellert & Elbro, 2015b).

Endvidere har vi undersøgt, om det også styrker forudsigelsen af *tidlige* afkodningsvanskeligheder (dvs. afkodningsvanskeligheder *i midten og slutningen af 1. klasse*), hvis man foruden Dynamisk Afkodningstest A i slutningen af 0. klasse inddrager traditionelle test. Vi fandt, at forudsigelsen af afkodningsvanskeligheder *i midten af 1. klasse* blev mere præcis, hvis man foruden den dynamiske afkodningstest også inddrog test af bogstavkendskab og afkodningspræcision. Forudsigelsen af afkodningsvanskeligheder *i slutningen af 1. klasse* blev mere præcis, hvis man foruden den dynamiske afkodningstest også inkluderede test af bogstavkendskab og fonologisk opmærksomhed. (For en nærmere redegørelse for disse resultater henvises der igen til Gellert & Elbro, 2014; Gellert & Elbro, 2015b).

Sammenfattende viste den aktuelle langtidsundersøgelse, at den dynamiske afkodningstest i slutningen af 0. klasse var den stærkeste enkelte prædiktor af både tidlige afkodningsfærdigheder (i midten og slutningen af 1. klasse) og senere afkodningsfærdigheder (i slutningen af 2. klasse). Selv efter skrap kontrol med alle traditionelle prædiktorer bidrog denne test selvstændigt til at forudsige afkodningsvanskeligheder både i midten og slutningen af 1. klasse samt i slutningen af 2. klasse. Endvidere viste langtidsundersøgelsen, at selv om den dynamiske afkodningstest i slutningen af 0. klasse var den stærkeste enkelte prædiktor af både tidlige afkodningsfærdigheder (i midten og slutningen af 1. klasse) og senere afkodningsfærdigheder (i slut-

ningen af 2. klasse), kunne denne dynamiske test *ikke* helt erstatte de traditionelle test. Især testen af bogstavkendskab bidrog yderligere til forudsigelse af senere afkodningsvanskeligheder.

Bilag 3: Præcision i tidlig identifikation af risikoelever med kategorisering i 2. klasse på basis af Elbros ordlister og DVO-ordblindetesten ved forskellige niveauer af sensitivitet

De to nedenstående tabeller (27 og 28) viser med elevernes afkodningsstatus på basis af deres resultater på hhv. Elbros ordlister og DVO-ordblindetesten i slutningen af 2. klasse som afhængige variable for hvert af de tre tidlige testtidspunkter (hhv. slutningen af 0. klasse, midten af 1. klasse og slutningen af 1. klasse) ved forskellige niveauer af sensitivitet (hhv. 60, 70, 80 og 90 %), hvor stor en andel af falsk positive (falske alarmer) der følger med. Tabellerne er parallelle med tabel 22 med UVM-ordblindetesten som afhængig variabel og skal ses som supplement til denne tabel (jf. afsnit 6.1).

Forudsigelse af: Elbros ordlister med ord og nonord (afkodningseffektivitet slut 2. kl.)			Sensitivitet (raten af sandt positive)			
			,60	,70	,80	,90
Prædiktions-tidspunkt	Prædiktorer	AUC	Andel falsk positive	Andel falsk positive	Andel falsk positive	Andel falsk positive
Slut 0. kl.	Dynamisk afkodning + bogstavkendskab	,88	,06	,12	,33	,38
Midt 1. kl.	Afkodningspræcision + bogstavkendskab	,95	,01	,06	,06	,18
Slut 1. kl.	Afkodningseffektivitet	,97	,01	,05	,08	,12

Tabel 27. Præcision i tidlig identifikation af elever i risiko for alvorlige afkodningsvanskeligheder ved forskellige niveauer af sensitivitet og kategorisering på basis af Elbros ordlister (afkodningseffektivitet) i slutningen af 2. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).

Forudsigelse af: DVO-ordblindetest (afkodningspræcision + fon. opm. slut 2. kl.)			Sensitivitet (raten af sandt positive)			
			,60	,70	,80	,90
Prædiktions-tidspunkt	Prædiktorer	AUC	Andel falsk positive	Andel falsk positive	Andel falsk positive	Andel falsk positive
Slut 0. kl.	Dynamisk afkodning + bogstavkendskab	,84	,14	,22	,32	,34
Midt 1. kl.	Afkodningspræcision + bogstavkendskab	,86	,06	,17	,27	,44
Slut 1. kl.	Afkodningseffektivitet	,89	,10	,11	,25	,33

Tabel 28. Præcision i tidlig identifikation af elever i risiko for alvorlige afkodningsvanskeligheder ved forskellige niveauer af sensitivitet og kategorisering på basis af DVO-ordblindetest (afkodningspræcision og fonologisk opmærksomhed) i slutningen af 2. klasse.

Beregningerne er baseret på det fulde sample ($N = 158$).