

Hvordan hænger det sammen?

Om specifikke vanskeligheder med sprogforståelse i læsning
og inferensfærdigheders betydning for læseforståelse

Ph.d.-afhandling af Hanne Trebbien Daugaard

Institut for Nordiske Studier og Sprogvidenskab

Københavns Universitet

Maj 2015

Vejleder: Carsten Elbro

Indhold

Abstract	5
Resumé	6
Forord	7
Introduktion	8
Kapitel 1. Specifikke vanskeligheder med sprogforståelse i læsning	11
Dårlige forståere	11
Den enkle læsemodel	11
Tre slags læsevanskeligheder	12
Mulige årsager til dårlig sprogforståelse i læsning	14
Ordforråd og ordkendskab	15
Grammatiske færdigheder	16
Viden om teksters struktur	18
Monitorering	19
Inferensfærdigheder	20
Arbejdshukommelse	21
Sammenfatning	21
Kapitel 2. Afgrænsning af gruppen af dårlige forståere	23
Indledende valg i afgrænsningen	24
Afgrænsning fra udviklingsforstyrrelser og andre kendte vanskeligheder	24
Afgrænsning fra læsning på et andet- eller fremmedsprog	25
Kontrol med udvalgte færdigheder	25
Afgrænsning i klinisk praksis	26
Undersøgelse af nøglefærdighederne	26
Undersøgelse af læseforståelse	26
Undersøgelse af afkodningsfærdigheder	27
Absolut eller relativ dårlig læseforståelse	30
Konsekvenser af forskellige grænsedragninger	30
Grænsedragning i overensstemmelse med den generelle definition af dårlige forståere	33
Sammenfatning	34
Kapitel 3: Stabilitet af profilen 'dårlig forstår' på tværs af tid og testbetingelser	36
Indledning	36
Stabilitet over tid	37
Stabilitet på tværs af test	40
Opsamling	47
Formål	47
Metode	48
Design	48

Deltagere	48
Materialer	49
Fremgangsmåde	51
Resultater	52
Udvælgelse af dårlige og gode forståere ved første testning	52
Stabiliteten af 'dårlig forstår'-profilen.....	55
Kilder til ustabilitet	57
Diskussion.....	59
Konklusion	63
Kapitel 4: Hvilke aspekter af inferensfærdigheder er vigtige for læseforståelse?	65
Indledning.....	65
Modalitet	69
Videnstype	72
Sproglige signaler om inferens i teksten	76
Formål.....	78
Metode.....	79
Design	79
Deltagere	79
Materialer	79
Fremgangsmåde	81
Resultater	81
Modalitet	85
Videnstype	86
Sproglig signalering.....	87
Diskussion.....	88
Konklusion	92
Kapitel 5: Samspillet mellem inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse	94
Indledning.....	94
Ordforråd og læseforståelse.....	94
Arbejdshukommelse og læseforståelse	96
Inferensfærdigheder og ordforråd i læseforståelse	97
Inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse	100
Opsamling.....	102
Formål.....	102
Metode.....	103
Design	103
Deltagere	103
Materialer	103
Fremgangsmåde	104
Resultater	105
Diskussion.....	108

Konklusion	112
Kapitel 6: Er detaljefokus en hæmsko for inferensfærdigheder?	113
Indledning.....	113
Forventning om sammenhæng	113
Tendens til at fokusere på detaljer.....	114
The Embedded Figures Test	116
Formål.....	117
Metode	118
Design	118
Deltagere	118
Materialer	118
Fremgangsmåde	119
Resultater	120
Diskussion.....	122
Konklusion	125
Kapitel 7: Sammenfatning med perspektiver	126
Hvem er dårlige forståere?	126
Inferensfærdighedens betydning for læseforståelse.....	128
Årsager til dårlig forståelse.....	132
Forskningsperspektiver	134
Undervisningsperspektiver.....	135
Referencer	137
Bilag.....	150

Abstract

This dissertation is concerned with students with specific reading comprehension difficulties (poor comprehenders) and the ability to draw inferences while reading (inference making). The dissertation encompasses empirical studies based on test results from students in the 6th grade.

In the dissertation poor comprehenders are defined as students with poor text comprehension which cannot be attributed to poor word reading abilities (decoding). Current research into possible reasons for poor reading comprehension is reviewed. Possible reasons are components of language comprehension: vocabulary and word knowledge, grammatical skills, knowledge of text structure, comprehension monitoring, and inference making, as well as the general ability to temporarily store and manipulate information in short-term memory (working memory). The review shows strong evidence for regarding problems with inference making as a reason for poor text comprehension.

The practical delimitation of poor comprehenders involves a number of choices and presents some challenges. These choices and challenges are discussed by reference to current research, and a method for delimitation is proposed. The proposed method entails both cut-off values for results on tests of decoding and reading comprehension *and* establishment of a discrepancy between decoding level and comprehension level.

Test results from 211 students provide a basis for the delimitation of 21 poor comprehenders and 26 good comprehenders and the stability of comprehension difficulties is investigated across a period of 4-5 months and across a group test and an individually administered test of reading comprehension (parallel as to content). The stability is moderate. Differences in conditions associated to group tests and individually-administered tests are likely to contribute to instability.

Inference making abilities are examined with data from 54 students at different decoding- and comprehension levels. Results suggest the ability of inference making is generalised (across visual modality and language), but results also suggest that only language-based inference making is important for text comprehension.

The type of knowledge (academic knowledge or knowledge based on common experience) which must be activated during inference making seems to be of significance for differences in reading comprehension. It seems the ability to activate academic knowledge goes along with good general text comprehension.

Some inferences are signalled in texts by means of noun references or connectives. The results show a correlation between ability to make signalled inferences and non-signalled inferences, but results also indicate that each of these abilities is important for good reading comprehension.

Inference making contributes to reading comprehension independent of vocabulary and working memory. Results point, moreover, to an interplay between vocabulary and inference making. Inference making carries the effect of vocabulary on comprehension. Working memory resources do not seem to be critical to this relationship.

Whether a tendency to focus on details (also called ‘weak central coherence’) is contrary to good inference making skills is tested. Tendency to focus on details is examined with *the Embedded Figures Test*. Results do *not* suggest that a tendency to focus on details is contrary to good inference making skills. As a matter of fact, there is a positive correlation between the two. However, this correlation can be explained by variation in other skills such as nonverbal problem-solving and vocabulary.

The results are summarised in a final discussion and perspectives for future research and intervention are outlined.

Resumé

Denne afhandling handler om elever med specifikke vanskeligheder med sprogforståelse i læsning (dårlige forståere) og om evnen til at drage slutninger, mens man læser (inferensfærdigheder). Afhandlingen fremlægger empiriske undersøgelser med testresultater fra elever på 6. klassetrin.

I afhandlingen defineres dårlige forståere som elever med dårlig tekstforståelse, der ikke skyldes problemer med at afkode de enkelte ord i teksten. Eksisterende forskning i mulige årsager til dårlig sprogforståelse i læsning gennemgås. Det er komponenter i sprogforståelse: Ordforråd og ordkendskab, grammatiske færdigheder, viden om teksters struktur, monitorering af egen forståelse og inferensfærdigheder. Og det er en generel evne til at holde information i korttidshukommelsen, samtidig med at den bearbejdes (arbejdshukommelse). Gennemgangen viser bl.a., at der er et stærkt belæg for at opfatte problemer med inferensfærdigheder som årsag til dårlig tekstforståelse.

Praktisk afgrænsning af dårlige forståere er forbundet med en række valg og udfordringer. Disse diskuteres med udgangspunkt i eksisterende forskning, og der gives bl.a. et forslag til konkret grænsedragning mellem dårlige forståere og andre læsere på baggrund af resultater fra afkodnings- og læseforståelsestest. Grænsedragningen indebærer simple grænseværdier for henholdsvis læseforståelse og afkodning og en sikring af et misforhold (en diskrepans) i niveau i de to færdigheder for dårlige forståere.

Testresultater fra 211 elever danner dernæst grundlag for en afgrænsning af 21 dårlige forståere og 26 gode forståere, og stabiliteten af forståelsesvanskelighederne undersøges over en periode på 4-5 måneder og på tværs af en gruppetest og en individuelt afviklet test af læseforståelse med parallelt indhold. Stabiliteten er moderat. Forskelle i betingelserne for gruppetest og individuelle test er sandsynligvis medvirkende til ustabilitet.

Inferensfærdigheder undersøges med data fra 54 elever på forskellige afkodnings- og forståelsesniveauer. Resultaterne fra den aktuelle undersøgelse tyder på, at der findes en tvær-modal inferensfærdighed (på tværs af visuel og sproglig modalitet), men resultaterne tyder *også* på, at det kun er inferensdragning i sproglig modalitet, der har betydning for læseres forståelse af tekster.

Den type af viden (emnespecifik viden eller almenmenneskelige erfaringer), der skal aktiveres ved inferenser, ser ud til at kunne have en betydning for forskelle i læseforståelse. Evne til at mobilisere emnespecifik viden synes at følges med generel god tekstforståelse.

Nogle inferenser er indikeret i teksten ved sproglige signaler som henvisende substantiver med bestemthedsendelse og konnektiver. Undersøgelsen viser, at der er sammenhæng mellem læseres færdigheder i at drage inferenser med og uden disse signaler, men tyder også på, at færdighederne hver især betyder noget for god læseforståelse.

Inferensfærdigheder har en selvstændig betydning for læseforståelse, når man kontrollerer for betydningen af ordforråd og arbejdshukommelse for læseforståelse. Resultaterne tyder desuden på, at ordforråd og inferensfærdigheder spiller sammen i læseforståelse. Betydningen af ordforråd for læseforståelse synes at gå igennem inferensfærdigheder. Arbejdshukommelsesressourcer synes ikke at være afgørende i denne sammenhæng.

Det afprøves, om en tendens til at fokusere på detaljer (eller 'svag central kohærens') er i modstrid med gode inferensfærdigheder. Detaljefokus testes med *the Embedded Figures Test*. Resultaterne tyder *ikke* på, at et detaljefokus er i modstrid med gode inferensfærdigheder. Der synes tværtimod at være en positiv sammenhæng mellem detaljefokus og inferensfærdigheder, der dog kan forklares med forskelle i andre færdigheder som evne til problemløsning og ordforråd.

Resultaterne sammenfattes og diskuteres, og perspektiver for fremtidig forskning og undervisning skitseres.

Forord

Ph.d.-projektet om forståelsesvanskeligheder og inferensfærdigheder i læseforståelse er en del af det kollektive forskningsprojekt *Language and Cognition – Perspectives from Impairment* (LaCPI) og er finansieret af Det Frie Forskningsråd og Det Humanistiske Fakultet, Københavns Universitet.

En række personer har haft betydning for projektforsløbet og for afhandlingens færdiggørelse, og jeg vil gerne benytte lejligheden til at takke dem.

En stor tak til Elisabeth Engberg-Pedersen og Rikke Vang Christensen for at skabe rum til mit projekt i LaCPI-projektet. Jeg er også taknemmelig for de mange studentermedhjælperes omhyggelige arbejde med at indsamle og indtaste data og ikke mindst for de deltagende elevers velvilje til at gennemføre de mange opgaver og tålmodigt svare på mine spørgsmål. Også en stor tak til elevernes forældre og lærere for at gøre testningen mulig.

Som led i mit stipendium var jeg på et forskningsophold ved Lancaster Universitet, som har haft stor betydning for mig og for mit arbejde med afhandlingen. Tak til Kate Cain for inspiration og beredvillighed til at dele sin store viden og mange idéer. Tak til Nicola Currie, Emma James, Liza van den Bosch, Liam Blything og Robin Gillmour.

Tak til Jane Oakhill og Courtenay Norbury for gode råd og opmuntring undervejs i processen.

Tak til Anne-Mette Veber Nielsen og Holger Juul for sparring og kommentarer til dele af afhandlingen. Og tak for lån af skulder, Anne-Mette.

Jeg har mødt mange fantastiske ph.d.-studerende på INSS og fakultetet, siden jeg startede på mit projekt. Tak for hjælp og opbakning til Susanne Kemp, Krista Stinne Rasmussen, Randi Skovbjerg Sørensen, Philip Thomas Lavender og Niels Grotum Sørensen.

Jeg skylder også min familie og Sebastian Daugaard en stor tak.

Tak til venner og til Per.

Sidst, men ikke mindst, tak for konstruktiv vejledning til min vejleder Carsten Elbro.

Amager, d. 29. maj 2015

Hanne Trebbien Daugaard

Introduktion

Vi læser tekster hver dag og hele tiden, også når vi ikke tænker over det. Inden vi er gået ud af døren om morgenen, har vi allerede læst en masse tekst. Vi har læst teksten bag på havregrynspakken og etiketten på hårshampooen; vi har læst i vores kalendere; og vi har læst glemte regninger og mødeindkaldelser. Vi har læst det sure opslag i opgangen (for 117. gang). Vi har læst diverse sms-beskeder og opdateringer på de sociale medier – og vi er ikke engang nået til indbakken eller dagens nyheder endnu. Når vi kommunikerer med andre, med offentlige institutioner, med rejseselskabet, med renovationsfirmaet, med banken, med arkitekten og med advokaten, sker det – før eller siden – skriftligt. Næsten alle uddannelser og jobs kræver læse- og skrivefærdigheder. Den enkeltes handlemuligheder indskrænkes eller vanskeliggøres altså, hvis ikke han eller hun kan læse og skrive på et vist niveau. Læsevanskeligheder kan derfor have store konsekvenser for den enkelte.

Nogle læsevanskeligheder er velundersøgt og relativt veldefinerede. Det gælder fx ordblindhed (dysleksi). Der er bred enighed om, at ordblindhed er en specifik indlæringsvanskelighed, og at ordblindes kernevanskelighed i læsning er afkodningen af de enkelte ord: Den ordblinde har svært ved at forbinde de enkelte bogstaver og bogstavfølger med deres rette sproglyd (Lyon, Shaywitz & Shaywitz, 2003; Elbro, 2007). Har man ordblinddevanskeligheder, er man berettiget til udredning, undervisning og hjælpemidler. Forskningen har identificeret vigtige forudsætninger for afkodningsfærdigheder, og mange effektundersøgelser har dokumenteret, hvordan nogle undervisningsmetoder virker bedre end andre (National Reading Panel, 2000) – og dette arbejde fortsætter.

Forståelsvanskeligheder i læsning er endnu ikke så grundigt undersøgt, og der er i praksis ikke tradition for at undersøge elevers sprogforståelse i læsning tidligt og sætte ind med målrettet intervention ved vanskeligheder (Leach, Scarborough & Rescorla, 2003; Cain & Oakhill, 2004). Det er ikke desto mindre dokumenteret, at forståelsvanskeligheder kan have uddannelsesmæssige konsekvenser (Ricketts, Sperring & Nation, 2014; Cain & Oakhill, 2006), og forskning i sprogforståelse og forståelsvanskeligheder i læsning har afdækket en række vigtige komponenter i sprogforståelse, der har betydning for tekstforståelse. Der er ligeledes effektundersøgelser, der kan danne grundlag for anvisninger til, hvordan man kan undervise i sprogforståelse i læsning med succes (se fx Oakhill, Cain, & Elbro, 2015).

Første del af denne afhandling handler om elever med specifikke vanskeligheder med sprogforståelse i læsning (dårlige forståere). Hvordan kan dårlige forståere defineres i forhold til andre læsere med læsevanskeligheder? Hvilke mulige årsager er der til dårlig sprogforståelse i læsning? Hvordan kan man konkret afgrænse dårlige forståere? Hvilke udfordringer er der, når man følger disse elever med specifikke vanskeligheder med sprogforståelse i læsning? Kapitel 1 giver en introduktion til forskellige typer af læsevanskeligheder og en generel definition af dårlige forståere. Dernæst følger en introduktion til de forskellige færdigheder, som kan volde problemer for dårlige forståere, og til det belæg, der er for at opfatte disse problemer som årsager til elevernes dårlige læseforståelse. I kapitel 2 gennemgår og diskuterer jeg udvalgte dimensioner i den praktiske afgrænsning af specifikke vanskeligheder med sprogforståelse i læsning med udgangspunkt i

fremgangsmåder i forskningslitteraturen. Jeg giver et forslag til en fremgangsmåde, der kan have relevans for praksis, og et konkret forslag til grænsedragning, der fører til en afgrænsning, der er i overensstemmelse med den generelle definition af dårlige forståere. Kapitel 3 handler om, hvilken stabilitet af specifikke vanskeligheder med sprogforståelse i læsning man kan forvente over tid og på tværs af forskellige læseforståelsestest og hvilke faktorer, der kan forventes at påvirke stabiliteten. Kapitlet fremlægger en empirisk undersøgelse af stabiliteten for en gruppe 6.-klasseelever med forståelsesvanskeligheder over en kort periode og på tværs af to indholdsparallelle test afviklet under forskellige testbetingelser.

Der kan næppe herske tvivl om, at tekstforståelse må være en kompleks kognitiv aktivitet. Fælles for teorier om tekstforståelse er, at de antager, at læseren konstruerer en mental repræsentation af tekstens forestillingsindhold, mens – og eventuelt efter – at han eller hun læser teksten. Én teori om tekstforståelse (Kintsch, 1998) antager, at der er repræsentationer på flere niveauer i spil; læseren konstruerer både en repræsentation baseret udelukkende på tekstens eksplicite informationer og en repræsentation, der også indeholder implicit udtrykt information og baggrundsviden. Kintsch kalder den sidstnævnte repræsentation for en *situationsmodel*. En anden teori om tekstforståelse (Johnson-Laird, 1989) fremhæver, at læsere netop ikke husker tekstens informationer ordret, og opererer derfor *ikke* med en antagelse om en grundlæggende repræsentation, der består af tekstens eksplicite informationer. Johnson-Laird kalder læserens mentale repræsentation af teksten (inkl. implicitte informationer og baggrundsviden) for en *mental model*.

Et kardinalpunkt i læserens konstruktion af en sammenhængende og meningsfuld mental repræsentation af tekstens indhold er, at læseren drager inferenser – han eller hun må knytte forbindelser på tværs af tekstens sætninger og bidrage til teksten med information, for at teksten hænger sammen og giver mening. Dette behov kan følgende tekst give et indblik i:

Vi lå i Aden i en uges tid. Vi kunne ikke få lov til at sejle, fordi der var problemer med noget last, som skulle losses. Jeg kunne ikke få fat på rederiet. Det skulle foregå via telegram, men vi måtte ikke bruge radiosenderen. Vi tændte for den en enkelt gang, og med det samme satte to små krigsskibe kurs mod os, så vi slukkede igen [...].

Det er her afgørende for forståelsen af teksten, at læseren fx infererer, hvor handlingen finder sted. Ord i teksten som *sejle*, *last*, *losses* og *rederiet* tyder på, at historien foregår på et containerskib. For at få teksten til at hænge sammen må læseren også inferere, *hvad* det er, der skal foregå via telegram, og *hvad* det er, som personerne i teksten tændte for en enkelt gang. Den hidtidige forskning har kun i begrænset omfang belyst, hvilke aspekter af inferensfærdigheder, der er særligt vigtige for læsers tekstforståelse. Er det selve det at kunne drage en følgeslutning, der er afgørende? Har den type af viden, det er nødvendigt at aktivere for at drage inferensen, en betydning? Er det springende punkt at kunne udnytte sproglige signaler om inferens i teksten? Det er heller ikke undersøgt grundigt, hvordan inferensfærdigheder spiller sammen med andre komponenter i læseforståelse, selv om der er grunde til at forvente, at inferensfærdigheder indgår i et komplekst samspil med andre færdigheder. Hvordan spiller inferensfærdigheder fx sammen med ordforråd og arbejdshukommelse i læseforståelse? Og er et fokus på detaljer i modstrid med gode inferensfærdigheder?

Anden del af afhandlingen drejer sig om betydningen af forskellige aspekter af inferensfærdigheder for læseforståelse og om relationen mellem inferensfærdigheder og andre færdigheder i læseforståelse. Kapitel 4, 5 og 6 fremlægger empiriske undersøgelser af inferensfærdigheder hos et bredt udvalg af elever fra 6. klassetrin med jævn spredning i læseforståelse. I kapitel 4 undersøger jeg betydningen af forskellige aspekter af inferensfærdigheder for læseforståelse. Det første aspekt er modalitet. Her undersøger jeg betydningen af at kunne drage inferenser fra tekst (sprogligt materiale) over for betydningen af at kunne drage inferenser fra billedhistorier (visuelt materiale). Det næste aspekt er, hvilken type af viden det er nødvendigt at aktivere for at drage en inferens. Her undersøger jeg betydningen af at kunne aktivere emnespecifik viden over for viden om almenmenneskelige forhold. Det tredje aspekt er tilstedeværelsen af et sprogligt signal om inferens i teksten. Her undersøger jeg betydningen af at kunne drage inferenser, der er signaleret i teksten, og inferenser, der ikke er signaleret i teksten. Kapitel 5 handler om relationen mellem inferensfærdigheder og andre færdigheder i læseforståelse. Jeg undersøger først betydningen af inferensfærdigheder for læseforståelse i forhold til betydningen af ordforråd og arbejdshukommelse. Dernæst undersøger jeg relationen mellem inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse. I kapitel 6 undersøger jeg, om en tendens til at fokusere på detaljer kan have betydning for læseres inferensfærdigheder.

Afhandlingen afrundes med en sammenfatning og diskussion af resultaterne (kapitel 7), hvor jeg også fremhæver perspektiver, som jeg anser for at være centrale for fremtidige undersøgelser og for identifikation og undervisning af dårlige forståere i praksis.

I de efterfølgende bilag findes eksempler på opgaver fra de eksperimentelle testmaterialer, der er blevet udviklet til undersøgelserne i afhandlingen. Til brug for en nærmere granskning af afhandlingen findes desuden i tillæg til afhandlingen en selvstændig bilagssamling, der indeholder de eksperimentelle testmaterialer i deres fulde længde og oversættelser af de engelske materialer, der blev brugt til at undersøge henholdsvis læseforståelse og detaljefokus. Den selvstændige bilagssamling opbevares hos mig.

Det skal nævnes, at jeg bruger 'læseforståelse' synonymt med 'tekstforståelse', hvis ikke andet er anført.

Kapitel 1. Specifikke vanskeligheder med sprogforståelse i læsning

Dårlige forståere

Den enkle læsemodel

God tekstforståelse beror på gode færdigheder i både *afkodning* og *sprogforståelse* (Hoover & Gough, 1990; Gough, Hoover, Peterson, 1996). Afkodning er alle de processer, der fører til identifikation af skrevne ord. Det kan altså være identifikation via de enkelte bogstavers lyd, via analogier med kendte ord, eller ved direkte genkendelse af den unikke bogstavfølge ("ordbilledet") (Ehri, 2005). Sprogforståelse er alle de processer, der leder fra de afkodede (eller hørte) ord frem til et sammenhængende forestillingsindhold. Det er processer som aktivering af ordenes betydning, analyse af sætningsopbygning, integrering af de enkelte sætningers betydning og inferensdragning (tankeslutninger) på tværs af de enkelte ord og sætninger (Perfetti, Landi & Oakhill, 2005). Selv om der er visse forskelle på talt og skrevet sprog og dermed på forståelsens vilkår i de to sproglige modaliteter, indebærer forståelse af talt og skrevet sprog grundlæggende de samme sprogforståelsesprocesser. Der er tale om ét og samme sprog, uanset om det er talt eller skrevet, og læse- og lytteforståelse er nært forbundne, når man først er blevet en god afkoder (Gough et al., 1996).

Den enkle læsemodel (Læseforståelse = afkodning x sprogforståelse) bygger på mange slags iagttagelser, som viser uafhængighed mellem de to hovedkomponenter i læsning (Gough et al., 1996; Oakhill et al., 2015). For det første er der dobbeltdissociation for læsevanskeligheder. Det vil sige, at der for begge komponenters vedkommende findes personer, der har vanskeligheder med denne ene komponent, men ikke med den anden komponent (fx Catts, Adlof & Weismer, 2006). Vanskeligheder med afkodning eller sprogforståelse kan altså eksistere uafhængigt af færdighedsniveauet i den anden komponent. Se også de følgende afsnit herunder om de tre slags læsevanskeligheder, der kan udledes fra den enkle læsemodel.

For det andet er der forskellige sproglige færdigheder, der forudsiger afkodning og sprogforståelse over tid. Der knytter sig altså et relativt selvstændigt sæt af prædiktorer til hver af de to komponenter. For afkodning er de mest betydningsfulde prædiktorer bogstavkendskab og fonologisk opmærksomhed. Blandt kendte tidlige prædiktorer for forståelse er ordforråd og kendskab til teksters struktur (Muter, Hulme, Snowling & Stevenson, 2004; Oakhill & Cain, 2012).

For det tredje kan man iagttage, at den relative betydning af de to komponenter for læseforståelse generelt ændres over tid i populationen af læsere. Tidligt i læseudviklingen er der meget stor forskel på elevernes afkodningsfærdigheder, og en stor del af variationen i tekstforståelse forklares derfor af variationen i afkodningsfærdigheder. Senere i læseudviklingen forklarer sprogforståelse gradvist

en større del af variationen i læseforståelse, efterhånden som læserne automatiserer deres afkodningsfærdigheder, og teksterne bliver indholdsmæssigt mere krævende (Gough et al., 1996; Chen & Vellutino, 1997).

For det fjerde er afkodning og sprogforståelse i *forskellig* grad associeret med andre kognitive færdigheder og viden. Forståelse er fx i høj grad korreleret med baggrundsviden, det samme gælder ikke afkodning (Gough et al., 1996). Forståelse og komponenter i sprogforståelse er desuden ofte stærkere korreleret med generelle IQ-mål, end afkodningsfærdigheder er (fx Cutting & Scarborough, 1997; Oakhill & Cain, 2012; Rowe, Miller, Ebenstein & Thompson, 2012).

Tre slags læsevanskeligheder

Ud fra den enkle læsemodel kan det forudsiges, at der må være tre slags læsevanskeligheder (Catts et al., 2006): Specifikke vanskeligheder med afkodning (fx ordblindhed), specifikke vanskeligheder med sprogforståelse i læsning ("dårlig sprogforståelse") og vanskeligheder med begge komponenter ("sammensatte vanskeligheder").

Den enkle læsemodel kan således lede til en generel definition af specifikke vanskeligheder med sprogforståelse i læsning: 'Vanskeligheder med tekstforståelse, som ikke skyldes dårlig afkodning'.

Ordblindhed

Ordblindhed er velbeskrevet. Ordblindhed (dysleksi) er en specifik indlæringsvanskelighed og kan beskrives som vanskeligheder med at lære at udnytte skriftens lydprincip (Elbro, 2007). Skriftens lydprincip dikterer, at der til hvert bogstav hører én eller flere bestemte sproglyde. Det kan være standardudtaler og betingede udtaler (til bogstavet p hører fx en standardudtale [p], som i ordet *pille*, og en betinget udtale [b] som i *spille og kop*). Den ordblinde har svært ved at tildele de korrekte sproglyde til bogstaverne og danne den korrekte syntese af lydene. Særligt læsning af nye ord er svært, fordi skriftens lydprincip er den væsentligste nøgle til afkodning af ord, man ikke har set på skrift før.

Den ordblinde kommer ofte langsomt i gang med at læse, og det er almindeligt at opleve vedvarende problemer med læsehastighed og stavning. Den internationale ordblindeforening benytter følgende definition, der også inkluderer en beskrivelse af vanskelighedernes typiske årsag og desuden kendetegn, der kan beskrives som følger af de primære vanskeligheder:

Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge. (Lyon, Shaywitz & Shaywitz, 2003, side 2).

Definitionen understreger læse- og stavevanskelighedernes udspring i en fonologisk svaghed. Underliggende fonologiske problemer findes hos næsten alle ordblinde (Ramus, Rosen, Dakin, Day, Castellote, White & Frith, 2003; Vellutino, Fletcher, Snowling & Scanlon, 2004).

Som det også fremgår af den internationale definition, kan de dårlige afkodningsfærdigheder få konsekvenser for den ordblindes tekstforståelse. Hvis afkodning af de enkelte ord i en tekst kræver mange mentale ressourcer af læseren, kan det medføre, at læseren har begrænsede ressourcer i overskud til at konstruere en sammenhængende, meningsfuld repræsentation af teksten (Perfetti, 1985). Over tid kan svage afkodningsfærdigheder indirekte påvirke udvikling af andre færdigheder og viden, der er afgørende for sprogforståelse i læsning. Begrænset læseomfang og/eller læsning med begrænset udbytte kan fx hæmme ordforrådsudviklingen (Stanovich, 1993; Cunningham & Stanovich, 2001). Nogle ordblinde har foruden fonologiske problemer også en langsom benævnelseshastighed på test, hvor man fx skal benævne (kendte) billeder så hurtigt som muligt. Sådanne dobbelte vanskeligheder følges med endnu større afkodningsvanskeligheder end rene fonologiske problemer (Elbro, 2007). Ordblindes *kerne vanskelighed* i læsning er dog altid knyttet til afkodningskomponenten i den enkle læsemodel.

Vanskeligheder med sprogforståelse i læsning

Specifikke vanskeligheder med sprogforståelse i læsning er som nævnt langt mindre velbeskrevne end ordblindhed. Det kan skyldes mange ting. Først og fremmest er forståelse ganske givet en yderst kompleks færdighed, og sprogforståelse afhænger formentlig af en lang række sproglige, pragmatiske, kognitive og sociale færdigheder. Vanskeligheder med sprogforståelse i læsning er dermed næppe en lige så (relativ) homogen indlæringsvanskelighed, som ordblindhed er.

Desuden viser vanskeligheder med sprogforståelse i læsning sig ikke så tidligt som ordblindhed, fordi kravene til sprogforståelse er meget beskedne i starten af elevernes læseudvikling, hvor der er fokus på tilegnelse af afkodningsfærdigheder. Kravene er fx langt under det niveau, man forventer af skolebegynderens forståelse af talt sprog. Da sprogforståelseskomponenten – i overensstemmelse med den udvikling der forudsiges af den enkle læsemodel – har lille betydning for variationen i læseforståelse i de første skoleår, er det derfor også forventeligt, at undersøgelser, der forsøger at afgrænse dårlige forståere i de tidlige skoleår, ofte finder meget få læsere med denne profil (fx Spencer, Quinn & Wagner, 2014; Rønberg & Petersen, 2015a).

Imidlertid kan man i mange tilfælde identificere svagheder i sprogforståelse i talt sprog hos børn, der endnu ikke har lært at læse, men som senere viser sig at have specifikke forståelsesvanskeligheder i læsning (Elwér, Gustafson, Byrne, Olson, Keenan & Samuelsson, 2015). Dette er i overensstemmelse med den enkle læsemodel, hvor sprogforståelseskomponenten indgår i forståelsen af både talt og skrevet sprog. Børn, der er blevet identificeret med specifikke forståelsesvanskeligheder i læsning, vil da ofte også have problemer, når de lytter til en oplæst tekst i stedet for at læse den selv (Cain & Oakhill, 1999). Dog er det oplagt, at sprogforståelsesvanskeligheder for nogle børns vedkommende kan komme forskelligt til udtryk, når man ser på deres forståelse af talt sprog og skrevet sprog på grund af de særlige karakteristika, der knytter sig til skrevet sprog og talt sprog.

Talt sprog er karakteriseret ved at foregå 'her og nu', og talt sprog involverer en modtager, der ansigt til ansigt responderer på det sagte. I det talte sprog bruges intonationen støttende og kan bidrage med selvstændig betydning til det sagte. Samtalepartnerne deler ofte en fysisk kontekst, som de kan inddrage umiddelbart uden først at beskrive denne kontekst detaljeret, og de kan understrege og supplere det sagte med blikretning og gestik. Skrevet sprog er derimod karakteriseret ved at indgå i 'enmandsaktiviteter'. Afsender og modtager er oftest alene, når de henholdsvis skriver eller læser. De deler sjældent en fysisk kontekst, og det er derfor nødvendigt for afsenderen at være eksplicit og detaljeret for at sikre sig, at modtageren forstår hans eller hendes budskab. Til gengæld har afsenderen af skrevet sprog typisk også god tid til at vælge sine ord med omhu sammenlignet med afsenderen af talt sprog. Måske som en konsekvens heraf indebærer skrevet sprog typisk længere og mere komplekse sætninger samt mindre frekvente og mere avancerede ord end talt sprog. På en række punkter kan skrevet sprog dermed forekomme mere krævende for sprogforståelsen end talt sprog. Men omvendt har skrevet sprog som oftest det fortrin frem talt sprog, at det er tilgængeligt i længere tid for modtageren. Modtageren (læseren) kan dermed gå tilbage i teksten, genlæse udvalgte afsnit og ord, lægge teksten fra sig og genoptage læsningen på et senere tidspunkt. Tilgængeligheden af teksten kan fx være med til at mindske krav til hukommelsen, og kan dermed ad den vej støtte sprogforståelsen (Garton & Prat, 1989; Elbro, 2014; Cain, 2010).

Sammensatte vanskeligheder i læsning

Sammensatte vanskeligheder i læsning ('garden variety' ifølge Gough et al.) er kombinationer af afkodnings- og sprogforståelsesvanskeligheder. Mange børn med tidlige sprog- og kommunikationsvanskeligheder udvikler sammensatte vanskeligheder i læsning (Ricketts, 2011). I en undersøgelse af Catts, Fey, Tomblin & Zhang (2002), hvor man fulgte et stort sample af førskolebørn med sprogvanskeligheder indtil 10-årsalderen, havde ca. 35% af disse børn sammensatte vanskeligheder i læsning, 15% havde specifikke forståelsesvanskeligheder i læsning og 10% havde specifikke afkodningsvanskeligheder. Børn med tidlige sprogvanskeligheder forventes særligt at være i risiko for sammensatte læsevanskeligheder, hvis de foruden fonologiske problemer også har problemer med sprogets semantik (betydning) og/eller syntaks (sætningsdannelse og -forståelse) (Bishop & Snowling, 2004). Hvis sprogvanskeligheder ledsages af andre kognitive (nonverbale) vanskeligheder, er der risiko for endnu mere gennemgribende læsevanskeligheder (Catts et al., 2002).

Mulige årsager til dårlig sprogforståelse i læsning

Undersøgelser har belyst forskellige psykolingvistiske årsager til specifikke vanskeligheder med sprogforståelse i læsning (fx Cain & Oakhill, 1999; Nation, Clarke, Marshall & Durand, 2004). De mest velbeskrevne problemer hos dårlige forståere er med ordforråd og ordkendskab, grammatisk

viden og grammatisk opmærksomhed, viden om teksters struktur, højere kognitive færdigheder som monitorering og inferensdragning, og med arbejdshukommelse.

Dette afsnit beskriver, hvordan de enkelte færdigheder støtter tekstforståelse, og det fremlægger belæg for at opfatte problemer med de enkelte færdigheder som årsager til dårlig sprogforståelse i læsning. Afsnittet præsenterer resultater fra sammenligningsundersøgelser med dårlige og gode forståere, resultater fra langtidsundersøgelser, og resultater fra effektundersøgelser med dårlige forståere og/eller elever generelt.

Ordforråd og ordkendskab

Der kan ikke sættes spørgsmålstegn ved, at ordforråd er af afgørende betydning for tekstforståelse. Forståelse af en tekst er simpelthen ikke mulig, hvis ikke læseren har viden om i hvert fald nogle af ordbetydningerne i teksten (Nation, 2009). Flere undersøgelser har da også påvist, at mange dårlige forståere har begrænsninger i ordforrådet.

Det er for det første, hvor mange ord eleverne kender til (ordforrådets bredde), der kan være begrænset hos dårlige forståere. Ordforrådets bredde vurderes ofte med en billedudpegningstest, hvor eleven i hver opgave hører et ord og skal pege på dét af flere mulige billeder, som bedst passer til ordet. Spencer, Quinn & Wagner (2014) afgrænsede dårlige forståere i tre store kohorter (data fra tre skoleår) med elever på både 1., 2. og 3. klassetrin. De fandt som tidligere nævnt relativt få dårlige forståere (prævalensen var max 2,25%), men en stor del af de elever, der blev identificeret havde dårligt ordforråd målt med en billedudpegningstest (1. og 2. klassetrin). På 2. klassetrin var det i hvert af de tre skoleår lidt over halvdelen af de dårlige forståere, der havde dårligt ordforråd.

Begrænsninger i ordforrådet hos dårlige forståere kan også gælde ordforrådets dybde, det vil sige hvor *meget* eleverne ved om de enkelte ord (hvor mange semantiske detaljer, de har knyttet til ordbetydningerne) og hvor mange forbindelser (semantiske relationer), de har etableret mellem ordene (Nation, Clarke, Snowling, 2002; Ricketts, Nation & Bishop, 2007; Nation et al., 2004). Hvor meget eleverne ved om ord, undersøges ofte med opgaver, hvor eleverne skal definere udvalgte ord. Jo flere korrekte semantiske detaljer elevernes definitioner inkluderer, des større kendskab til ordet. Nation et al. (2004) fandt, at der var signifikant forskel på, hvor meget viden dårlige og gode¹ forståere havde om ord i en sådan test med definitionsopgaver. På individniveau scorede 8 ud af 23 dårlige forståere (35%) mindst en standardafvigelse under gennemsnittet for deres aldersgruppe på testen. I sammenligning hermed scorede 3 ud af 22 gode forståere (14%) under -1 standardafvigelse.

Kendskab til relationer mellem ordbetydninger er blandt andet blevet undersøgt med opgaver, hvor eleverne skal forklare, hvordan forskellige ord (fx *rød* – *grøntsag* – *tomat*) er relateret. Med en standardiseret test af kendskab til relationer mellem ordbetydninger fandt Nation et al. (2004), at tre ud af 23 dårlige forståere (13%) scorede mere end -1 standardafvigelse fra gennemsnittet for deres aldersgruppe. Ingen gode forståere scorede under gennemsnittet.

¹ Gode forståere kan i undersøgelsen af Nation et al. (2004) betegnes som 'gennemsnitlige til gode forståere', og Nation et al. betegner dem i deres artikel som 'kontrolbørn'. Eleverne kunne have læseforståelse svarende til en z-score på -0,33 på den normerede test, men gruppens gennemsnit lå tilsvarende over middel.

Betydningen af ordforråd for tekstforståelse understøttes af effektundersøgelser (Beck, Perfetti & McKeown, 1982; Vadasy, Sanders & Herrera, accepteret). Imidlertid er effekten af at undervise i specifikke ord begrænset på generel tekstforståelse (Elleman, Lindo, Morphy & Compton, 2009; National Reading Panel, 2000). Undervisning i specifikke ord er desuden tidskrævende, og det er usandsynligt, at man kan undervise elever med dårligt ordforråd direkte i *alle* de ord, der skal til for at rykke et dårligt ordforråd væsentligt. Et alternativ, eller supplement til undervisning i specifikke ord, kan derfor være at forsøge at gøre eleverne interesserede i at opsøge og undersøge nye ord på egen hånd (Beck & McKeown, 2013). En anden mulighed er at lære elever, hvordan de kan udnytte konteksten omkring et nyt ord og viden om verden til at slutte sig til (inferere) betydningen af de nye ord, de møder i tekster. Læsning af tekster er en meget vigtig kilde til udvidelse af ordforrådet (Cunningham & Stanovich, 2001), men dårlige forståere kan have sværere ved at udlede ordbetydninger fra tekst end gode forståere, når der er en vis afstand mellem de relevante informationer i teksten (Cain, Oakhill & Lemmon, 2004). Der er netop undersøgelser, der indikerer, at det er muligt at lære elever at blive bedre til at udnytte kontekst og viden til at udlede nye ords betydning (Cain, 2007a). Det gælder også for elever med dårligt ordforråd (Nash & Snowling, 2006).

Relationen mellem ordforråd, inferensfærdigheder og læseforståelse behandles yderligere i kapitel 5.

Grammatiske færdigheder

Det synes oplagt i en eller anden udstrækning at forvente en årsagssammenhæng mellem grammatiske færdigheder og tekstforståelse. Tidligt i sprogudviklingen vil der fx være grammatiske konstruktioner, som et barn endnu ikke har tilegnet sig (Cain, 2007b). Overfor talt eller skrevet sprog, hvori disse konstruktioner indgår, synes barnet alt andet lige at være bedre stillet med end uden en forståelse af disse konstruktioner. Det er ikke nok at kunne forstå de enkelte ord i en sætning – den grammatiske konstruktion, de indgår i, har stor betydning for meningsindholdet.

Nogle undersøgelser har fundet, at dårlige forståere har relativ dårlig grammatisk forståelse (Nation et al., 2004; Li & Kirby, 2014). Der må dog tages forbehold for konklusioner baseret på flere af disse undersøgelser på grund af den måde, man har undersøgt 'grammatisk forståelse' på. Fx anvendte Nation et al. (2004) og Nation, Cocksey, Taylor og Bishop (2010) en test med gentagelse af sætninger, der øgedes i længde og kompleksitet, som udtryk for grammatisk forståelse. Nation et al. (2004) fandt, at 11 ud af 22 dårlige forståere, der prøvede testen, scorede mere end 1 standardafvigelse under gennemsnittet for aldersgruppen. Til sammenligning scorede tre ud af 22 gode forståere under denne grænse. Nation et al. (2010) fandt, at de dårlige forståeres dårlige resultater på testen var synlige allerede ved læseundervisningens start, hvilket kunne tyde på, at svagheder i grammatisk forståelse kan være en årsag til dårlig læseforståelse. De anvendte test med gentagelse af sætninger af tiltagende længde kræver dog ud over grammatisk forståelse (og eventuelt ordkendskab) også en del af deltagerens verbale arbejdshukommelse (se nedenfor) i takt med, at mere og mere verbalt materiale skal holdes hukommelsen, mens deltageren lytter til stimuli. Gentagelsestest med verbalt materiale, der præsenteres i tiltagende længde, anvendes netop også

som test af verbal korttids- og arbejdshukommelse. Dette forbehold modererer de konklusioner, man kan drage om dårlige forståeres grammatiske forståelse ud fra undersøgelserne af Nation og kolleger.

Andre test er også blevet anvendt som en indikator på grammatisk forståelse i undersøgelser med dårlige forståere. Ofte skal læseren udpege ét af fire billeder, der matcher en kompleks sætning, som læseren hører en enkelt gang. Stothard og Hulme (1992) fandt, at dårlige forståere scorede signifikant lavere på en sådan test af grammatisk forståelse (TROG) end en aldersmatchet gruppe af gode af forståere, men på niveau med en gruppe yngre elever, der var matchet med de dårlige forståere på tekstforståelse. Stothard og Hulme behandlede selv testresultaterne som et generelt udtryk for forståelse af talt sprog (lytteforståelse). Det forhold understreger, at det kan være svært at skelne test af grammatisk forståelse fra test af generel sprogforståelse. Den samme test af grammatisk forståelse blev anvendt af Oakhill og Cain (2012) i en langtidsundersøgelse. De fandt, at grammatisk forståelse ikke forudsagde senere læseforståelse, når man kontrollerede for den autoregressive effekt af læseforståelse og effekten af andre komponenter i sprogforståelse som monitorering og inferensfærdigheder.

Ofte er det dog også et andet aspekt ved grammatiske færdigheder end grammatisk forståelse, nemlig grammatisk *opmærksomhed*, der er blevet undersøgt i forbindelse med læseforståelsesvanskeligheder. Grammatisk opmærksomhed er evnen til at manipulere og reflektere over de grammatiske strukturer i sproget (Cain, 2007b). Der kan både være tale om syntaktiske og morfologiske strukturer (henholdsvis sætningers og ords opbygning). Nation og Snowling (2000) undersøgte dårlige og gode forståeres syntaktiske opmærksomhed med en opgave, hvor eleverne skulle lytte til sætninger, hvori ordene var blevet tilfældigt kombineret, og derefter rette dem til korrekte sætninger (fx: *trommen aben af slået blev* → *trommen blev slået af aben*). Nation og Snowling fandt, at dårlige forståere havde signifikant dårligere opmærksomhed på syntaks end gode forståere.

Undersøgelser af morfologisk opmærksomhed hos dårlige forståere har fokuseret på flere forskellige aspekter af morfologi. Nation et al. (2004) undersøgte med en eksperimentel test dårlige og gode forståeres evne til at bøje verber korrekt i datid. Eleverne skulle indsætte den korrekte datidsbøjning af et verbum i udfyldningsopgaver (fx: *I dag skinner solen. I går ____ solen*). Dårlige forståere var som gruppe betragtet signifikant dårligere end gode forståere til dette. Tong, Deacon, Kirby, Cain og Parrila (2011) undersøgte foruden bøjning af ord også afledning af ord. De sammenlignede færdigheder hos dårlige forståere, gennemsnitlige forståere og gode forståere. De fandt, at dårlige forståere *ikke* adskilte sig fra deres jævnaldrende på opgaver, der krævede bøjning af ord (fx *kører* – *kørte*), men at dårlige forståere i 5. klasse scorede dårligere end de to andre grupper af forståere på opgaver, der krævede afledning af et ord fra én ordklasse til et ord i en anden ordklasse (fx *kører* – *kørsel*) – dette resultat var også signifikant efter statistisk kontrol for forskelle i ordforråd. Grupperne adskilte sig dog ikke på afledningsopgaverne på samme måde to år tidligere. I 3. klasse var der ikke forskel på dårlige forståere og gennemsnitlige forståere, men gode forståere var signifikant bedre end de to andre grupper. Det resultat *kan* indikere, at problemer med afledning kan være en konsekvens af, snarere end en årsag til, dårlig læseforståelse. I en longitudinel undersøgelse fandt Muter et al. (2004) dog et selvstændigt bidrag fra et sammensat mål for tidlig opmærksomhed på morfologiske og syntaktiske strukturer til læseforståelse ét år senere.

På tværs af effektundersøgelser med læsere med og uden forskellige læsevanskeligheder er der indikationer på, at undervisning i morfologisk opmærksomhed kan have en effekt på forståelsesmål som læseforståelsestest (Bowers, Kirby & Deacon, 2010). Det er dog sandsynligt, at en del af denne effekt kan tilskrives forbedrede afkodningsfærdigheder. Morfologisk analyse kan netop anvendes i afkodningsundervisning med succes (Bowers et al., 2010) og elever med afkodningsvanskeligheder kan have gavn af undervisning i morfologisk analyse (Arnbak & Elbro, 2000). Der er dog også undersøgelser, der tyder på, at morfologisk undervisning kan påvirke sprogforståelsen i læsning indirekte via ordkendskabet, hvis den morfologiske analyse kombineres med strategier til at udlede betydningen af andre ord ud fra deres sammensætning af morfemer (Baumann, Edwards, Font, Tereshinski, Kame'enui & Olejnik, 2002; Baumann, Edwards, Boland, Olejnik & Kame'enui, 2003; Tomesen & Arnoutse, 1998).

Viden om teksters struktur

Der knytter sig forskellige typiske strukturer til tekster i forskellige genrer. Har læseren kendskab til de typiske strukturer i den type tekst, som han eller hun læser, har læseren en vigtig støtte for forståelse i dette kendskab. Læseren ved da nemlig på forhånd, hvilken slags informationer det er sandsynligt at finde i denne type tekst, og hvor informationerne sandsynligvis vil være placeret i teksten.

Det er oplagt, at viden om teksters struktur til dels er en effekt af eksponering for tekster (Perfetti, 1984). Dårlige forståere har vist sig at have mindre viden om teksters struktur end gode forståere, og de producerer også mindre strukturerede historier end gode forståere (Cain & Oakhill, 1996; Cain, 2003; Yuill & Oakhill, 1991). Cain og Oakhill (1996) undersøgte fx dårlige forståeres viden om teksters struktur ved at stille spørgsmål som: *Hvad kan en histories titel fortælle dig? Hvad kan en histories begyndelse (eller slutning) fortælle dig?* Sammen med puslehistorier (historier, hvor man blander sætningerne i en kort historie, og beder eleverne samle sætningerne i en god rækkefølge) kan sådanne spørgsmål desuden forudsige senere læseforståelse – selv når man tager højde for den autoregressive effekt af forståelse og betydningen af ordforråd og generelle sproglige færdigheder (sidstnævnte blev undersøgt med en standardiseret test af 'verbal IQ') (Oakhill & Cain, 2012).

Williams og kolleger har gennemført en række effektundersøgelser med tekststrukturer i fagtekster i fokus (Williams, Hall & Lauer, 2004; Williams, Hall, Lauer, Stafford, DeSisto, deCani, 2005; Williams, Nubla-Kung, Pollini, Stafford, Garcia & Snyder, 2007; Williams, Stafford, Lauer, Hall & Pollini, 2009). Information i fagtekster kan fx være struktureret som en beskrivelse eller rækkefølge. Det er også almindeligt, at information i fagtekster er struktureret omkring årsagsforhold, et problem og dets løsning eller ligheder og forskelle (Meyer & Freedle, 1984). Williams og kolleger underviste fx elever i at genkende og bruge årsag-følge strukturer og strukturer for ligheder og forskelle. Eleverne fik bl.a. indføring i signalord (fx *derfor*, *fordi* og *ligesom*, *i modsætning til*) og lærte at udfylde grafiske modeller, der synliggjorde strukturerne. Der var generelt effekt af undervisningen; eleverne lærte at genkende strukturerne i tekster. På flere mål

af overføring til forståelse havde eleverne desuden signifikant bedre forståelse af nye tekster med strukturerne end kontrolgrupperne (Williams et al., 2007; 2009).

Monitorering

Læseren, der monitorerer sin forståelse, er opmærksom på, i hvilken grad han eller hun faktisk forstår teksten. Hvis læseren opdager, at forståelsen halter – at det læste af den ene eller anden grund ikke giver mening – har læseren mulighed for at gøre noget for at komme frem til en bedre forståelse af teksten. Det kan fx være at slå ord op, genlæse det netop læste, gå længere tilbage i teksten og læse et tidligere afsnit igen eller hoppe frem i teksten og lede efter information, der kan hjælpe.

Dårlige forståere monitorerer deres forståelse i mindre grad end gode forståere (Oakhill, Hart & Samols, 2005; de Sousa & Oakhill, 1996; Ehrlich, Remond & Tardieu, 1999). Monitorering kan undersøges eksperimentelt ved at bede eleverne om at tage stilling til, om en række forskellige tekster bør redigeres, eller om de er gode nok, som de er. I en del af teksterne er der så indlejret fejl som uoverensstemmelser mellem to informationer i teksten eller uoverensstemmelser mellem information i teksten og almindelig viden. Fx læste eleverne i undersøgelsen af Oakhill et al. (2005) en tekst om muldvarper. Eleverne læste bl.a., at muldvarpen ser meget dårligt. Senere i teksten stod der, at muldvarpen bruger sit gode syn til at finde føde til sine unger. Elever, der monitorerede deres forståelse af teksten effektivt, sendte teksten til redigering på grund af denne uoverensstemmelse mellem tekstens informationer om muldvarpens syn. I en undersøgelse af Ehrlich et al. (1999), hvor fejlene i teksten bestod af forkert brug af pronominer, inddrog man også online mål for monitorering (man målte den hastighed, hvormed eleverne læste pronominerne i teksten og den umiddelbart efterfølgende tekst samt antallet af situationer, hvor eleverne gik tilbage i teksten i forbindelse med læsning af et forkert pronomen). Dårlige forståere sænkede læsehastigheden, når de stødte på et forkert pronomen, men ikke så markant som gode forståere, som i øvrigt også sænkede hastigheden, mens de læste den umiddelbart efterfølgende tekst. Og dårlige forståere gjorde markant mindre brug af muligheden for at gå tilbage i teksten og genlæse. Undersøgelsen viste også, at dårlige forståere havde tendens til at overestimere deres egen forståelse i sammenligning med gode forståere ved læsning af tekster *uden* uoverensstemmelser.

Monitorering forudsiger læseforståelse over tid. Det gælder også, når der kontrolleres for den autoregressive effekt af læseforståelse og betydningen af ordforråd og generelle sproglige færdigheder (Oakhill & Cain, 2012). I en omfattende litteraturundersøgelse (National Reading Panel, 2000) fandt man 20 effektundersøgelser, der havde undersøgt effekten af at undervise elever i at blive mere opmærksomme på eventuelle problemer med at forstå, det de læste. Undervisningen bestod primært i, at underviserne *demonstrerede* opmærksomhed på forståelse, og eleverne lærte desuden, *hvad* de kunne gøre i tilfælde, hvor der var noget i en tekst, de ikke forstod. I fem af syv undersøgelser, hvor man undersøgte, om der var overføringseffekt til læseforståelse af en sådan undervisning, fandt man, at eleverne forbedrede deres læseforståelse.

Inferensfærdigheder

For at opnå en sammenhængende og meningsfuld forståelse af det læste, må læseren skabe forbindelser på tværs af tekstens dele (ord, sætninger og afsnit) og bidrage med informationer, der ikke fremgår eksplicit af teksten – læseren må drage inferenser. Inferensdragning er et gennemgående krav ved læsning af tekster. Selv en simpel tekst som den følgende kræver, at man drager mindst én inferens for at opnå sammenhæng i tekstens informationer:

Solveig ønskede sig en baby. Solveig begyndte at se sig om efter en kæreste.

Der kan selvfølgelig være mange grunde til, at Solveig ser sig om efter en kæreste, men læseren må ud fra tekstens informationer og relevant baggrundsviden gå ud fra, at handlingen har noget at gøre med Solveigs ønske om en baby: Læseren må drage en inferens om et årsagsforhold.

En lang række undersøgelser har dokumenteret, at dårlige forståere er dårligere til at drage inferenser end gode forståere (fx Oakhill, 1982;1984; Cain & Oakhill, 1999; Bowyer-Crane & Snowling, 2005; Catts et al., 2006). Inferensfærdigheder undersøges typisk ved at lade elever læse eller lytte til historier og derefter stille inferenskrævende spørgsmål til det læste eller hørte (fx: *Hvorfor begyndte Solveig at se sig om en kæreste?*). Dårlige forståere er markant dårligere end gode forståere til at svare på sådanne inferenskrævende spørgsmål. Til gengæld er dårlige forståere ikke nødvendigvis dårligere til at besvare spørgsmål til teksten, der ikke kræver inferenser (fx: *Hvad gjorde Solveig?*) (Cain & Oakhill, 1999). Catts et al. (2006) undersøgte dårlige forståeres inferensdragning fra historier, som eleverne lyttede til. Catts et al. anvendte besvarelser på ikke-inferentielle spørgsmål til teksten som en kontrolvariabel. Efter kontrol for elevernes evne til at besvare sådanne spørgsmål, var der fortsat signifikant forskel på dårlige og gode forståeres evne til at besvare inferenskrævende spørgsmål, når de informationer i teksten, der skulle forbindes, blot ikke fulgte umiddelbart efter hinanden i teksten.

Baggrundsviden er vigtig for læseforståelse generelt (Catts, 2009), og baggrundsviden er også yderst vigtig i forbindelse med inferensfærdigheder. Dårlige inferensfærdigheder skyldes dog ikke i alle tilfælde manglende viden. Cain, Oakhill, Barnes & Bryant (2001) sammenlignede inferensfærdigheder hos gode og dårlige forståere i en undersøgelse, hvor de sikrede sig, at eleverne havde den viden, det var nødvendigt at trække på i inferenserne. Dårlige forståere havde også her signifikant dårligere inferensfærdigheder end gode forståere.

Inferensfærdigheder forudsiger læseforståelse over tid, også når man tager højde for den autoregressive effekt af forståelse og betydningen af ordforråd og generelle sproglige færdigheder (Oakhill & Cain, 2012). Effektundersøgelser har også vist effekt på læseforståelse af at undervise elever i drage inferenser. Det gælder både for dårlige forståere (Yuill & Joscelyne, 1988; Yuill & Oakhill, 1988; McGee & Johnson, 2003) og for læsere generelt (Hansen, 1981; McMaster, 2012; Elbro & Buch-Iversen, 2013).

Der er således et stærkt belæg for inferensfærdigheders betydning for dårlige forståeres vanskeligheder og for forskelle i læseforståelse blandt læsere generelt. I denne afhandling er der netop fokus på inferensfærdigheder (kapitel 4,5 og 6). Kapitel 4 handler fx om forskellige aspekter ved inferensfærdigheder, der kan spille en *særlig* rolle for tekstforståelse.

Arbejdshukommelse

En række almene psykologiske færdigheder som opmærksomhed, refleksionsniveau (IQ) og almen viden er korreleret med læseforståelse (fx Keenan & Meenan, 2014; Catts, 2009).

Arbejdshukommelse er en generel psykologisk mekanisme, der har fået meget opmærksomhed i undersøgelser af dårlige forståere. Arbejdshukommelse er evnen til at lagre information midlertidigt (i korttidshukommelsen), samtidig med at denne eller anden information bearbejdes.

Arbejdshukommelsen er dermed involveret i mange kognitive aktiviteter, som fx sprogforståelse og problemløsning (Baddeley, 1992).

Mange undersøgelser har dokumenteret relativt lave præstationer på test af arbejdshukommelse hos dårlige forståere i sammenligning med gode forståere (fx De Beni, Palladino, Pazzaglia & Cornoldi, 1998; Oakhill, Hart & Samols, 2005). Flere af de specifikke komponenter i sprogforståelse, som monitorering og inferensfærdigheder, er afhængige af, at læseren kan holde information umiddelbart tilgængeligt i hukommelsen simultant med, at informationen bearbejdes. Men arbejdshukommelse kan ikke fuldt ud forklare betydningen af monitorering og inferensfærdigheder for læseforståelse (Cain, Oakhill & Bryant, 2004). Der er selvstændige bidrag fra disse komponenter til læseforståelse efter kontrol for forskelle i arbejdshukommelse.

I langtidsundersøgelsen af Oakhill og Cain (2012) var arbejdshukommelsesressourcer ikke blandt de variable, der kunne forudsige læseforståelse over tid (som ordforråd, viden om teksters struktur, monitorering og inferensfærdigheder). I en metaundersøgelse af effekten af arbejdshukommelse konkluderede Melby-Lervåg og Hulme (2013) da også, at der ikke er signifikant overføringseffekt af arbejdshukommelsestræning til sproglige færdigheder.

Relationen mellem arbejdshukommelse og inferensfærdigheder i læseforståelse behandles yderligere i kapitel 5.

Sammenfatning

Ifølge den enkle læsemodel er god tekstforståelse afhængig af både gode afkodningsfærdigheder og god sprogforståelse. Denne antagelse underbygges af en lang række iagttagelser og undersøgelser. Det er blandt andet forekomsten af forskellige slags læsevanskeligheder: Nogle læsere har specifikke vanskeligheder med afkodning (ordblindhed), andre læsere har specifikke vanskeligheder med sprogforståelse i læsning, og andre læsere igen har sammensatte vanskeligheder, det vil sige vanskeligheder med både afkodning og sprogforståelse. Elever med specifikke vanskeligheder med sprogforståelse i læsning kaldes også 'dårlige forståere'. Ud fra den enkle læsemodel kan man udlede en generel definition af dårlige forståere: 'Vanskeligheder med tekstforståelse der ikke skyldes dårlig afkodning'. Hvor ordblindhed er velbeskrevet, er vanskeligheder med sprogforståelse endnu ikke så velundersøgt. En række undersøgelser har dog dokumenteret mulige årsager af psykologisk art til forståelsesvanskeligheder. Disse undersøgelser inkluderer sammenligninger af dårlige og gode forståere, langtidsundersøgelser og

effektundersøgelser. De bedst afdækkede mulige årsager til dårlig sprogforståelse i læsning er problemer med ordforråd og ordkendskab, dårlige grammatiske færdigheder, begrænset viden om teksters struktur, begrænset monitorering af egen forståelse og dårlige inferensfærdigheder. En del undersøgelser har også interesseret sig for arbejdshukommelsesproblemer som mulig årsag til dårlig sprogforståelse i læsning. Der er i forskellig grad belæg for de mulige årsager til forståelsesvanskeligheder. Belægget synes at være særligt stærkt for ordforråd og inferensfærdigheder. For inferensfærdigheder er der fx adskillige sammenligningsundersøgelser, der har dokumenteret forskelle på gode og dårlige forståere; inferensfærdigheder forudsiger læseforståelse over tid selv efter kontrol for andre relevante færdigheder som bl.a. ordforråd; og en række effektundersøgelser har vist, at forbedrede inferensfærdigheder kan føre til bedre læseforståelse.

Hvordan man rent praktisk afgrænser dårlige forståere fra andre læsere og hvilken stabilitet, der er for forståelsesvanskelighederne, er emnet i de følgende kapitler. Dernæst vendes fokus i de efterfølgende kapitler mod inferensfærdigheder og deres betydning for læseforståelse.

Kapitel 2. Afgrænsning af gruppen af dårlige forståere

Ud fra den enkle læsemodel kan specifikke vanskeligheder med sprogforståelse i læsning som nævnt defineres som 'vanskeligheder med tekstforståelse, som ikke skyldes dårlig afkodning'. Denne definition er så generel, at den rejser en lang række spørgsmål i forbindelse med den praktiske afgrænsning af gruppen af dårlige forståere. I forskningslitteraturen finder man da også mange forskellige metoder til afgrænsning af dårlige forståere. De forskellige afgrænsningsmetoder er kun sjældent eksplicit begrundet i forskningsartiklerne, men særligt en række indledende valg i afgrænsningen kan ofte forstås meningsfuldt i sammenhæng med de særlige forskningsinteresser i undersøgelse. Det er fx, hvilke baggrundsinformationer og hvilke andre færdighedsmål ud over test af læseforståelse og afkodningsfærdigheder man bringer i spil i afgrænsningen. I de følgende afsnit bliver der givet eksempler på forskelle i de indledende valg i afgrænsning af dårlige forståere i konkrete undersøgelser, og det bliver uddybet, hvordan undersøgelsesnes forskellige formål kan ses afspejlet i nogle af disse valg. De mange forskellige fremgangsmåder i afgrænsning af grupper af dårlige forståere i forskning kan udgøre et problem for overføring til klinisk praksis, og i et selvstændigt afsnit skitserer jeg derfor principper for en fremgangsmåde, der kan være hensigtsmæssig i praksis.

De anvendte færdighedsmål for læseforståelse og afkodning er naturligvis også en faktor, der varierer på tværs af undersøgelser. Hvilke testmaterialer, man har til rådighed til at undersøge elevers afkodning og tekstforståelse med, kan have konsekvenser for mulighederne for at identificere dårlige forståere i såvel forskning som praksis. Hvordan læseforståelse og afkodning er blevet undersøgt i forbindelse med afgrænsning af dårlige forståere, er emnet for de efterfølgende afsnit, og jeg fremhæver egenskaber ved færdighedsmålene, der synes at kunne være afgørende for en hensigtsmæssig afgrænsning af dårlige forståere.

Sidst, men ikke mindst er der en helt central forskel på tværs af undersøgelser, når det kommer til, hvordan man trækker grænsen mellem dårlige forståere og andre læsere ud fra deltagernes resultater på test af læseforståelse og afkodning. Også her giver jeg eksempler på forskellige metoder, der anvendes i forskning, og jeg illustrerer konsekvenserne af forskellige afgrænsningsmetoder grafisk. Afslutningsvis giver jeg et forslag til en afgrænsningsmetode, der er i bedst mulig overensstemmelse med den generelle definition af specifikke vanskeligheder med sprogforståelse i læsning.

Indledende valg i afgrænsningen

Afgrænsning fra udviklingsforstyrrelser og andre kendte vanskeligheder

Som beskrevet tidligere er elever med en udviklingsforstyrrelse i form af sprogvanskeligheder særligt udsat for vanskeligheder med sprogforståelse i læsning (Simkin & Conti-Ramsden, 2006; Kelso, Fletcher & Lee, 2007; Ricketts et al., 2011). Ligeledes kan elever med generelle kognitive vanskeligheder (lav IQ) forventes at have tilsvarende store vanskeligheder med sprogforståelse i læsning (Harlaar, Hayiou, Thomas & Plomin, 2005; Wingerden; Segers, van Balkom & Verhoeven, 2014). Der kan også være en overrepræsentation af problemer med sprogforståelse og eventuelle *særlige* problematikker i forbindelse med læsning forbundet med diagnoser som ADHD og autismespektrum-forstyrrelser og med medfødte eller erhvervede neurologiske eller sensoriske skader (Miller, Keenan, Betjemann, Willcutt, Pennington & Olson, 2013; Norbury & Nation, 2011; Barnes, Johnston & Dennis, 2007; Kelly & Barak-Cikoja, 2007).

Ved undersøgelser af specifikke vanskeligheder med sprogforståelse i læsning er det mest almindeligt at selekttere deltagere blandt elever, der ikke kendetegnes ved udviklingsforstyrrelser eller andre kendte vanskeligheder, som åbenlyst kan være årsag til eventuelle problemer med læsning (fx Ricketts et al., 2011). For at kunne udlede konklusioner om karakteristika mm. ved dårlige forståere i de tilfælde, hvor der *ikke* er en oplagt årsag til problemer med sprogforståelse i læsning, etablerer man i mange tilfælde deltagergrupper uden sådanne kendte vanskeligheder. Men der er ikke en ensrettet praksis, når det kommer til sådanne eksklusionskriterier i forskningslitteraturen. For nogle undersøgelser gælder det, at et fravalg af eksklusionskriterier for kendte vanskeligheder kan begrundes i forskningsspørgsmålet. Fx undersøgte Catts et al. (2006), hvordan svagheder i forskellige sproglige færdigheder er associeret med forskellige læsevanskeligheder over tid fra børnehaveklasse til 8. klassetrin. I overensstemmelse med dette formål var der ikke noget eksklusionskriterium, hvad angik sprogvanskeligheder i udvælgelsen af deltagere. Deltagere med sprogvanskeligheder var faktisk oversamlet². Der *var* derimod et eksklusionskriterium for deltagere, der ved det sidste undersøgelsestidspunkt (8. klasse) havde en nonverbal IQ-score på -2 standardafvigelser. Dette eksklusionskriterium blev ikke begrundet eller kommenteret, men den valgte grænseværdi svarer til en IQ-score på 70, som ofte angives som grænseværdi for mild retardering. Det er altså sandsynligt, at Catts et al. har ønsket at udelukke generel kognitiv funktionsnedsættelse som årsag til dårlig forståelse af talt og skrevet sprog blandt deltagerne i undersøgelsen. I en undersøgelse af Nation et al. (2002) var ét af formålene derimod *netop* at undersøge, hvor stor en andel af deltagerne i en gruppe af dårlige forståere, der havde generelle kognitive vanskeligheder. I overensstemmelse med dette formål anvendte man ikke et eksklusionskriterium for lav IQ.

² For at kunne generalisere deres resultater til populationen af læsere tog Catts et al. højde for overrepræsentationen af elever med sprogvanskeligheder ved at anvende vægtede scorer. De vægtede scorer var baseret på forekomsten af sprogvanskeligheder i en epidemiologisk undersøgelse.

Afgrænsning fra læsning på et andet- eller fremmedsprog

Begrænset erfaring med det sprog, man læser på, kan føre til dårlig forståelse af det læste. Elever med begrænset erfaring med dansk sprog vil fx oftere støde på ord, faste udtryk og sætningskonstruktioner, som de ikke kender, end elever med mere erfaring med dansk sprog. Elever med et andet første sprog end dansk vil som gruppe derfor overordnet set være i særlig risiko for sprogforståelsesproblemer i læsning i sammenligning med deres jævnaldrende med dansk som førstesprog (Elbro, 2007). Markant færre erfaringer med det relevante sprog er dog ikke nødvendigvis det, der gør sig gældende for andre læsere med forståelsesvanskeligheder. For at opnå indsigt i forhold forbundet med mere *overraskende* forståelsesvanskeligheder er det derfor i forskning almindeligt at afgrænse specifikke forståelsesvanskeligheder blandt børn, der læser på deres førstesprog (fx Keenan, Hua, Meenan, Pennington, Willcutt & Olson, 2014; Ricketts et al., 2011). Man undersøger altså specifikke forståelsesvanskeligheder hos en relativt homogen gruppe, hvor man på forhånd har sikret sig, at læserne ikke har dårlig læseforståelse af den grund, at de læser på et andet sprog end deres første sprog.

Kontrol med udvalgte færdigheder

Det er som nævnt velkendt, at begrænsninger i ordforråd og ordkendskab kan give vanskeligheder med sprogforståelsen i læsning. I undersøgelser af den mulige betydning af *andre* færdigheders betydning for læseforståelse har man af denne grund ofte på forhånd udelukket deltagere med et ordforråd, som ikke var alderssvarende, eller man har kontrolleret for ordforrådets indflydelse ad statistisk vej. I udvælgelsen af deltagere til en undersøgelse af sammenhængen mellem inferensfærdigheder og læseforståelse screenede Oakhill og Cain (1999) fx med en test, der foruden at kræve afkodningsfærdighed også krævede viden om ords betydning. Man ekskluderede elever med meget høje og meget lave resultater på denne test, inden man etablerede grupper af gode og dårlige forståere. Hermed blev variation i elevernes ordforråd begrænset, og grupperne havde da også sammenlignelige middelværdier i ordforråd. Man kunne således undersøge effekten af inferensfærdigheder isoleret fra effekten af ordforråd.

I andre undersøgelser finder man som nævnt – og ikke overraskende – at dårlige forståere bl.a. er karakteriseret ved begrænsninger i ordforrådet. Det gælder fx i en undersøgelse af Nation et al. (2004). I denne undersøgelse var hovedinteressen at vurdere elevernes grundlæggende sproglige færdigheder (som semantik, morfologi og syntaks) og på baggrund af disse testresultater anslå, hvor stor en andel af de dårlige forståere i udtrækket, der kunne siges at have milde og/eller oversete sproglige vanskeligheder. Der blev derfor ikke kontrolleret for forskelle i ordforråd i udvælgelsen af deltagere. Blandt andre færdigheder, der ses inddraget i afgrænsningen af dårlige forståere i forskning, er forskellige nonverbale færdigheder. Tong et al. (2011) tilstræbte fx, at dårlige forståeres vanskeligheder var overraskende ikke bare set i lyset af deres afkodningsfærdigheder, men også set i lyset af deres nonverbale færdigheder. Dette valg syntes dog ikke at være specifikt begrundet i undersøgelsens særlige forskningsinteresse, som var forskelle i dårlige og gode forståeres morfologiske færdigheder.

Afgrænsning i klinisk praksis

De fremgangsmåder, der anvendes til at identificere dårlige forståere i forskning, indebærer et problem for identifikation af dårlige forståere i klinisk praksis. Forskningsmetoderne munder alle ud i meget selekterede grupper af dårlige forståere – og der er dermed sandsynligvis elever med læseforståelsesvanskeligheder, som ikke bliver inkluderet. I forskningen kan man tillade sig en sådan selektion, og relativt homogene elevgrupper er desuden en nødvendighed, for at det bliver praktisk muligt at drage konklusioner om mulige årsager til dårlig læseforståelse. Men i praksis, hvor det overordnede formål med afgrænsningen er at blive i stand til at intervenere målrettet over for læseforståelsesvanskeligheder, må det bærende princip i afgrænsningen være, at man finder frem til så mange som muligt af de læsere, der har specifikke forståelsesvanskeligheder i læsning.

I den generelle definition af dårlige forståere indgår der ikke andre færdigheder end *tekstforståelse* og *afkodning*. Selektion på andre test end disse nøglefærdigheder, fx test af ordforråd og nonverbal IQ, er derfor ikke nødvendig eller hensigtsmæssig i en afgrænsning af dårlige forståere i praksis: Man ønsker at opdage specifikke forståelsesvanskeligheder i læsning hos en elev, uanset om eleven fx har relativt begrænset ordforråd eller ej, og uanset om eleven ligger under eller over gennemsnittet på en test af nonverbale færdigheder.

Aktiv brug af baggrundsoplysninger kan medvirke til, at man kan koncentrere sin afgrænsningsprocedure om specifikke forståelsesvanskeligheder blandt elever *uden* kendte vanskeligheder – det vil sige elever, hvis vanskeligheder *ikke* er indikeret i baggrundsoplysningerne. Det synes oplagt, at baggrundsoplysninger om kendte sproglige og kognitive vanskeligheder, der kan have betydning for elevers sprogforståelse, i sig selv kan berettige til en monitorering og eventuel udredning af elevens læseforståelse. På lignende vis kan man for at koncentrere sin afgrænsning om elever med *overraskende* vanskeligheder med sprogforståelse i læsning vælge at afgrænse dårlige forståere fra læsere med mulige andetsprogsvanskeligheder. Et sådan eksklusionskriterium giver i praksis dog kun mening i situationer, hvor der enten samtidig foretages en nærmere undersøgelse af de flersprogede elevers læseforståelse, eller hvor der i forvejen er fokus på deres sprogforståelse i talt og skrevet dansk sprog.

Undersøgelse af nøglefærdighederne

Undersøgelse af læseforståelse

En læseforståelsestest, der skal kunne bidrage til afgrænsningen af dårlige forståere, må være følsom for forskelle i sprogforståelse. Der er dog variation i, hvor store krav læseforståelsestest stiller til sprogforståelse (fx målt ved test af lytteforståelse) (Cutting & Scarborough, 2006; Keenan, Betjemann & Olson, 2008), og dårlige forståere har (ikke overraskende) større problemer på de test,

der er mest følsomme for sprogforståelse, og mindre problemer på de test, hvor præstationen i høj grad er afhængig af afkodningsfærdigheder (Nation & Snowling, 1997).

Desuden må det regnes for højst sandsynligt, at forskellige læseforståelsestest kræver mere eller mindre af de forskellige delfærdigheder i sprogforståelse, og denne variation kan også forventes at påvirke, hvilke dårlige forståere der kan identificeres med testene. Hvis en test fx stiller høje krav til inferensfærdigheder, må man forvente, at elever, der har dårlige inferensfærdigheder, let identificeres som dårlige forståere. Hvis en test derimod stiller få krav til inferensfærdigheder, er det mindre sandsynligt at finde dårlige forståere, der har dårlige inferensfærdigheder. En undersøgelse af Bowyer-Crane og Snowling (2005) med 10 dårlige og 10 gode forståere understøtter denne formodning om forskelle på forskellige tests krav til de forskellige komponenter i sprogforståelse. Der blev anvendt to læseforståelsestest i undersøgelsen: *Neale Analysis of Reading Ability – Revised* (NARA) (Neale & Whetton, 1997) og *Wechsler Objective Reading Dimensions* (WORD) (Rust, Golombok & Trickey, 1993). NARA blev anvendt i udvælgelsen af dårlige og gode forståere, men alle deltagere gennemførte også WORD. Bowyer-Crane og Snowling fandt, at gode forståere som gruppe præsterede bedre end dårlige forståere på begge test, men at dårlige forståere præsterede bedre på WORD sammenlignet med NARA, og at gode forståere præsterede bedre på NARA end på WORD. Bowyer-Crane og Snowling argumenterede for, at forskellene i gruppernes scorer på tværs af test kunne have at gøre med indholdsmæssige forskelle i testene. I en analyse af opgaverne i de to testmaterialer fandt Bowyer-Crane og Snowling netop, at testene i forskellig grad stillede krav til inferensfærdigheder: NARA indeholdt overordnet flere inferenskrævende opgaver, og fordelingen af forskellige slags inferenser var desuden forskellig på tværs af de to test. Lignende 'skævheder' må naturligvis også kunne gøre sig gældende på tværs af læseforståelsestest, når det gælder andre komponenter i sprogforståelse som fx ordforråd – der kan være relativt mange eller få avancerede ordbetydninger i en læseforståelsestest.

I kapitel 3 behandles forskelle på læseforståelsestest yderligere.

Undersøgelse af afkodningsfærdigheder

Afkodning af ord i kontekst eller isolerede ord

I flere af Cain og Oakhills undersøgelser (fx 1999) er det primære afkodningsmål i afgrænsningen af dårlige forståere ofte udledt fra læseforståelsestesten. Undervejs i deltagerens højtlesning af teksten er fejllæsninger blevet noteret (og korrigeret), og afkodningsmålet afspejler andelen af disse fejllæsninger. Spooner, Baddeley & Gathercole (2004) har påpeget, at det kan give problemer, at målene for afkodning og forståelse *opfattes* som uafhængige. I den testprocedure, der er beskrevet i manualen til den anvendte test (fx NARA), afbrydes testen på baggrund af et vist antal fejllæsninger i en tekst³. Spooner et al.'s bekymring var bl.a., at elever med svage afkodningsfærdigheder dermed ikke får lejlighed til at demonstrere deres *forståelse* fuldt ud. Men også uden anvendelse af

³ Cain og Oakhill (2006) beskriver, hvordan de i afgrænsningen af dårlige forståere tager højde for det antal af tekster, som eleverne når at læse, før stopkriteriet (antal fejllæste ord i en tekst) nås. Dette gøres netop for at imødegå uhensigtsmæssige konsekvenser af testproceduren.

stopkriterier synes et afkodningsmål baseret på læsning af ord i kontekst at have nogle ulemper i sammenligning med afkodning af isolerede ord. Fx kan den meningsfulde kontekst, som ordene indgår i (tekst, afsnit eller blot en sætning), potentielt hjælpe den lidt usikre afkoder til et godt resultat, fordi vedkommende kan slutte sig til, hvilke ord der forekommer sandsynlige i netop denne meningskontekst. Deltageren, der derimod har vanskeligheder med teksten på grund af dårlig sprogforståelse, har til gengæld ikke samme støtte i konteksten (Nation & Snowling, 1998). For det første kan kvaliteten af kontekstens støtte forventes at være stærkt forringet for denne læser, fordi forståelsen af konteksten netop er mangelfuld eller fejlagtig. For det andet kan det være, at vedkommende med dårlig sprogforståelse har dårlige forudsætninger for at *udnytte* konteksten på grund af begrænsede inferensfærdigheder. Et tredje problem kan være, at læseren, der bruger mange mentale ressourcer på at forstå en tekst, umiddelbart har færre ressourcer til rådighed for afkodningen ved eventuelle ukendte eller svære ord i teksten. På den måde kan dårlig forståelse lede til fejl i afkodningen. Selv hvis man udleder et afkodningsmål fra andre tekster end det, som giver anledning til en vurdering af forståelsen, kan graden, hvormed en læser forstår den tekst, der læses, stadig forventes at påvirke afkodningsscoren. Der er altså en del usikkerheder forbundet med at anvende læsning af ord i kontekst som afkodningsmål i afgrænsningen af dårlige forståere.

Afkodning af nonord eller rigtige ord

Afkodning er som nævnt alle de processer, der leder frem til genkendelse af et skrevet ord. Afkodning undersøges ofte enten med lister med nonord, som deltageren skal læse højt (ved en fonologisk strategi) (fx Nation & Snowling, 1999) eller lister med rigtige ord, som deltageren skal læse højt (muligt ved umiddelbar genkendelse af den specifikke bogstavfølge, hvis læseren har set ordet før) (fx Verhoeven, 2013). I afgrænsningen af dårlige forståere er det således muligt at afdække afkodningsfærdigheder med enten den ene eller den anden slags test – eller med en kombination.

Catts et al. (2006) anvendte fx et afkodningsmål sammensat af resultater på en test med højtlæsning af isolerede ord og en test med højtlæsning af nonord. Valget af et sammensat mål blev dog ikke begrundet.

Keenan et al. (2014) afprøvede og sammenlignede afgrænsninger af dårlige forståere, hvor en diskrepans med læseforståelse enten blev fundet ud fra et mål for afkodning af nonord (to test) eller et mål for afkodning af rigtige ord (to test). En afgrænsning, hvor diskrepans med læseforståelse blev fundet ud fra afkodning af nonord, resulterede i, at flere deltagere blev identificeret som dårlige forståere, end når diskrepansen med læseforståelse blev fundet ud fra afkodning af rigtige ord, trods det at grænseværdi for læseforståelse og kriteriet for størrelsen af diskrepans med afkodning var det samme. Rønberg og Petersen (2015a) fandt en lignende reduktion i antallet af dårlige forståere, når de tilføjede et kriterium for ortografisk kodning (estimat for afkodning af rigtige ord) til en oprindelig afgrænsning af dårlige forståere, hvor et mål for fonologisk kodning (estimat for afkodning af nonord) var blevet anvendt. Hvis man vælger at afgrænse dårlige forståere ved hjælp af afkodning af nonord frem for rigtige ord, kan man altså forvente en større gruppe af dårlige forståere. Men det store spørgsmål er imidlertid, hvordan dårlige forståere afgrænset med det ene mål frem for det andet *adskiller* sig. Et bud er, at de bl.a. vil adskille sig på ordforråd.

Læsning af rigtige ord stiller fx et krav til ordforråd, som ikke er til stede ved læsning af nonord – der netop er karakteriseret ved ikke at have en betydning. Korrelationen mellem afkodning af nonord og ordforråd er generelt svagere end korrelationen mellem afkodning af rigtige ord og ordforråd – særligt ved læsning af uregelmæssigt stavede ord (Ricketts et al., 2007; Ouellette & Beers, 2010). Når man afgrænser dårlige forståere med mål for afkodning af nonord, synes der altså at være større sandsynlighed for at få flere elever med et begrænset ordforråd med i gruppen af dårlige forståere.

Hvis man udelukkende anvender test af læsning af nonord i afgrænsningen af dårlige forståere, må man dog også forvente, at en del af de afgrænsede deltageres vanskeligheder med læseforståelse kan bunde i, at deltagerne endnu ikke har etableret præcise ortografiske repræsentationer, som er nødvendige for præcis og flydende afkodning. Det kan en undersøgelse af Nation og Snowling (1998) ses som et eksempel på. Nation og Snowling anvendte udelukkende læsning af nonord som afkodningsmål i deres afgrænsning af dårlige forståere. Alle de afgrænsede dårlige forståere faldt desuden inden for det, Nation og Snowling kaldte et alderssvarende område på et andet mål for afkodning, nemlig højtlesning af tekster, men gruppen var ikke helt matchet med kontrolgruppen på dette mål, sådan som det var tilfældet for læsning af nonord. Faktisk var der en marginal signifikant forskel på de to grupper, der svarede til en halv standardafvigelse på dette andet afkodningsmål, der involverede læsning af rigtige ord. Nation og Snowling fandt på denne baggrund ikke overraskende, at gruppen med de laveste læseforståelsesscorer, det vil sige de afgrænsede dårlige forståere, var dårligere end en kontrolgruppe af normale læsere til at læse uregelmæssigt stavede og lavfrekvente ord. Da hensigten med afgrænsningen af dårlige forståere netop er at finde frem til deltagere, hvis vanskeligheder i læseforståelse *ikke* kan forklares af dårlige afkodningsfærdigheder, kan det være et argument for at sammensætte afkodningsmålet i afgrænsningen af *både* test med nonord *og* rigtige ord, som det fx sås hos Catts et al. (2006). Herved får man et mere fuldgældigt udtryk for elevernes samlede afkodningsfærdigheder.

Præcision eller effektivitet i afkodning

Betydningen af *effektive* afkodningsfærdigheder for læseforståelse synes at have været overset i mange hidtidige afgrænsninger af dårlige forståere, hvor afkodningsscoren ofte har været baseret på en optælling af korrekt læste ord (præcision) og ikke har inddraget hastighed (fx Nation & Snowling, 1998; Cain & Oakhill, 1999). Tong et al. (2011) inddrager dog et effektivitetsmål for afkodningsfærdigheder, og der synes at være gode grunde til at kombinere information om afkodningens præcision med afkodningens hastighed. Afkodningsfærdigheder, der endnu ikke er automatiserede, kan nemlig kræve ekstra mentale ressourcer under tekstlæsning. Disse ressourcer kan så ikke bruges på forståelsen af teksten (Perfetti, 1985), og konsekvensen kan være et dårligt resultat på en læseforståelsestest. Det er altså relevant at sikre, at det ikke er manglende effektivitet i afkodningen, der afstedkommer, at elever afgrænses som dårlige forståere. Hvis man supplerer information om præcision i afkodning med afkodningshastighed i afgrænsningen af dårlige forståere som hos Tong et al. (2011) kan det give en sikkerhed for, at de afgrænsede dårlige forståere besidder automatiserede afkodningsfærdigheder og altså ikke først og fremmest er uøvede

læsere, der bruger mange ressourcer på afkodningen af de enkelte ord og derfor opnår lave læseforståelsesscorer.

Absolut eller relativ dårlig læseforståelse

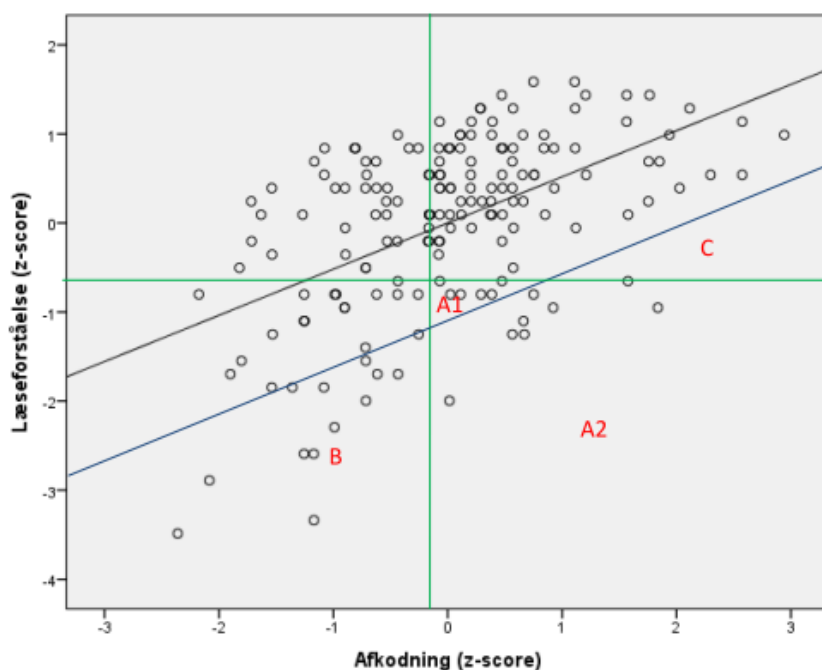
Der er ingen naturgiven grænse for, hvornår man har specifikke forståelsesvanskeligheder. Det angives ofte, at ca. 10% af en årgang kan betegnes som dårlige forståere, men eftersom både læseforståelse og afkodningsfærdigheder er færdigheder, der er normalfordelt, beror sådanne opgørelser af prævalens i bund og grund på en arbitrær fastsættelse af forskellige grænseværdier. Undersøgelser, der finder forskellig prævalens af specifikke forståelsesvanskeligheder, benytter da også ofte forskellige grænseværdier for læseforståelse og afkodning (se fx Spencer et al. (2014), Catts et al. (2006) og Ehrlich & Remond (1997) for forskellige grænseværdier for læseforståelsesscorer, der svarer til hhv. 5-, 25- og 30-percentilen).

En anden og mere væsentlig forskel på grænsedragningerne er, om grænser for færdighedsniveauerne i læseforståelse og afkodning er absolutte (*cut-off definition*), eller om udgangspunktet er at sikre et vist misforhold (en diskrepans) mellem færdighedsniveauerne i læseforståelse og afkodning, altså at sikre at elevernes læseforståelse er dårlig relativt til deres afkodningsfærdigheder (*diskrepansdefinition*).

Jeg giver i dette afsnit eksempler på undersøgelser, der har anvendt enten absolutte grænser (simple grænseværdier) eller diskrepansgrænser og også undersøgelser, hvori man har anvendt kombinationer, og jeg konkretiserer konsekvenser af de forskellige grænsedragninger ved hjælp af et datasæt. Jeg vil desuden argumentere for, at en hensigtsmæssig afgrænsning af dårlige forståere, der er i overensstemmelse med den generelle definition af specifikke vanskeligheder med sprogforståelse i læsning (vanskeligheder med tekstforståelse, som ikke skyldes dårlig afkodning), indeholder elementer fra begge metoder til grænsedragning.

Konsekvenser af forskellige grænsedragninger

Figur 1 viser resultater for en gruppe 6.-klasseelever, der har deltaget i læseforståelses- og afkodningstest. Disse data behandles i kapitel 3, hvor der også er en redegørelse for de anvendte mål. Datapunkterne i figur 1 vidner om, at der er en sammenhæng mellem afkodningsfærdigheder og læseforståelse ($r = 0,52$, $p < 0,01$). Højere afkodningsscorer følges generelt af højere læseforståelsesscorer. Elevernes scorer er i figuren standardiseret (omregnet til z-scorer).



Figur 1. Elevgrupper ved forskellige afgrænsninger af dårlige forståere.

I figur 1 kan man se grænseværdier fra Catts et al. (2006) og Tong et al. (2011) anvendt på det aktuelle udtræk af 6.-klasseelever. Catts et al. (2006) anvendte en *cut-off definition* baseret på simple grænseværdier for læseforståelse og afkodning. De definerede dårlige forståere som deltagere med læseforståelsesscorer under 25-percentilen og afkodningsfærdigheder over 40-percentilen (grønne grænser i figur 1).

Tong et al. (2011) anvendte derimod en *diskrepansdefinition*. Tong et al. sammensatte et prædiktormål på basis af afkodningsscorer (og deltagernes alder og nonverbale færdigheder). De brugte dernæst lineær regression til at finde frem til deltagere, der havde en diskrepans mellem afkodning og læseforståelse, altså specifikke forståelsesvanskeligheder. Dårlige forståere blev afgrænset som deltagere, der afveg med en bestemt afstand fra regressionslinjen (de læseforståelsesscorer som man kunne forvente på baggrund af den lineære sammenhæng mellem prædiktormålet og læseforståelse). Nærmere bestemt blev dårlige forståere afgrænset som deltagere under den nedre grænse for 80%-konfidensintervallet omkring regressionslinjen.

I figur 1 er regressionslinjen for læseforståelse ud fra afkodningsfærdigheder markeret (den sorte skrå linje). Regressionslinjen viser de forventede læseforståelsesscorer for deltagere med forskellige afkodningsscorer. Datapunkter, der ligger på eller tæt ved denne linje, repræsenterer altså deltagere med læseforståelse på et niveau, der matcher deres afkodningsfærdigheder – uanset om de generelt kan siges at være svage læsere (lave scorer på begge akser), gennemsnitlige læsere (scorer tæt ved gennemsnittet 0 på begge akser) eller stærke læsere (høje scorer på begge akser). Under regressionslinjen finder man deltagere med læseforståelsesscorer, der i den ene eller anden grad er lavere, end man kunne forvente ud fra deltagernes afkodningsfærdigheder. Dårlige forståere kan derfor defineres med udgangspunkt i datapunkternes afvigelse fra regressionslinjen. Under og

parallelt med regressionslinjen er der indsat den samme diskrepansgrænse, som Tong et al. (2011) anvendte (blå grænse)⁴.

Med *cut-off definitionen* afgrænses elever i felterne A1 og A2 som dårlige forståere. Alle disse elever har vanskeligheder med tekstforståelse i den udstrækning, at de er blandt den dårligste fjerdedel i læseforståelse. For eleverne tæt ved både læseforståelses- og afkodningsgrænsen (svarende til elever i feltet A1) kan man dog sætte spørgsmålstejn ved, om forståelsesvanskelighederne i læsning i alle tilfælde er *specifikke*; for én af eleverne svarer diskrepansen mellem afkodning og læseforståelse fx til blot lidt mere end 0,5 standardafvigelse. For elever i feltet A1 kan forskellen mellem læseforståelse og afkodning altså synes beskedent. Læseforståelse og afkodning forekommer at være på et nogenlunde sammenligneligt niveau.

Deltagerne i feltet A2 har til gengæld væsentligt dårligere læseforståelse end afkodning. Her er der med rette tale om problemer med tekstforståelse, der ikke kan forklares med dårlig afkodning. Dette synes imidlertid også at være tilfældet for deltagere i feltet B. De har væsentlig dårligere læseforståelse end afkodning. Men disse deltagere afgrænses ikke som dårlige forståere ved brug af de simple grænseværdier, der her er sat.

Med *diskrepansdefinitionen* bliver elever i feltet B derimod afgrænset som dårlige forståere. Det samme gør elever i A2 og C. For alle disse elever gælder, at der er en væsentlig diskrepans mellem læseforståelse og afkodning – i overensstemmelse med den generelle definition af dårlige forståere ud fra den enkle læsemodel. Elever i A1 bliver dermed *ikke* afgrænset som dårlige forståere med diskrepansdefinitionen.

I den aktuelle elevgruppe er der ingen datapunkter i feltet C, men elever i dette felt ville med diskrepansdefinitionen som nævnt også være blevet afgrænset som dårlige forståere. Som det ses af figur 1, ville elever i dette felt imidlertid kunne have gennemsnitlige læseforståelsesscorer. Afgrænsning af elever i feltet C er dermed *ikke* i overensstemmelse med definitionen af specifikke forståelsesvanskeligheder i læsning; der vil ikke være tale om *vanskeligheder* med tekstforståelse i forhold til elevernes jævnaldrende.

Keenan et al. (2014) anvendte, hvad man kan kalde en *kombinationsdefinition*. De opererede med en øvre grænseværdi for læseforståelse og kombinerede den med et krav om en minimumsdiskrepans mellem læseforståelse og afkodningsfærdigheder for hver enkelt deltager. Kriteriet var, at deltagerens læseforståelsesscore lå mindst 1 standardafvigelse under deltagerens afkodningsscore. Keenan et al. sikrede altså, at alle deltagere i gruppen af dårlige forståere havde vanskeligheder med tekstforståelse i forhold til jævnaldrende, og at der samtidig tilsyneladende var tale om vanskeligheder, som ikke kunne forklares med afkodningsfærdigheder.

Cain og Oakhill (1999) anvendte, hvad man kan kalde en *fuldgyldig kombinationsdefinition*. Definitionen var til dels en cut-off definition, fordi en øvre grænseværdi for læseforståelse og en nedre grænseværdi blev sat for afkodning. I stedet for at vurdere niveau i læseforståelse og afkodning relativt til et begrænset udvalg af elever benyttede man et normeret testmateriale, der gjorde det muligt at identificere deltagere med læseforståelsesscorer, der lå under det forventede

⁴ Man kan dog hævde, at Tong et al. indirekte opererede med en nedre grænseværdi for afkodningsfærdigheder, fordi de kun udvalgte dårlige forståere til videre testning blandt de deltagere, der kunne matches på prædiktorscorer med gode forståere (deltagere med læseforståelsesscorer over den øvre grænse for 80%-konfidensintervallet omkring regressionslinjen), og dette var ikke muligt for dårlige forståere med prædiktorscorer under -1 standardafvigelse.

niveau for deltagere på deres alder (opgjort i måneder), men desuden havde alderssvarende ('age-appropriate') afkodningsfærdigheder⁵. Men definitionen var til dels også en diskrepansdefinition i lighed med definitionen hos Tong et al. (2011). Diskrepansen mellem færdighedsniveauerne blev blot fastlagt ved en anden fremgangsmåde end lineær regression, nemlig ud fra det normerede testmateriale: Læseforståelsesscoren for hver enkelt dårlig forståer var på et niveau, der svarede til eller lå under et niveau på 6 måneder under elevens afkodningsniveau.

Da den øvre grænse for læseforståelse hos Cain og Oakhill (1999) imidlertid var lig med *gennemsnittet* for aldersgruppen, var der dog – som i undersøgelsen af Tong et al. (2011) – det væsentlige problem ved denne afgrænsning, at elever med gennemsnitlig læseforståelse for alderstrinnet kunne blive afgrænset som 'dårlige forståere' – vel at mærke hvis afkodningsfærdighederne var mindst seks måneder over den ikke nærmere specificerede nedre grænse for afkodningsfærdigheder.

Grænsedragning i overensstemmelse med den generelle definition af dårlige forståere

Hverken en grænsedragning med simple grænseværdier eller en grænsedragning med et rent fokus på diskrepans mellem afkodning og læseforståelse synes at kunne afgrænse dårlige forståere helt i overensstemmelse med den generelle definition af dårlige forståere, der kan udledes af den enkle læsemodel: 'Vanskeligheder med tekstforståelse der ikke skyldes dårlig afkodning'.

En kombination af de to tilgange synes at være det bedste bud på en grænsedragning i overensstemmelse med den generelle definition. For det første må en øvre simpel grænseværdi for læseforståelse sættes for at sikre, at de udpegede dårligere forståere reelt har lav læseforståelse sammenlignet med deres jævnaldrende. Det er oplagt, at denne grænseværdi som minimum må ligge et stykke under gennemsnittet. Elever med god eller gennemsnitlig læseforståelse bør naturligvis ikke kunne afgrænses som dårlige forståere. For at sikre at den dårlige tekstforståelse ikke kan forklares ud fra afkodningsniveauet, kan der dernæst etableres en diskrepansgrænse. En diskrepansgrænse kan fx etableres ved lineær regression som hos Tong et al. (2011).

Tilbage står en overvejelse om, hvorvidt elever skal kunne blive identificeret som dårlige forståere, selv om de *også* har alvorlige vanskeligheder med afkodning. I princippet kan deres forståelse jo stadig være ringere, end man skulle forvente ud fra deres afkodning. Elever med både afkodnings- og forståelsesvanskeligheder beskrives ifølge den enkle læsemodel som elever med *sammensatte vanskeligheder*. En nedre grænseværdi for afkodning kan derfor føjes til afgrænsningen af dårlige forståere. En sådan grænseværdi kan fx ligge ved -1 standardafvigelse. En score på -1 standardafvigelse svarer til en standardscore på 85, som ofte bliver anført som en grænse for vanskeligheder på andre færdighedsmål.

⁵ Grænseværdi(er) for 'alderssvarende afkodningsfærdigheder' i definitionen hos Cain og Oakhill (1999) blev ikke nærmere præciseret, men alderssvarende afkodningsfærdigheder må forventes at dække over et betydeligt spænd – standardafvigelsen i den afgrænsede gruppe af dårlige forståere var på 7 år og 4 måneder (til sammenligning var standardafvigelsen for læseforståelse 3 år og 9 måneder).

Det er vigtigt at understrege, at afgrænsning af dårlige forståere fra deltagere, der *tilsyneladende* har mere sammensatte vanskeligheder, på ingen måde betyder, at elevers sammensatte vanskeligheder ikke udgør reelle problemer eller er vigtige at tage sig af. Tværtimod, disse vanskeligheder er vidtrækkende og kan forventes at kræve en omfattende og langvarig intervention. Men det er ikke rimeligt at karakterisere disse deltageres vanskeligheder som et *særligt* sprogforståelsesproblem. Det er fx muligt, at sammenhængen mellem læseforståelse og afkodning ikke er lineær i den nedre del af afkodningsskalaen. Der kan måske være en nedre grænse for afkodning, hvorunder forståelsen bliver markant påvirket af afkodningsproblemerne. Desværre er der ikke data til at belyse denne mulighed i den nærværende undersøgelse.

Sammenfatning

Ud fra den enkle læsemodel kan man udlede en generel definition af dårlige forståere: 'Vanskeligheder med tekstforståelse der ikke skyldes dårlig afkodning'. Spørgsmålet melder sig nu, hvordan man kommer fra denne generelle definition til en faktisk afgrænsning af en gruppe dårlige forståere. I forskningen er mange fremgangsmåder blevet anvendt. For at blive i stand til at udlede konklusioner om mulige årsager til dårlig sprogforståelse i læsning, afgrænser man som oftest så homogene grupper af dårlige forståere (og gode forståere), som det er muligt. Det kan indebære, at man ikke inkluderer elever med kendte sproglige og andre kognitive vanskeligheder, og det kan indebære, at man matcher sine grupper på udvalgte færdigheder – fx at man matcher sine grupper på ordforråd, hvis man ønsker at isolere og undersøge effekten af variation i inferensfærdigheder. Hvor sådanne fremgangsmåder er hensigtsmæssige i forskning, gælder der dog andre hensyn i praksis. Her er formålet ikke at udlede konklusioner om årsager til dårlig sprogforståelse, men at blive opmærksom på så mange som muligt af de elever, der har vanskeligheder med sprogforståelse i læsning. I en afgrænsning, der skal kunne anvendes i praksis, synes det derfor ikke relevant at inddrage kontrolvariable som fx ordforråd. Til gengæld synes det oplagt, at man som et indledende trin i afgrænsningen anvender baggrundsoplysninger om eksempelvis sproglige vanskeligheder til at identificere elever, der med stor sandsynlighed kan have eller få læseforståelsesvanskeligheder. Herefter kan man ved hjælp af resultater på læseforståelses- og afkodningstest afgrænse dårlige forståere, hvis vanskeligheder *ikke* er indikeret i baggrundsoplysningerne.

Testmaterialerne, der anvendes i afgrænsningen, kan have betydning for, hvilke elever der afgrænses som dårlige forståere – og om det overhovedet er muligt at afgrænse dårlige forståere. Det synes for det første at være afgørende, at læseforståelsestesten er følsom for forskelle i sprogforståelse, og at den stiller krav til de *forskellige* færdigheder, der har vist sig at have betydning for sprogforståelse i læsning – som fx ordforråd og inferensfærdigheder. Hvilke(n) test man anvender til at undersøge afkodningsfærdigheder synes også at kunne have betydning. Ved test af afkodning af ord i kontekst risikerer man en påvirkning fra enten god eller dårlig forståelse af det læste til afkodning af de enkelte ord. Ved afkodning af henholdsvis isolerede nonord eller rigtige ord synes det muligt, at man kun delvist får indblik i de samlede afkodningsfærdigheder. I nogle

undersøgelser af dårlige forståere kombinerer man da også mål for afkodning af nonord og afkodning af rigtige ord i afgrænsningen. I nyere undersøgelser er man desuden begyndt at inddrage ikke bare mål for præcision i afkodning, men også hastighed. Ved at kombinere mål for effektivitet i afkodning af både nonord og rigtige ord, synes det muligt at sikre, at de afgrænsede dårlige forståere har automatiserede afkodningsfærdigheder – og at det dermed ikke er *usikre* afkodningsfærdigheder, der giver sig udslag i elevernes lave læseforståelsesscorer.

Hvorvidt den konkrete grænsedragning hviler på simple grænseværdier eller en diskrepansgrænse, kan også have konsekvenser for, hvilke elever der bliver afgrænset. Ved udelukkende brug af simple grænseværdier sikrer man ikke nødvendigvis, at læseforståelsen hos alle elever er markant ringere end deres afkodningsfærdigheder. Ved udelukkende brug af en diskrepansgrænse sikrer man dette, men man risikerer til gengæld at afgrænse elever med god læseforståelse og elever med afkodningsvanskeligheder som dårlige forståere. En kombination af de to metoder til grænsedragning synes derfor at være den, der fører til en afgrænsning, der er i bedst overensstemmelse med den generelle definition af dårlige forståere. En øvre simpel grænseværdi for læseforståelse kan sikre, at de afgrænsede dårlige forståere reelt har dårlig læseforståelse i sammenligning med deres jævnaldrende. En diskrepansgrænse sikrer, at den dårlige læseforståelse ikke udelukkende er en konsekvens af afkodningsniveauet. En nedre simpel grænseværdi for afkodningsfærdigheder kan desuden afgrænse dårlige forståere fra elever med sammensatte vanskeligheder i læsning.

Kapitel 3: Stabilitet af profilen 'dårlig forstår' på tværs af tid og testbetingelser

Indledning

I kapitel 2 blev det bekræftet, hvordan man meningsfuldt kan afgrænse dårlige forståere. Der er en enkel teoretisk ramme (den enkle læsemodel), som forklarer, hvad dårlige forståere i princippet har vanskeligheder med. Kapitel 2 viste bl.a. også, hvordan den praktiske identifikation af dårlige forståere til en vis grad kan falde forskelligt ud, alt efter hvordan man operationaliserer misforholdet mellem læseforståelse og afkodning, og det blev pointeret, at indholdet i den valgte læseforståelsestest (umiddelbare krav til afkodning og de forskellige komponenter i sprogforståelse), samt den eller de valgte afkodningstest også kan forventes at have indflydelse på hvilke læsere, der afgrænses som dårlige forståere.

Et andet spørgsmål er *stabiliteten* af profilen 'dårlig forstår'. Hvis det skal give praktisk mening at betegne en given læser som 'dårlig forstår', så skal profilen være forholdsvis robust og stabil. Det skal være en fast egenskab ved læseren, at forståelsen er klart ringere end afkodningen. Læseren må ikke være 'dårlig forstår' den ene dag og fx normal læser den næste dag – eller måneden efter. Antagelsen om en stabilitet for læsevanskeligheder præger både forskning og praksis. Baseret på de færdigheder og eventuelle vanskeligheder, der aktuelt kan observeres hos personer på en test, der afvikles under et sæt af særlige betingelser, forsøger man i forskning at forudsige personernes læseforståelse over tid og finde korrelater til dårlig læseforståelse, og i praksis ønsker man at intervenere med det formål at forebygge eller afhjælpe vanskelighederne.

Færdigheder udvikles (heldigvis) over tid, og man må derfor forvente en vis ustabilitet, også for profilen 'dårlig forstår', over længere perioder. Det samme kan i princippet også være tilfældet over lidt kortere perioder på fx nogle måneder. Sådanne kortere perioder mellem testgange er hyppige i både praksis og forskning. I praksis udtages en elev måske til detaljeret afdækning af færdigheder på baggrund af en enkelt overordnet test. De to eller flere testdage vil ofte være forskudt i tid. I forskning hvor man indhenter mange forskellige testresultater fra de deltagende elever, er det af praktiske tids- og energihensyn også almindeligt at sprede den omfattende testning ud over en periode og at antage, at elevens færdigheder er uforandrede over denne periode, og at resultaterne, der indhentes fra forskellige testgange, derfor kan sammenlignes direkte uden at tage højde for udvikling. Men hvor stor stabilitet kan man egentlig forvente over sådanne kortere perioder?

Det er også almindeligt at antage en høj stabilitet for den enkelte elevs færdigheder på tværs af de mange forskellige læseforståelsestest, der er i anvendelse. Men eftersom forskelle i testen indhold, som det blev beskrevet i kapitel 2, må forventes at bidrage til, at lidt forskellige elevgrupper identificeres som dårlige forståere, er det givet, at disse forskelle også kan føre til en

vis grad af ustabilitet for profilen på tværs af to forskellige test. Det er imidlertid ikke kun indholdet, der varierer på tværs af test, også testenes formater og de betingelser, som deltagerne skal forstå og demonstrere deres forståelse under, varierer. Nogle test afvikles fx som gruppetest, hvor andre afvikles som individuelle test, og forskellige typiske betingelser kendetegner gruppetest og individuelle test, men det er ikke almindeligt at anføre variation i testbetingelser som forklaring på forskelle i testresultater på tværs af test. Tværtimod bruges gruppetest og individuelle test ofte i kombination. Man bliver måske opmærksom på en elev med vanskeligheder i en screening med gruppebaserede test, og screeningsresultatet følges så op med individuel testning, når vanskelighederne skal undersøges nærmere. Kan de forskellige måder, man afvikler læseforståelsestest på, have en effekt på stabiliteten af profilen 'dårlig forstår'?

Hvilken indflydelse, det kan have på stabiliteten af profilen 'dårlig forstår', at man kigger på testresultater over tid og samtidig tillader variationer i testbetingelserne, som er typiske i både praksis og forskning, er emnet i kapitel 3. Stabiliteten vil blive undersøgt på tværs af to tidspunkter og med test, der er indholdsmæssigt parallelle, men afvikles som henholdsvis gruppetest og individuel test med hver deres sæt af karakteristika. Men først præsenteres undersøgelser, der kan give indtryk af stabiliteten på tværs af tid målt med samme test. Her er der en del eksisterende forskning. Dernæst præsenteres forskningsresultater, der kan give en idé om stabiliteten på tværs af test, og om hvordan forskelle i såvel indhold, format og testbetingelse muligvis kan påvirke stabiliteten af profilen 'dårlig forstår'.

Stabilitet over tid

Deltagernes alder og klassetrin er en afgørende faktor for, hvor mange deltagere med specifikke forståelsesvanskeligheder det er sandsynligt at finde. Som beskrevet i kapitel 1 er variationen i tekstforståelse på de første klassetrin stort set udelukkende bestemt af variationen i afkodning. Derfor er det meget svært om ikke umuligt at påpege specifikke forståelsesvanskeligheder i læsning, før børn er i stand til at læse ord nogenlunde præcist og flydende (Nation et al., 2010).

Efterhånden som eleverne kommer i gang med at afkode, og teksternes kompleksitet og krav til anvendelse af læsefærdighederne øges, bliver forståelsesvanskeligheder synlige. Fjerde klasse angives ofte som et kritisk tidspunkt/skel, hvor formålet med læsning for alvor går fra at være 'at lære at læse' til at være 'at læse for at lære', dvs. forståelse og videnstilegnelse, og hvor nogle elever, der ikke tidligere har haft problemer med læsning, begynder at score lavt på test af læseforståelse. Dette fænomen har Chall (1983) betegnet som et skred i læseforståelse i 4. klasse (the fourth grade slump). I overensstemmelse med et skred i læseforståelse omkring 4. klassetrin viser retrospektive undersøgelser, at deltagere med specifikke læseforståelsesvanskeligheder i 4. klasse kan have haft gennemsnitlige scorer på læseforståelsestest på tidligere klassetrin (Elwér, Gustafson, Byrne, Olson, Keenan, Samuelsson, 2015), og undersøgelser, der retrospektivt har fulgt udviklingen hos grupper af dårlige forståere og gode forståere, indikerer, at disse grupper ikke adskilte sig overbevisende på et tidligere tidspunkt (Nation et al, 2010; Elwér et al., 2015). Nation et al. (2010) identificerede fx 15 dårlige forståere og 15 gode/gennemsnitlige forståere i 8-årsalderen. Da børnene også var blevet testet i 6- og 7-årsalderen, kunne man undersøge, i hvilken

grad grupperne var forskellige på disse tidspunkter. Da børnene var 7 år, var de to grupper også signifikant forskellige i læseforståelse. Da de var 6 år, var det ikke alle deltagere, der kunne afkode godt nok til at deltage i læseforståelsestesten, men der var en ikke-signifikant forskel i læseforståelse ($p = 0,08$) for dem, der kunne (10 dårlige forståere og 12 gode forståere). Og selv om forskellen på gruppernes middelværdier var stor, både da børnene var 6 og 7 år ($d = 0,8$ og $d = 0,95$), var den væsentligt mindre end i 8-årsalderen ($d = 3,81$). Lignende forskelle i effektstørrelser mellem dårlige og gode forståere på tidligere klassetrin ses hos Elwér et al. (2015). At forskellene i læseforståelse for grupperne af dårlige og gode forståere var mindre på tidligere klassetrin, er i overensstemmelse med den simple læsemodel, der siger, at sammenhængen mellem læseforståelse og sprogforståelse bliver stærkere, jo ældre børnene er, men en del af forklaringen på, at forskellene blev mindre, da man fulgte eleverne tilbage i tid, må også skyldes regressionseffekt. Regressionseffekt må altid medregnes, når man udvælger deltagere med ekstreme scorere (det vil sige meget høje eller lave scorere) og følger dem over tid (enten bagud eller fremad). Den betegner en generel tendens til, at deltagere med meget lave og meget høje scorere på en test højst sandsynligt vil score lidt tættere på gennemsnittet næste gang, hvis man følger dem over tid. Det sker, fordi sandsynligheden for, at deltagere med meget lave scorere havde en særlig dårlig dag, den dag de blev udvalgt som *meget* dårlige, er relativ høj, og omvendt er sandsynligheden relativ høj for, at deltagere med meget høje scorere havde en særlig god dag, den dag de blev udvalgt som *meget* gode (Rost, 2007; Elbro & Poulsen, 2015).

Catts et al. (2006) anvendte flere sammenlagte mål for læseforståelse, hvorved de begrænsede betydningen af de enkelte tests pålidelighed og regressionseffekten. Resultaterne viste samme tendens, som hos Nation et al. (2010). Dårlige forståere blev identificeret på 8. klassetrin, og gruppen var signifikant forskellig fra en gruppe typiske læsere i læseforståelse. De to grupper var også signifikant forskellige i læseforståelse i 4. og 2. klasse, men forskellen var mindre i 2. og 4. klasse ($d = 0,97$ og $d = 1,66$) end i 8. klasse ($d = 3,72$).

De nævnte undersøgelser indikerer altså, at der er en vis stabilitet, men også en individuel udvikling, der kan betyde, at nogle børn går fra ikke at have vanskeligheder på ét tidspunkt til at have vanskeligheder på et andet tidspunkt – eller omvendt. Hverken Nation et al. (2010) eller Catts et al. (2006) opgør dog, hvor mange af de identificerede dårlige forståere, eller for den sags skyld gode forståere eller typiske læsere, der rent faktisk ville falde i samme kategori på de tidligere undersøgelsestidspunkter, hvilket kunne give et konkret mål for stabiliteten retrospektivt. Stabiliteten fremad i tid synes dog også at være endnu mere interessant: I hvor høj grad kan vi regne med, at deltagere identificeret som dårlige forståere på ét tidspunkt vil blive gen-identificeret som dårlige forståere på senere tidspunkter? Hvad viser undersøgelser, der følger dårlige og gode forståere fremad i tid?

Catts, Compton, Tomblin & Bridges (2012) undersøgte både, hvornår forskellige typer læsevanskeligheder opstår og stabiliteten af dem over tid (2., 4., 8. og 10. klassetrin). Catts et al. (2012) anvendte samme oprindelige elevgruppe som Catts et al. (2006), og der er derfor et delvist overlap i deltagere og testmaterialer mellem de to undersøgelser. For at identificere dårlige forståere anvendte man mål for afkodning og læseforståelse, der var baseret på de forskellige afkodnings- og læseforståelsestest i testbatteriet. Elever med tidlige sproglige og/eller kognitive vanskeligheder var overrepræsenteret i elevgruppen, så for at vurdere sandsynligheden for de forskellige typer af

læsevanskeligheder og deres stabilitet på tværs af klassetrin i populationen, korrigerede man for overrepræsentationen af børn med tidlige vanskeligheder ved at anvende vægtede scorer. Catts et al. (2012) fandt, at afkodningsvanskeligheder og sammensatte læsevanskeligheder var repræsenteret i 2. klasse, men at der ikke var læsere med specifikke læseforståelsesvanskeligheder på dette tidlige tidspunkt. Dårlige forståere blev først identificeret i 4. klasse eller senere. Når en læser havde skiftet kategori, fra fx 'typisk læser' til 'dårlig forstår', var der generelt tendens til, at læseren blev i denne kategori fremover. For dårlige forståere fandt Catts et al. en stabilitet på 75% mellem 4. og 8. klassetrin. Det vil sige, at 75% af dårlige forståere i 4. klasse også blev identificeret som dårlige forståere i 8. klasse. Syv procent flyttede til gruppen uden vanskeligheder, og 19% flyttede til gruppen med sammensatte vanskeligheder i samme tidsrum. Mellem 8. og 10. klasse var stabiliteten for dårlige forståere 95% (5% flyttede til gruppen med sammensatte vanskeligheder). Graden af stabilitet for dårlige forståere på tværs af klassetrin svarede dermed til den stabilitet, man fandt for specifikke afkodningsvanskeligheder (hhv. 73% og 91% på de nævnte to tidspunkter). Der var altså tale om en meget høj stabilitet.

Cain og Oakhill (2006) identificerede dårlige og gode forståere i alderen 7-8 år. Tre år senere gentestede de 17 dårlige forståere og 19 gode forståere med den samme læseforståelsestest. Ved gentestningen havde én af de dårlige forståere alderssvarende læseforståelse, og tre af de gode forståere havde nu læseforståelse under alderssvarende niveau. Stabiliteten af kendetegnet 'dårlig forstår' synes altså også her meget høj (94%). Stabiliteten af kendetegnet 'god forstår' var lidt lavere, men stadig høj (84%). Men disse mål for stabiliteten synes dog ved nærmere eftersyn at kunne være sat for højt. Ved gentestning anvendte Cain og Oakhill (2006) nemlig ikke de samme strenge kriterier for dårlige forståere og gode forståere som ved første testtidspunkt, og det sætter spørgsmålstegn ved den opgivne grad af stabilitet. Ved det første testtidspunkt var det et kriterium for dårlige forståere, at deres forståelsesalder lå mindst 12 måneder under deres kronologiske alder og også 12 måneder under deres afkodningsalder. Ved andet testtidspunkt kvalificerede en forståelsesscore til betegnelsen 'dårlig forstår', hvis den svarede til et niveau under kronologisk alder – uanset hvor lidt eller meget under⁶. Kriteriet for betegnelsen 'god forstår' synes også at variere mellem første og andet testtidspunkt. Ved det første testtidspunkt var det et kriterium for gode forståere, at deres læseforståelsesscore lå på eller over det niveau, der kunne forventes på baggrund af afkodningsfærdighederne. Ved andet testtidspunkt var der ikke angivet kriterier for betegnelsen 'god forstår'. Her synes deltagere, som oprindeligt var identificeret som gode forståere, at forblive i kategorien, såfremt de ikke havde læseforståelsesscorer under deres kronologiske alder. Det betyder, at det ikke kan formodes, at den høje stabilitet ville gælde, hvis man i stedet havde undersøgt stabiliteten af profilerne på tværs af test med de oprindelige (strengere) definitioner af dårlige og gode forståere ved gentestningen.

Cornoldi, De Beni og Pazzaglia (1996) fandt en vis stabilitet for en gruppe af 12 dårlige forståere, som blev identificeret da de var 11 år og fulgt gennem to år. Cornoldi et al. (1996)

⁶ Cain og Oakhill oplyste dog, at majoriteten af eleverne havde mindst alderssvarende afkodningsfærdigheder ved andet testtidspunkt, og at en stor del af deltagerne i gruppen ved andet testtidspunkt opnåede afkodningsscorer, der lå på et niveau mere end 12 måneder over deres alder. Det er derfor muligt, at størstedelen af de dårlige forståere ved andet testtidspunkt faktisk havde en forskel mellem afkodningsfærdigheder og læseforståelse der svarede til det ene af de to anvendte kriterier (en læseforståelsesalder 12 måneder under afkodningsalder) ved første testtidspunkt.

oplyste, at syv ud af de 12 dårlige forståere fortsat havde svære forståelsesvanskeligheder efter to år. Svære forståelsesvanskeligheder blev ved andet testtidspunkt defineret som en læseforståelsesscore på mindst 1,5 standardafvigelse under gennemsnittet for et normativt sample. Dermed synes stabiliteten umiddelbart at være 58%. Cornoldi et al. (1996) angav desuden, at vanskelighederne for yderligere tre deltagere kun var delvist afhjulpet idet deres læseforståelsesscorer ved det andet testtidspunkt lå omkring 1 standardafvigelse under gennemsnittet. I den oprindelige udvælgelse af dårlige forståere, da børnene var 11 år, var kriteriet for dårlig læseforståelse imidlertid -0,4 standardafvigelse og ikke -1,5 standardafvigelse. Det står tvivl om, hvordan den stabilitet, som Cornoldi et al. (1996) rapporterede, skal læses. Det synes desuden at kunne være en *anden* læseforståelsestest, der blev anvendt til at vurdere om deltagerne var dårlige forståere ved det andet testtidspunkt end ved det første undersøgelsestidspunkt. Dermed er den oplyste stabilitet muligvis en stabilitet, der ikke blot går på tværs af tid, men også på tværs af test. Der er altså flere ubekendte forbundet med Cornoldi og kollegers resultat.

Samlet set vidner undersøgelser, der har set på stabiliteten af profilen 'dårlig forstår' over tid, om en relativt høj stabilitet. Specifikke forståelsesvanskeligheder manifesterer sig på forskellige tidspunkter, ofte omkring 4. klasse eller senere, men når først en elev har vanskeligheder på en læseforståelsestest eller en kombination af læseforståelsestest, er det sandsynligt, at eleven vil have vanskeligheder på både kort og langt sigt, når tekstforståelse undersøges med det samme materiale/mål igen. Resultaterne fra langtidsundersøgelsen af Catts et al. (2012) tyder på, at stabiliteten bliver højere med tiden. Dermed er det også muligt, at variationen i stabilitet fra undersøgelse til undersøgelse til dels kan hænge sammen med, at der indgår elever fra forskellige aldersgrupper i undersøgelserne og/eller, at elever på samme alderstrin på tværs af undersøgelser ikke nødvendigvis har modtaget læseundervisning i samme omfang.

Stabilitet på tværs af test

Stabiliteten af profilen dårlig forstår på tværs af test er ikke undersøgt i samme grad som stabiliteten på tværs af tid med samme testmaterialer. Men en undersøgelse af Keenan et al. (2014) viser, at stabiliteten på tværs af læseforståelsestest kan være begrænset. Keenan et al. (2014) undersøgte stabiliteten af profilen 'dårlig forstår' i et sample med 1255 børn. Alle børn havde prøvet fire forskellige læseforståelsestest, og Keenan et al. (2014) lagde de udledte mål for læseforståelse sammen to og to og undersøgte stabiliteten på tværs af disse sammenlagte mål. Det ene sammenlagte mål var den gennemsnitlige score på to test, der i tidligere undersøgelser havde vist sig at være mest følsom for afkodningsfærdigheder, og det andet sammenlagte mål var den gennemsnitlige score på to test, der var mest følsom for sprogforståelse. Til at vurdere deltageres afkodning havde man både test af læsning af nonord og ord, men i stedet for at kombinere disse til én afkodningsscore, anvendte man hvert mål til at identificere elever med en diskrepans mellem afkodning og læseforståelse, hvorved man også kunne se på, om stabiliteten af dårlig forstår-profilen ændredes som funktion af, hvilket afkodningsmål der blev brugt i afgrænsningen. Keenan et al. (2014) undersøgte stabiliteten af profilen 'dårlig forstår' samlet for hele deltagergruppen,

men også for aldersopdelte grupper (en yngre gruppe med en gennemsnitsalder på 9,5 år og en ældre gruppe med en gennemsnitsalder på 13,9 år). Stabiliteten var generelt lidt højere for den ældre gruppe og generelt lidt højere, når læsning af nonord (i modsætning til læsning af ord) blev anvendt til at finde en diskrepans mellem afkodning og læseforståelse. Den højeste stabilitet for profilen dårlig forstår på tværs af de to sammenlagte mål for læseforståelse i den ældre gruppe var 25%. Den laveste stabilitet i den yngre gruppe var 7%. Foruden læseforståelsestest havde Keenan og kolleger også et sammenlagt mål for lytteforståelse med i undersøgelsen. Hvis dette mål blev anvendt i afgrænsningen af dårlige forståere, fandtes den højeste stabilitet af profilen mellem lytteforståelsesmålet og det mest sprogforståelsesfølsomme af de to sammenlagte læseforståelsesmål. Stabiliteten var da 41%. Den laveste stabilitet var igen i den yngre gruppe. Der var tale om stabiliteten på tværs af lytteforståelsesmålet og det mest afkodningsfølsomme læseforståelsesmål. Her var stabiliteten 6%.

Undersøgelsen af Keenan et al. (2014) viser altså, at stabiliteten på tværs af test i nogle tilfælde kan være meget begrænset. Spørgsmålet er nu, hvad den begrænsede stabilitet på tværs af test skyldes. I hvilken grad testene er enige med sig selv, det vil sige testens grad af pålidelighed, er en faktor, der kan bidrage til uoverensstemmelse i scorer – og dermed også til uoverensstemmelse i scorer på tværs af test. En test, der ikke i tilfredsstillende grad er enig med sig selv, kan heller ikke i tilstrækkelig grad være enig med en anden test. De fire test, som indgik i Keenan og kollegers undersøgelse, var alle udgivne testmaterialer, der anvendes i amerikansk praksis og forskning, og de publicerede reliabilitetskoefficienter spændte fra 0,64 til 0,98 (Keenan & Meenan, 2014). Det kan altså ikke udelukkes, at moderat testpålidelighed for nogle af testenes vedkommende kan være én af flere faktorer, der bidrog til den begrænsede stabilitet på tværs af testene.

Foruden begrænset testpålidelighed, må også indbyrdes forskelle på testmaterialerne være en væsentlig faktor, der kan bidrage til begrænset stabilitet på tværs af test. En oversigt over nogle forskelle og ligheder mellem de fire læseforståelsestest fremgår af tabel 1.

	Test	Tekstlængde	Svarmåde	Læsemåde	Tekst til rådighed ved løsning af opgaver eller ej
Mest følsom for sprogforståelse	GORT-3 Gray Oral Reading Test	85-150 ord	Multiple choice	Højtlæsning	Tekst til rådighed
	QRI-3 Qualitative Reading Inventory	250-785 ord	Genfortælling, besvarelse af åbne spørgsmål	Højtlæsning	Tekst ikke til rådighed
Mest følsom for afkodning	PIAT Peabody Individual Achievement Test	Enkelt-sætninger	Billed-udpegning	Stillelæsning	Tekst ikke til rådighed
	WPJC-3 Woodcock-Johnson Passage Comprehension	1-2 sætninger	Udfyldning af manglende ord i teksten (cloze)	Stillelæsning	Tekst til rådighed

Tabel 1. En analyse og oversigt over de fire læseforståelsestest som indgik i undersøgelser af Keenan et al. (2014) og Keenan og Meenan (2014). Tabellen viser punkter, hvorpå de fire test er tydeligt forskellige.

Det ses af tabel 1, at der for det første er forskel på, i hvilken grad testene har vist sig at være følsomme for afkodning og sprogforståelse. Testene er desuden forskellige, når det kommer til tekstlængde, svarmåde, læsemåde, og hvorvidt deltageren har teksten til rådighed under opgaveløsningen. Tekstlængde kan siges at være knyttet til eller have betydning for teksternes *indhold*, svarmåden er afhængig af *testformatet*, og læsemåden, det vil her sige om deltageren læser for sig selv (stillelæsning) eller læser højt, kan ses som en del af den *testbetingelse*, som deltageren skal forstå og demonstrere sin forståelse under. Det samme gælder spørgsmålet om, hvorvidt deltageren har teksten til rådighed eller ej under opgaveløsningen. I det følgende vil overvejelser om betydningen af forskelle i indhold, testformat og testbetingelse for stabilitet af profilen dårlig forstået på tværs af test blive præsenteret. Jeg tager udgangspunkt i forskellene på de fire test i undersøgelsen af Keenan og kolleger (alle blev afviklet som individuelle test) og supplerer med andre variationer, der kan være relevante i forbindelse med en distinktion mellem gruppetestning og individuel testning.

Forskelle i indhold

De fire test, som Keenan og kolleger anvendte, var som nævnt i forskellig grad følsomme for sprogforståelse og afkodning, hvilket i sig selv potentielt kan bidrage til ustabilitet på tværs af test. Foruden disse forskelle i følsomhed for afkodning og sprogforståelse er det også oplagt, at testene på grund af forskelle i indhold adskiller sig lidt eller meget i følsomhed for forskellige komponenter i sprogforståelse. Variationer i teksternes længde i de fire test hos Keenan og kolleger kan som

nævnt indikere, at testene er indholdsmæssigt forskellige og derved stiller forskellige krav til læserens sprogforståelse. Som det fremgår af tabel 1, er der meget stor forskel på, hvor lange teksterne, som eleverne læser i de enkelte test, er. I én test indgår fx kun enkeltsætninger, hvor en anden test består af tekster på op til 785 ord. Der er sammenfald mellem test med længere tekster og test, der er mest følsomme for sprogforståelse, og der er sammenfald mellem test med tekster, der består af 1-2 sætninger, og test, der er mest følsomme for afkodning. I flere undersøgelser ser man en tendens til at test med korte tekster er mere følsomme for afkodning end test med længere tekster og omvendt (Keenan, Betjemann & Olson, 2008; Andreassen & Bråten, 2010). Men tekstlængde kan også meget vel tænkes at have betydning for testens relative følsomhed for specifikke komponenter i sprogforståelse. I en test, hvor hver af teksterne består af bare en enkelt sætning, er det fx ganske enkelt ikke muligt at have opgaver, der stiller krav til læseren om at drage inferenser, der integrerer betydning fra flere sætninger. Keenan og kolleger undersøgte imidlertid ikke testenes følsomhed for forskellige komponenter i sprogforståelse. Det havde fx været interessant at se, i hvilken grad de forskellige test korrelerede med komponenter som ordforråd og inferensfærdigheder. Forskelle her ville kunne være med til at forklare en del af den begrænsede stabilitet på tværs af testene: Som nævnt i kapitel 2 må man fx forvente, at elever, der har udprægede vanskeligheder med inferenser, scorer lavere på en læseforståelsestest med store inferenskrav end på en læseforståelsestest med lave eller ingen inferenskrav.

Forskelle i testformat

At den specifikke tests format, det vil sige opgavernes konkrete udformning, kan have en indflydelse på læseforståelse, har flere undersøgelser sat fokus på (fx Nation & Snowling, 1997; Francis, Fletcher, Catts & Tomblin, 2005), og de fire testmaterialer i Keenan og kollegers undersøgelser er da også meget forskellige. De inkluderer: en test med et multiple choice format; en test hvor genfortælling af den læste tekst og besvarelse af åbne spørgsmål til teksten kombineres til en forståelsesscore; en test med billedudpegning; og en test med clozeformat, hvor læseren udfylder blanke felter i teksten med passende ord. I en række undersøgelser har man fundet indikationer på, at en test med clozeformat kan være mere følsom for afkodningsfærdigheder end en test med et format, der inkluderer åbne spørgsmål (Nation & Snowling, 1997; Gellert & Elbro, 2013). Men betydningen af testformatet er ikke fuldstændig uafhængig af indholdsmæssige overvejelser i testkonstruktionen. Således kan også en læseforståelsestest med clozeformat konstrueres, så den er følsom for sprogforståelse (Spear-Swerling, 2004), og så den i udstrakt grad korrelerer med mere konventionelle læseforståelsestest med åbne spørgsmål (Gellert & Elbro, 2013).

Identiske testformater er på ingen måde en garanti for, at test er lige følsomme for forskellige komponenter i tekstforståelse (Andreassen & Bråten, 2010) og dermed heller ikke for stabilitet, men at forskelle i testformat potentielt kan have selvstændig betydning for stabiliteten af profilen 'dårlig forstår' vidner en undersøgelse af Spooner et al. (2004, undersøgelse 3) om. Her blev dårlige og gode forståere udvalgt på baggrund af første del af den tidligere nævnte læseforståelsestest NARA, som er en test med åbne spørgsmål. Grundet udvælgelseskriterierne adskilte de to grupper sig signifikant i læseforståelse på denne test. Eleverne prøvede også anden del af NARA, men i en version hvor Spooner og kolleger havde modificeret spørgsmålene, så de var blevet til sandt-falsk

spørgsmål. På disse scorer adskilte de to grupper af forståere sig *ikke* signifikant. Gode forståere scorede på samme niveau som på testen med åbne spørgsmål, men dårlige forståere havde tilsyneladende gavn af det alternative testformat og øgede deres gennemsnit betydeligt på testen med sandt-falsk spørgsmål. Testformatet med sandt-falsk spørgsmål syntes altså at være lettere for de dårlige forståere end formatet med åbne spørgsmål, og det i en sådan grad, at forskellen på grupperne udviskedes.

Forskelle i testbetingelse

En forskel i testbetingelse, der har vist sig at have betydning for resultatet på en læseforståelsestest, er, hvorvidt læseren har teksten til rådighed under opgavebesvarelsen eller ej. Hvis læseren *ikke* har adgang til teksten, som det er tilfældet i to af testene hos Keenan og kolleger, bliver scoren på testen i højere grad afhængig af arbejdshukommelsesressourcer (Keenan et al., 2014; Andreassen & Bråten, 2010). Som det blev skitseret tidligere, er der nogle testbetingelser, der hyppigt knytter sig til gruppetestning og individuel testning. I en gruppetest af læseforståelse stillelæser eleverne teksterne og besvarer opgaver skriftligt. Typisk er der en øvre tidsgrænse for opgavebesvarelsen, men eleverne arbejder sig som oftest igennem opgaverne i deres eget tempo – og måske også i deres egen rækkefølge. Heroverfor indebærer en individuel test af læseforståelse derimod ofte, at eleven læser højt for testlederen, hvilket giver mulighed for at korrigere læsefejl. Sådan en procedure, hvor der gives individuel tilpasset støtte til afkodningen, anvendes hyppigt i undersøgelser med dårlige forståere, hvor man ønsker at reducere eller helt fjerne indflydelse fra afkodningsfærdigheder til forståelse. Der er typisk ingen øvre tidsgrænse på den individuelle læseforståelsestest, og rækkefølge og tempo bestemmes i høj grad af testleder. Det er også almindeligt, at testleder stiller opgaverne/spørgsmål mundtligt, og at eleven besvarer dem mundtligt.

I det følgende vil der være fokus på forskelle på de to skitserede betingelser, der kan knytte sig til gruppetestning og individuel testning, nemlig: Om deltagerne stillelæser eller højt læser teksterne, om der gives støtte til afkodningen, om testen afvikles med eller uden en tidsgrænse, og hvem der har det primære ansvar for opgaveløsningens forløb.

Stillelæsning eller højt læsning

I gruppetest vil der næsten altid være tale om, at deltageren skal stillelæse teksterne i testen. I individuel testning er der til gengæld mulighed for, at deltageren højt læser teksterne. Højt læsning er i tidligere undersøgelser blevet forbundet med lidt højere forståelsesscorer end stillelæsning (Elgart, 1978), og måske gælder denne tendens særligt for svage læsere frem for gennemsnitlige og gode læsere, der i en anden undersøgelse opnåede samme forståelsesniveau, uanset om de stillelæste eller højt læste (Swalm, 1972). En mulig årsag til højt læsnings betydning for forståelsen kan være, at læseren ved højt læsning i modsætning til stillelæsning er tvunget til at give alle ord i teksten opmærksomhed (Elgart, 1978; Swalm, 1972).

Det er værd at lægge mærke til, at de to test, der i Keenans undersøgelser er mest følsom for sprogforståelse, inkluderer højt læsning, hvorimod de to test, der er mest følsom for afkodning,

inkluderer stillelæsning. En meta-analyse af Garcia og Cain (2014) med data fra 110 undersøgelser viste netop også, at læsemåden kan påvirke sammenhængen mellem læseforståelse og afkodningsfærdigheder, men tydede også på, at betydningen af læsemåden varierer med alder. Garcia og Cain fandt på tværs af de 110 undersøgelser, at der var en stærkere sammenhæng mellem tekstforståelse og afkodningsfærdigheder, når der var tale om læseforståelsestest, der inkluderede stillelæsning, end når der var tale om læseforståelsestest, der inkluderede højtlesning. Ved opsplitning i to aldersgrupper var forskellen signifikant for deltagere under 10 år, mens der var tale om en ikke-signifikant tendens for deltagere over 10 år. De to aldersgrupper (under og over 10 år) blev dannet, fordi man generelt fandt en markant reduktion i sammenhængen mellem afkodningsfærdigheder og læseforståelse for elever over 10 år. Hvordan stillelæsning frem for højtlesning bidrager til følsomhed for afkodning er dog ikke umiddelbart klart. I nogle tilfælde kan svaret måske findes i et andet aspekt, som ofte knytter sig til højtlesning over for stillelæsning. Nemlig om deltageren modtager støtte til afkodningen af svære ord. Det synes ikke at have været tilfældet i Keenan og kollegers undersøgelse, men støtte til afkodning er en ikke sjældent anvendt testprocedure (fx i den før omtalte test af læseforståelse NARA), når man tester tekstforståelse for at afgrænse en gruppe dårlige forståere.

Med eller uden støtte til afkodning

Individuel testning, hvor deltageren højtleser, giver mulighed for at yde støtte til afkodning af svære ord. Rationalet er, at denne støtte øger sandsynligheden for at afdække forskelle i forståelse af teksten og ikke forskelle, der fx skyldes fejlagtig afkodning af et centralt ord i teksten. Det samme er ikke muligt i gruppetest, hvor deltageren stillelæser. Han eller hun må forsøge at forstå teksten på baggrund af evt. fejlaftkodede ord og uden den støtte, det er at vide, at man får de rigtige ord oplyst ved afkodningsproblemer. I Garcia og Cains metaundersøgelse (2014) undersøgte man betydningen af, om deltagerne fik hjælp til afkodning af de enkelte ord i teksterne i læseforståelsestesten eller ej. For det samlede sample var der dog ikke signifikant forskel på sammenhængen mellem afkodningsfærdigheder og læseforståelse i test med og uden hjælp til afkodning af teksterne, og ved opsplitning i de to aldersgrupper fandt man overraskende, at sammenhængen for de yngre deltagere (under 10 år) overordnet set var *stærkere* i undersøgelser, hvor deltagerne fik hjælp til afkodningen. Garcia og Cain foreslog, at dette overraskende resultat måske kunne skyldes, at afkodningshjælpen for yngre relativt utrænede læsere virker som en afbrydelse i forståelsesprocessen.

Med eller uden tidsgrænse

En del læseforståelsestest, og typisk gruppetest, bliver afviklet med en øvre grænse for deltagernes tidsforbrug. Hvorvidt en læseforståelsestest gennemføres under et tidspres kan måske også være vigtig for resultatet på testen og en potentiel faktor med betydning for ustabilitet på tværs af læseforståelsestest. Walczyk (1993; 1995) har foreslået, at læseren under et tidspres ikke i samme grad kan kompensere for eventuelle svagheder (der normalt ikke er synlige) i automatiserede

delkomponenter i læseforståelse/processer med betydning for læseforståelse som fx benævnelse og adgang til ordbetydninger (undersøgt ved en overbegrebstest), og at et tidspres derfor kan ændre på den sammenhæng, man kan observere mellem disse processer og tekstforståelse. I forlængelse af Walczyks hypotese kan man forestille sig, at stabiliteten på tværs af test med og uden tidsgrænse kan påvirkes som en konsekvens af forskelle i disse sammenhænge. Det er dog et vigtigt forbehold, at det tidspres som Walczyk undersøgte læseforståelse under, blev skabt ved, at teksterne kun blev vist i ganske kort tid for læserne. Det er ikke den form for tidspres læsere oftest møder – og slet ikke i en læseforståelsestest. I læseforståelsestest, der skal kunne anvendes i praksis, vil et tidspres typisk udgøres af en øvre tidsgrænse på testen. Desuden kan de korrelationer, som Walczyk bruger til at underbygge hypotesen, alternativt tolkes sådan, at de først og fremmest viser, at hastighedsmål generelt korrelerer bedre med (andre) hastighedsmål, end de korrelerer med præcisionsmål. De mål for processer, der var signifikant stærkere korreleret med scorer på læseforståelsestest med tidspres end scorer på læseforståelsestest uden tidspres, var nemlig *responstiderne*, hvorimod der ikke på samme måde var en tendens for sammenhængen mellem præcisionsmål og scorer på læseforståelsestest i de to betingelser (se fx Walczyk, 1995, tabel 3, side 405).

Da afkodning netop er en automatiseret delfærdighed i læsning, er afkodning også et eksempel på en delkomponent eller proces, hvor læseren ifølge Walczyk kan kompensere for svagheder, så længe han eller hun *ikke* læser under et præstationspres. Ud fra Walczyks hypotese kunne man derfor forvente, at tilstedeværelse af et tidspres i læseforståelsestest ville påvirke (øge) den sammenhæng, man kan observere mellem afkodning og læseforståelse. Rønberg og Petersen (2015b) fandt netop, at afkodning (målt med en test af ortografisk kodning) var stærkere korreleret med læseforståelse, når læseforståelse blev opgjort som antal rigtige pr. minut, end når antal rigtige blev opgjort. Det kan her nævnes, at Walczyks egne resultater muligvis delvist afspejler sammenhænge mellem afkodning og læseforståelse, da alle de anvendte test af automatiserede processer inkluderede afkodning af ord, og der ikke blev kontrolleret for denne delfærdighed. Men i Garcia og Cains (2014) metaundersøgelse var der imidlertid *ikke* tegn på, at variabelen 'tidsgrænse/ingen tidsgrænse' modererede korrelationen mellem afkodning og læseforståelse. Ideen om, at et tidspres kan influere på resultatet på en læseforståelsestest, synes dog relevant i mere generel forstand. Det er oplagt, at mennesker kan opfatte – og reagere på – tidspres forskelligt. Og hvis læsere reagerer forskelligt under tidspres, kan disse forskelle i adfærd måske føre til en grad af ustabilitet på tværs af læseforståelsestest, der afvikles med og uden en tidsgrænse.

Ansvar for opgaveløsningens forløb

Under individuel testning er arbejdet med tekst og opgaver i høj grad (med)styret af testlederen: Testleder fører deltageren igennem teksterne i en bestemt rækkefølge, testleder læser opgaverne højt i en bestemt rækkefølge, og eleven kommer gennem opgaverne i et tempo, der delvist sættes af testlederen. Hvis deltageren er meget længe om at besvare et spørgsmål, måske fordi det er for svært, spørger testleder, om deltageren vil gætte på noget, og man går videre til næste opgave. Heroverfor synes gruppetestning, hvor eleven læser og arbejder for sig selv, i høj grad at sætte krav til, at deltageren styrer igennem tekst og opgaver selv på bedst mulig måde. En rækkefølge er foreslået, og en tidsramme er lagt ud, men resten er op til deltageren. I sådanne situationer kan det,

som Vidal-Abarca, Salmeron & Mana (2011) kalder *selvregulerende adfærd*, måske være vigtig. Vidal-Abarca et al. (2011) kalder det fx selvregulerende adfærd, når læseren, der kan gå frem og tilbage mellem tekst og spørgsmål/opgaveformuleringen i læseforståelsestesten, beslutter sig for enten at læse tekst eller spørgsmål først og beslutter, om besvarelse af et spørgsmål kræver, at han eller hun læser/gennemgår teksten endnu engang. Det er også selvregulerende adfærd, når læseren beslutter, hvornår en opgave er løst tilfredsstillende/godt nok til, at man vil fortsætte til næste opgave eller tekst (Vidal-Abarca et al., 2011). Sådanne forskelle i selvregulerende adfærd kan muligvis påvirke scoren på en læseforståelsestest. Vidal-Abarca og kolleger har fx vist, at det ofte giver det bedste resultat, hvis deltagerne læser teksten før spørgsmålene (Cerdán, Vidal-Abarca, Martínez, Gilabert & Gil, 2009).

Opsamling

Den publicerede forskning i stabiliteten af profilen 'dårlig forstår' tyder på, at stabiliteten kan svinge endog meget stærkt fra så lavt som 6% til helt op omkring 90%. Der ses generelt en høj stabilitet over tid, men variation i den anvendte test af læseforståelse synes at kunne medføre en høj grad af ustabilitet. Undersøgelser tyder på, at forskelle i testenes indhold og format kan være centrale. Hvorvidt forskelle på de betingelser, som testene afvikles under, kan have betydning, er ikke belyst i samme grad. Forskellige variationer i testbetingelse forekommer dog hyppigt på tværs af anvendte test i praksis og forskning, fx i forbindelse med gruppetestning over for individuel testning – ligesom testning i både praksis og forskning ofte også afvikles over en kortere periode. På den baggrund var det oplagt at undersøge stabiliteten af profilen 'dårlig forstår' på tværs af sådanne hyppigt forekommende variationer.

Formål

Det var formålet med den aktuelle undersøgelse at bidrage til indsigten i, om profilen 'dårlig forstår' er stabil, sådan at en læser, der defineres som dårlig forstår på ét tidspunkt under én bestemt betingelse, også har vanskeligheder med sprogforståelse i læsning på et andet tidspunkt under en anden betingelse. Den aktuelle undersøgelse skal ses som en første, dansk stikprøve på stabiliteten, når man tillader typiske, praktisk forekommende variationer. Undersøgelsen vurderede således stabiliteten af profilen over et interval på 4-5 måneder og med to læseforståelsestest, der var parallelle indholdsmæssigt; men hvor den ene blev gennemført som gruppetest med stillelæsning og med tidsgrænse; mens den anden blev gennemført som en individuel test med højtlesning og individuel støtte til afkodning og blev afviklet uden tidsgrænse. Forskningsspørgsmålene var følgende:

1. *Hvor stor stabilitet er der i klassifikationen af dårlige forståere over tid og over to typiske måder at gennemføre læseforståelsestest på?*

Som led i en indkredsning af mulige kilder til eventuel ustabilitet blev følgende spørgsmål stillet:

2a. *I hvilket omfang kan ustabilitet i klassifikationen skyldes begrænset pålidelighed af læseforståelsestestene?*

2b. *Er der indikationer på, at ustabilitet for profilen 'dårlig forstår' kan have at gøre med at de to betingelser varierer i følsomhed for afkodningsfærdigheder?*

Metode

Design

Der blev indhentet data ved to testtidspunkter, og stabiliteten af profilen 'dårlig forstår' blev undersøgt på tværs af disse tidspunkter med indholdsmæssigt parallelle test, der blev afviklet som henholdsvis gruppetest og individuel test.

Ved første testtidspunkt blev læseforståelse og afkodningsfærdigheder undersøgt med gruppetest i et uselekeret sample på 203 elever på 6. klassetrin. Af ressourcehensyn blev ikke alle elever testet igen ved andet testtidspunkt, hvor læseforståelse blev undersøgt med en individuel test. Et udsnit af de elever, der havde dansk som førstesprog og ingen kendte vanskeligheder, blev testet ved andet testtidspunkt. Blandt disse elever var en gruppe af dårlige forståere. Udvælgelsen af dårlige forståere var baseret på definitionen af dårlige forståere i kapitel 2. Blandt eleverne, der blev gentestet, var også en gruppe af gode forståere. De konkrete kriterier for udvælgelse af dårlige og gode forståere er beskrevet i det følgende. Yderligere inklusion af læsere på andre forståelses- og afkodningsniveauer (opåfaldende forståere og relativt svage afkodere) i gentestningen sikrede desuden et mindre samlet sample af elever med jævn spredning i læseniveau. Scorer fra dette sample med jævn spredning i læseniveau indgik i fuld-skala analyser uden gruppeinddeling som supplement til analyser med de to grupper af forståere.

Deltagere

I undersøgelsen deltog elever fra 13 klasser fordelt på seks skoler (fordelt på de københavnske bydele Østerbro, Valby, Frederiksberg og Christianshavn). I de 13 deltagende klasser gik der samlet 308 elever. 211 elever afleverede en tilladelse. Tre ønskede ikke, at deres besvarelser indgik i undersøgelsen. Den samlede svarprocent var dermed på 69%. Svarprocenten i de enkelte klasser spændte fra 47% til 96%. Ved første testtidspunkt var 203 elever med tilladelse tilstede, og for 197 af disse elever blev der indsamlet et komplet datasæt.

Materialer

Læseforståelse

Første testtidspunkt. Læseforståelse blev undersøgt med oversatte opgaver fra *The Diagnostic Reading Analysis* (form A) (Crumpler & McCarthy, 2004). Det engelske testmateriale består af en række tekster og spørgsmål af stigende sværhedsgrad og dækker et aldersspænd fra 6 til 16 år. Materialet er i udgangspunktet udformet som en individuel test, hvor deltageren højt læser tekster (udvalgt på baggrund af deltagerens alder og formodede niveau) og mundtligt besvarer spørgsmål til teksterne. For hver sværhedsgrad er der både en narrativ tekst og en fagtekst. Spørgsmålene til teksterne er ifølge testmanualen rettet mod hhv. inferentiell/opsummerende information, bogstavelig information fra teksten, forudsigelser og kendskab til bestemte ordbetydninger.

Oversættelser af otte fortløbende tekster (fire narrativer og fire fagtekster) med tilhørende illustrationer og spørgsmål blev trykt i et opgavehæfte til den aktuelle undersøgelse. De første opgaver i hæftet var de to opgaver, der i det engelske materiale var på et niveau tiltænkt deltagere på 10-11 år, og de sidste to opgaver i hæftet var opgaverne, der i det engelske materiale var på et niveau tiltænkt deltagere på 13-14 år. I den aktuelle undersøgelse blev testen afviklet som en gruppetest. Deltagerne (stille)læste teksterne og besvarede efter hver tekst skriftligt de efterfølgende spørgsmål. Der var en tidsgrænse på 30 minutter til at nå så mange tekster og spørgsmål som muligt. Der var i alt 46 spørgsmål i testen.

Besvarelsene blev scoret på baggrund af en tidligere udviklet manual, der gav eksempler på svar, der skulle tildeles hhv. 1 og 0 point (Daugaard & Elbro, 2013). Alle besvarelses blev scoret af samme person. På tværs af forskellige bedømmere har enigheden tidligere vist sig at være meget høj (mellem 97,3 og 99,5%) (Daugaard & Elbro, 2013). Testens interne homogenitet i den aktuelle undersøgelse var tilfredsstillende (Cronbach's Alfa = 0,75⁷). Testen findes i den selvstændige bilagssamling (Bilag A).

Andet testtidspunkt. Læseforståelse blev undersøgt med oversatte opgaver fra *The Diagnostic Reading Analysis* (form B) (Crumpler & McCarthy, 2004). De udvalgte otte tekster med tilhørende spørgsmål matchede de udvalgte tekster og opgaver fra *The Diagnostic Reading Analysis* (form A), der blev anvendt ved første testtidspunkt, i sværhedsgrad, indhold og form.

Testen blev afviklet som en individuel test. Deltageren læste teksterne højt. Hvis der opstod problemer med at afkode et ord, blev dette ord læst højt af tester for at minimere indflydelse fra afkodningsfærdigheder på testresultatet. Efter hver tekst stillede tester en række spørgsmål til teksten, som deltageren besvarede et efter et. Disse spørgsmål fremgik ikke af elevens hæfte, men selve teksten var til rådighed under opgavebesvarelsen, og eleven kunne kigge i den, før han eller hun besvarede et spørgsmål. I alt var der 46 spørgsmål i testen.

Besvarelsene fra læseforståelsestesten blev scoret ud fra transskriptioner af lydoptagelser og direkte fra lydoptagelser. Besvarelsene blev scoret på baggrund af en tidligere udviklet manual til

⁷ Beregning af Cronbach's Alfa er baseret på de første 30 items og besvarelses fra 198 børn (98% af 203 deltagere der prøvede testen), der enten nåede til eller længere end det 30. item i testen.

scoring af de samme opgaver, dog baseret på skriftlige besvarelser (Daugaard & Elbro, 2013), men med enkelte ændringer foretaget på baggrund af mundtlige svar, der kom frem i piloteringen af materialet anvendt som en test med højtlesning og mundtlig besvarelse af spørgsmålene. Alle besvarelser blev scoret af samme person. På tværs af forskellige bedømmere har enigheden tidligere vist sig at være meget høj (mellem 96,1 og 99,5%) (Daugaard & Elbro, 2013), når testen blev afviklet som skriftlig test. Testens interne homogenitet i den aktuelle undersøgelse var tilfredsstillende (Cronbach's Alfa = 0,73). Testen findes i den selvstændige bilagssamling (bilag B).

Tabel 2 opsummerer forskelle på de to testbetingelser, som de to læseforståelsestest blev afviklet under.

Tabel 2. Forskelle på de testbetingelser, som læseforståelsestestene ved de to testtidspunkter blev afviklet under. Testene var indholdsmæssigt parallelle

Første testtidspunkt	Andet testtidspunkt
Gruppetest	Individuel test
- Stillelæsning	- Højtlesning
- Deltager læser spørgsmål	- Deltager lytter til spørgsmål
- Skriftlig besvarelse	- Mundtlig besvarelse
- Tidsgrænse	- Ingen tidsgrænse
- Ingen støtte til afkodning	- Støtte til afkodning

Fonologisk kodning

Fonologisk kodning blev undersøgt med testen *Find det der lyder som et ord* (Nielsen & Petersen, 1993a). Testen undersøger deltagerens udnyttelse af skriftens lydprincip. I hver opgave skal deltageren tage stilling til, hvilket ud af fire nonord der ville lyde som et rigtigt ord, hvis man læste det højt, det vil sige hvilket nonord, der er homofon til et ord. Sådanne fire nonord kunne være *jir*, *jal*, *jyr* og *jor*, hvor kun *jor* lyder som et rigtigt ord (jord), hvis det læses højt. Testen har en tidsgrænse. Der er 38 opgaver og 5 minutter til at nå så mange opgaver som muligt. Testen korrelerer stærkt med læsning af nonord ($r = 0,81$) (Elbro, Nielsen & Petersen, 1994) og anvendes ofte som et alternativ til test af højtlesning af nonord. Split-half reliabiliteten har i en tidligere undersøgelse vist sig at være 0,98 (Elbro, Nielsen & Petersen, 1994). I den aktuelle undersøgelse var split-half reliabiliteten (ulige over for lige opgavenumre) 0,85⁸.

⁸ Baseret på besvarelser fra 203 elever.

Ortografisk viden

Ortografisk viden blev undersøgt med testen *Find det ord der er rigtigt stavet* (Nielsen & Petersen, 1993a). Testen undersøger deltagerens genkendelse af ords korrekte stavemåder. I hver opgave skal deltageren tage stilling til hvilken af fire forskellige stavemåder for et ord, der er den korrekte (fx *fesd*, *fæsth*, *fest* eller *fæsd*). Testen har en tidsgrænse. Der er 38 opgaver og 1 minut til at nå så mange opgaver som muligt. Testen kan anvendes som alternativ til højtlesning af ord. Blandt voksne ordblinde korrelerer testen moderat med stillelæsning af sætninger (Spearman's $\rho = 0,54$) og moderat til stærkt med stavning af lydrette og ikke-lydrette ord (Spearman's $\rho = 0,85$ og $0,62$), men ikke signifikant med stavning af nonsensord (Nielsen & Petersen, 1993b). I den aktuelle undersøgelse var split-half reliabiliteten (ulige over for lige opgavenumre) $0,92^9$.

Fremgangsmåde

Forud for testningen blev der via elevernes dansklærere uddelt skriftlige forespørgsler til elevernes forældre om at give tilladelse til, at deres børns opgavebesvarelser måtte indgå i undersøgelsen. Der blev ligeledes uddelt et spørgeskema bl.a. med spørgsmål om barnets første sprog og eventuelle kendte vanskeligheder.

Første testtidspunkt (T1). Testningen foregik i skoletiden og fandt sted i elevernes klasseværelse. I én klasse måtte testningen dog afvikles i skolens kantine-lokale. Testningen varede i alt 60 minutter. Testleder (jeg selv eller en studentermedhjælper) gennemgik instruktioner og eksempel-opgaver sammen med eleverne før hver test og besvarede elevernes eventuelle spørgsmål. En lærer, oftest elevernes dansklærer, var til stede under testningen i 12 ud af de 13 klasser. Som hovedregel sad elever, der enten ikke havde afleveret en tilladelse, eller hvis opgavebesvarelser ikke måtte indgå i undersøgelsen, sammen med deres klassekammerater i klasselokalet og løste opgaver. I to tilfælde sad disse elever dog efter aftale med læreren i et andet lokale, mens testningen blev gennemført.

Andet testtidspunkt (T2). Testningen foregik efter skoletid – som oftest i et lokale på deltagerens skole, men i nogle tilfælde blev testningen afviklet i et lokale på et fritidshjem eller på Københavns Universitet. Under testningen sad tester og deltager over for hinanden ved et bord med hver deres hæfte. Tester instruerede deltageren kort i opgavens form og forløb, og deltageren kunne stille spørgsmål inden de egentlige testopgaver blev påbegyndt. Testningen blev lydoptaget. Læseforståelsestesten var én af flere test i et større testbatteri, der blev afviklet over to testsessioner. Der var som hovedregel max 14 dage mellem disse to sessioner. Hver testsession varede 1,5 til 2 timer. Mellem testene i en session var der mulighed for pauser og forfriskninger. Deltagerne modtog et gavekort til en biografbillet som tak for deres deltagelse.

⁹ Baseret på besvarelser fra 202 elever.

Resultater

Blandt de 197 elever med positiv tilladelse og komplette datasæt var der 168 deltagere med dansk som førstesprog og ingen kendte vanskeligheder. Tabel 3 viser deskriptiv statistik for disse deltagere ved første testtidspunkt.

Tabel 3. Deskriptiv statistik for deltagere med dansk som første sprog og ingen kendte vanskeligheder ved T1

	Middelværdi	Standardafvigelse	Spænd
Læseforståelse (max 46)	33,38	6,70	10-44
Fonologisk kodning (max 38)	21,61	6,39	4-37
Ortografisk viden (max 38)	19,19	6,49	5-37

Note. N = 168.

Alle scorefordelinger for målene i tabel 3 havde acceptable værdier for skævhed (inden for ± 1).

Tabel 4. Korrelationer mellem læseforståelse, fonologisk kodning og ortografisk viden for deltagere med dansk som førstesprog og ingen kendte vanskeligheder ved T1

	Læseforståelse	Fonologisk kodning	Ortografisk viden
Læseforståelse			
Fonologisk kodning	0,43		
Ortografisk viden	0,45	0,44	

Note. N = 168. Alle korrelationer er signifikante, $p < 0,01$.

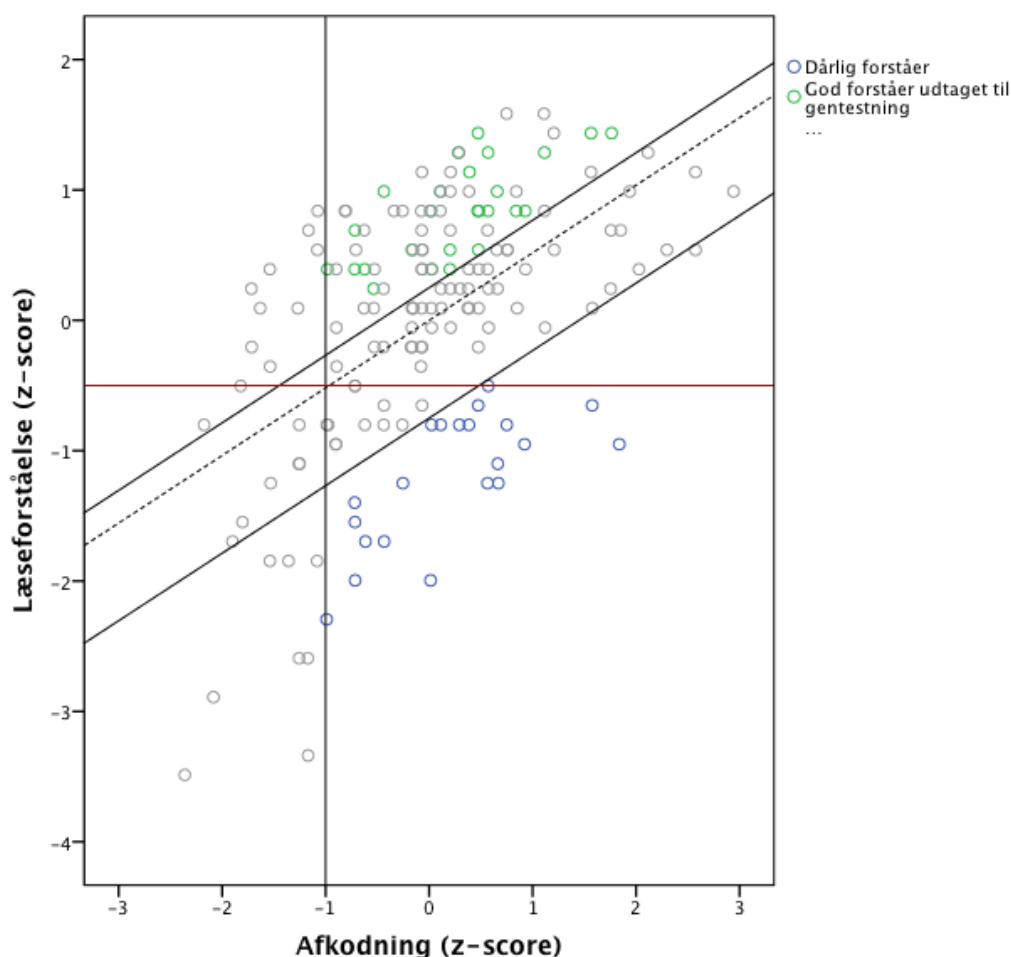
Af tabel 4 fremgår det, at alle mål ved T1 var signifikant korreleret. Som forventet korrelerede målet for læseforståelse med de to mål for delfærdigheder i afkodning, fonologisk kodning og ortografisk viden. Den moderate indbyrdes korrelation mellem mål for fonologisk kodning og ortografisk viden var i overensstemmelse med, at der netop er tale om mål for to delfærdigheder i afkodning. For at få et samlet mål for afkodningsfærdigheder blev scorene fra fonologisk kodning og ortografisk viden kombineret til én afkodningsscore. Scorene fra hver test blev standardiseret på de 168 deltagere og blev derefter lagt sammen. Disse afkodningsscorer blev dernæst standardiseret (z-score), hvorved den gennemsnitlige score blev udtrykt ved værdien 0 og standardafvigelsen lig med 1. Læseforståelsesscorerne blev ligeledes standardiseret (z-score). Det samlede afkodningsmål og læseforståelse korrelerede som forventet ($r = 0,52$, $p < 0,01$).

Udvælgelse af dårlige og gode forståere ved første testning

Dårlige forståere blev udvalgt i overensstemmelse med den udvælgelsesmetode, som blev foreslået i kapitel 2. For eleverne med dansk som førstesprog og ingen kendte vanskeligheder (tabel 3), blev

læseforståelse som funktion af afkodningsfærdigheder vurderet ved hjælp af lineær regression. I den lineære regressionsanalyse blev det samlede mål for afkodning brugt.

Figur 1 viser deltagernes resultater på læseforståelsestesten ved første testtidspunkt som funktion af deltagernes afkodningsfærdigheder. I figuren er dårlige forståere og gode forståere, der blev udtaget til gentestning, fremhævet. De konkrete grænsedragninger for deltagere med disse profiler beskrives herunder.



Figur 1. Resultater på test af læseforståelse som funktion af afkodningsfærdigheder ved T1 ($n = 168$). Figuren viser grænsedragninger for dårlige og gode forståere. Elever med disse profiler, som blev udvalgt til gentestning, er desuden markeret.

'Dårlige forståere'

Dårlige forståere blev defineret som deltagere med en læseforståelsesscore, der lå mere end 0,5 standardafvigelse fra middelværdien (vandret rød linje i figur 1), hvormed de var placeret blandt den dårligste tredjedel i læseforståelse, og som samtidigt lå mere end 0,75 standardafvigelse under den forventede score ud fra afkodningsscoren (udtrykt ved regressionslinjen, stiplede linje i figur 1) (dvs. at deres vanskeligheder med læseforståelse ikke kunne forklares af dårlig afkodning). De dårlige forståere måtte heller ikke have vanskeligheder med afkodning i det hele taget, så de

scorede mindst -1 standardafvigelse fra middelværdien i afkodning (lodret linje i figur 1). Der var 21 deltagere, der blev defineret som dårlige forståere ud fra disse tre grænsedragninger.

'Gode forståere'

Gode forståere blev defineret som deltagere med en læseforståelsesscore, der lå mere end 0,25 standardafvigelse over regressionslinjen, hvis afkodningsscoren samtidig lå over -1 standardafvigelse fra middelværdien. 21 gode forståere blev udvalgt fra de samme klasser som de dårlige forståere og sådan, at de bedst muligt matchede de udvalgte dårlige forståere på afkodningsfærdigheder. For at sikre en gruppe af en vis størrelse ved gentestning i tilfælde af frafald blev yderligere fem gode forståere udvalgt til gentestning.

I alt 62 deltagere (dårlige forståere, gode forståere og andre deltagere) blev efter første testtidspunkt udtaget til gentestning. Blandt disse elever faldt fire dårlige forståere fra, idet de enten ikke kunne eller ikke havde lyst til at deltage i gentestningen. Også fire elever, der hverken var dårlige eller gode forståere, faldt fra, idet de enten ikke kunne deltage i testning eller ikke mødte op til de aftalte tidspunkter. I alt deltog 54 deltagere dermed ved begge testtidspunkter. Tabel 5 viser deskriptiv statistik for disse 54 deltagere.

Tabel 5. Deskriptiv statistik for deskriptiv statistik for deltagere der deltog ved begge testtidspunkter

	Middelværdi	Standardafvigelse	Spænd
T1			
Læseforståelse (max 46)	32,94	7,27	16-43
Fonologisk kodning (max 38)	21,94	5,76	10-37
Ortografisk viden (max 38)	20,52	5,55	5-36

Note. N = 54.

Alle scorefordelinger for målene i tabel 5 havde acceptable værdier for skævhed (inden for +/-1,1).

De 54 deltagere, der deltog ved andet testtidspunkt, kunne med nogen rimelighed antages at være repræsentative for det store sample (n = 168). De 54 deltagere (M = 32,94, standardafvigelse = 7,27) var ikke signifikant forskellige fra de 114 deltagere (M = 33,58, standardafvigelse = 6,44), der ikke deltog i gentestningen, i læseforståelse ved første undersøgelsestidspunkt $t(166) = -5,5$, $p > 0,05$. Elever, der deltog i gentestning, og elever, der ikke deltog i gentestningen, adskilte sig heller ikke signifikant på målet for fonologisk kodning (n = 54: M = 21,94, standardafvigelse = 5,76; n = 114 = 21,45, standardafvigelse = 6,68) $t(166) = 0,50$, $p > 0,05$. De var heller ikke signifikant forskellige på målet for ortografisk viden (n = 54: M = 20,52, standardafvigelse = 5,54; n = 114: M = 18,56, standardafvigelse = 6,82) $t(166) = 1,84$, $p > 0,05$.

Stabiliteten af 'dårlig forstår'-profilen

Tabel 6 viser resultater for dårlige og gode forståere, der deltog i gentestningen.

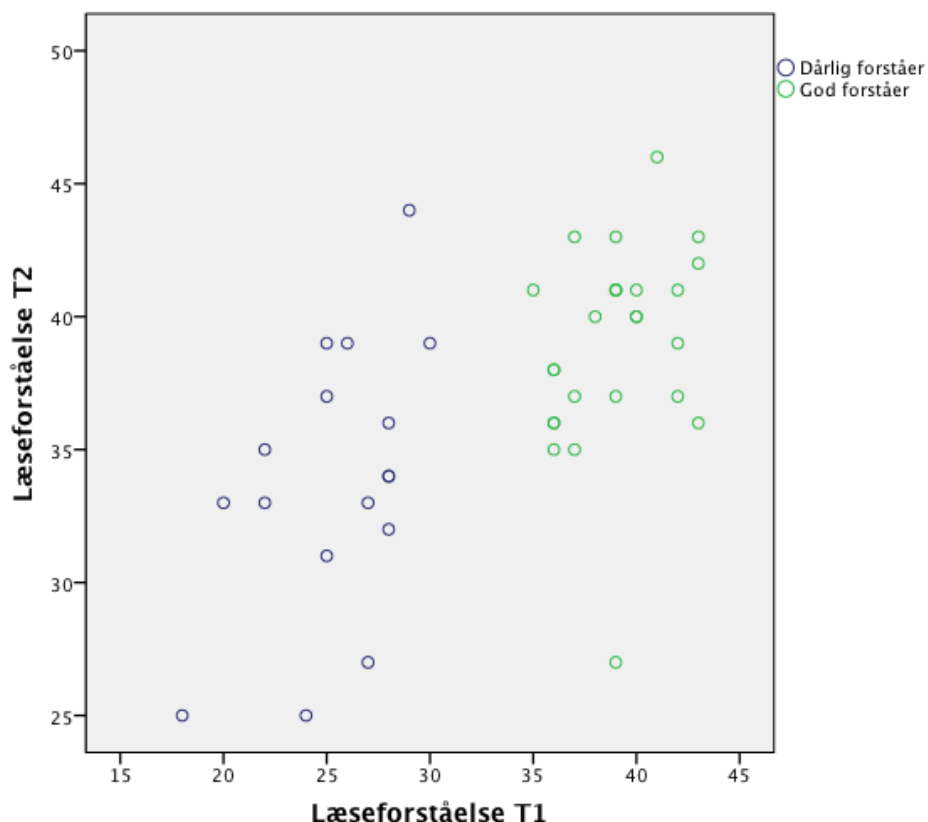
Tabel 6. Resultater fra første og andet testtidspunkt for dårlige og gode forståere

	Dårlige forståere (n = 17)				Gode forståere (n = 26)			
	Md	Q ₁	Q ₃	Spænd	Md	Q ₁	Q ₃	Spænd
T1								
Læseforståelse	26,00	23,00	28,00	18-30	39,00	36,75	41,25	35-43
Fonologisk kodning	23,00	20,50	25,50	13-33	22,50	19,00	26,25	13-37
Ortografisk viden	19,00	16,00	23,00	12-36	20,00	18,00	24,00	11-33
T2								
Læseforståelse	34,00	31,50	38,00	25-44	40,00	36,75	41,00	27-46

Grupperne var som bekendt selekteret, og variationen i scorerne er dermed begrænset. I det følgende anvendes der derfor ikke-parametrisk statistik (Mann Whitney test) i sammenligningen af de to gruppers centraltendenser.

Grupperne af dårlige og gode forståere, der deltog ved både første og andet testtidspunkt, var trods frafald fra gruppen af dårlige forståere ikke signifikant forskellige på scorer fra testene i fonologisk kodning, $U = 218$, $z = -0,08$, $p > 0,05$, og ortografisk viden, $U = 174$, $z = -1,17$, $p > 0,05$. Forskellen på de to gruppers læseforståelsesscorer ved T1 svarede fortsat til en meget stor effekt, $U = 0,00$, $z = -5,51$, $p < 0,01$, $r = -0,84$. Gruppernes læseforståelse ved T2 behandles i afsnittet om stabilitet på tværs af tid og testbetingelser nedenfor.

Det første spørgsmål var, hvor stor stabilitet der er i klassifikationen af dårlige forståere over tid og over to forskellige måder at gennemføre typiske læseforståelsestest på, gruppetestning og individuel testning. Resultaterne for dårlige og gode forståeres læseforståelse ved T2 fremgår af tabel 6. Ved T2 var læseforståelsesniveauet for de oprindeligt identificerede dårlige forståere (Median = 34) fortsat signifikant lavere end for de oprindeligt identificerede gode forståere (Median = 40), $U = 81$, $z = -3,49$, $p < 0,01$. Forskellen svarede til en stor effekt, $r = -0,53$. Som gruppe adskilte de udvalgte dårlige forståere sig dermed i høj grad fortsat fra gode forståere ved T2. Spændvidder og kvartiler ved T2 vidnede dog om et vist overlap. Figur 2 viser sammenhængen mellem læseforståelsesscorer for dårlige og gode forståere ved første og andet testtidspunkt og illustrerer overlappet mellem de to grupper, der opstod ved T2.



Figur 2. Dårlige og gode forståeres læseforståelsesscorer ved første og andet testtidspunkt.

Resultaterne i tabel 6 og illustrationen i figur 2 indikerer, at overlappet ved andet testtidspunkt primært opstod ved, at gennemsnit og spredning for gruppen af *dårlige forståere* øgedes fra T1 til T2 (på tværs af de to test). Blandt de *gode forståere* scorede én deltager markant dårligere på læseforståelsestesten ved T2 end ved T1 og desuden markant dårligere ved T2 end de andre gode forståere, men gruppens gennemsnit og spredning ændredes ikke væsentligt.

For at besvare spørgsmålet om stabiliteten af profilen 'dårlig forstår' blev det undersøgt, i hvilken grad deltagerne i de to distinkte grupper, der blev udvalgt ved T1, kunne genklassificeres i de respektive grupper ved T2. Det blev opgjort, hvilken scoreværdi ved T2 der afgrænsede de oprindeligt udvalgte dårlige og gode forståere bedst muligt fra hinanden. Tabel 7 viser den højest mulige grad af overensstemmelse med klassifikationen ved T1, man kunne opnå på baggrund af læseforståelsesscorerne for dårlige og gode forståere ved T2. Den højest mulige grad af overensstemmelse blev opnået, hvis grænsen blev trukket ved 34 rigtige på testen ved T2. Tabel 7 viser andelen af både dårlige og gode forståere identificeret ved T1, der havde enten til og med eller over 34 rigtige på testen ved T2.

Tabel 7. Dårlige og gode forståere identificeret ved T1, der klassificeres som dårlige og gode forståere ved T2, hvis 34 rigtige på læseforståelsestesten anvendes som grænseværdi

	Dårlig forstår ved T1	God forstår ved T1
T2 score ≤ 34	10 (23%)	1 (2%)
T2 score >34	7 (16%)	25 (58%)

I alt 81% af de 43 dårlige og gode forståere kunne klassificeres på samme måde ved T2 som ved T1. Den overordnede stabilitet var altså ganske god. Grupperne var langt fra distinkte ved andet testtidspunkt, men majoriteten af deltagerne i begge grupper kunne ved denne opgørelsesmåde placeres i samme kategori som ved første testtidspunkt. For 10 ud af de 17 dårlige forståere (59%) var der ved andet testtidspunkt overensstemmelse med klassifikationen fra første testtidspunkt, og for 25 af de 26 gode forståere (96%) var der overensstemmelse med klassifikationen fra første testtidspunkt. Der var dermed en betydelig stabilitet for dårlige forståere og en god stabilitet for gode forståere på tværs af de to testtidspunkter og de to måder at undersøge læseforståelse på.

Kilder til ustabilitet

Der blev stillet to spørgsmål som led i en indkredsning af mulige kilder til ustabilitet for profilen 'dårlig forstår'. Det første af disse var, i hvilket omfang ustabilitet i klassifikationen kan skyldes begrænset pålidelighed af læseforståelsestestene. For at vurdere testpålidelighedens mulige betydning for ustabiliteten blev overensstemmelsen mellem scorerne på de to test undersøgt for alle de deltagere, der havde prøvet begge test. Tabel 8 viser korrelationer mellem læseforståelses- og afkodningsmål ved T1 og læseforståelse ved T2 for de 54 deltagere, der deltog på begge testtidspunkter.

Tabel 8. Korrelationer mellem læseforståelse, fonologisk kodning og ortografisk viden ved T1 og læseforståelse ved T2

	Læseforståelse T1	Fonologisk kodning T1	Ortografisk viden T1
Læseforståelse T1			
Fonologisk kodning T1	0,30*		
Ortografisk viden T1	0,37**	0,17	
Læseforståelse T2	0,57**	0,16	0,06

Note. N = 54. * $p < 0,05$. ** $p < 0,01$.

Af tabellen fremgår det, at læseforståelse ved første og andet testtidspunkt var moderat til stærkt korreleret ($r = 0,57$). Der var altså en høj grad af overensstemmelse mellem scorerne på de to test, men sammenhængen var dog langt fra perfekt. Spørgsmålet var så, i hvor høj grad uoverensstemmelsen mellem de to testresultater kunne skyldes upålidelighed af testene. For at

estimere korrelationen mellem de to test, *hvis* begge tests havde haft perfekt pålidelighed, blev den attenuationskorrigerede korrelationskoefficient beregnet ved hjælp af følgende formel (Rost, 2007).

$$r_{x'y'} = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}}$$

Korrelationskoefficienten korrigeret for attenuation $r_{x'y'}$ fås ved at dividere den observerede korrelation r_{xy} med kvadratroden af produktet af koefficienterne for de to tests kendte pålidelighed, hhv. r_{xx} og r_{yy} . I beregningen af den attenuationskorrigerede korrelation blev testenes interne homogenitet (Cronbach's Alfa) anvendt som estimer for testenes pålidelighed. For læseforståelsestesten ved T1 var Cronbach's Alfa = 0,75. For læseforståelsestesten ved T2 var Cronbach's Alfa = 0,73. Den beregnede attenuationskorrigerede korrelationskoefficient var 0,77. Korrelationen korrigeret for attenuation indikerede dermed, at korrelationen mellem de to test fortsat ikke ville være perfekt, selv om pålideligheden for de enkelte test var det. Det giver nogen plads til at forvente, at faktorer ud over testenes pålidelighed bidrog til ustabilitet på tværs af de to test.

Det næste spørgsmål handlede om, hvorvidt forskelle i afkodningsfærdigheders indflydelse på læseforståelse kunne være en faktor, der bidrog til ustabilitet. Spørgsmålet var, om der er indikationer på, at ustabilitet for profilen 'dårlig forstår' kan have at gøre med, at læseforståelsestestene i forskellig grad er følsom for afkodningsfærdigheder. Faktorer som stillelæsning over for højtlesning, tidspres og støtte til afkodning af ord (der adskiller de to betingelser, som de to læseforståelsestest blev afviklet under) er netop blevet kædet sammen med, hvor følsom en læseforståelsestest er over for afkodningsfærdigheder. Hvis forståelse under den ene betingelse var mere følsom for afkodningsfærdigheder end under den anden, kunne det potentielt medvirke til uoverensstemmelse mellem scorer på de to test – og dermed til ustabilitet for profilen 'dårlig forstår'.

Som det fremgår af tabel 8, var scorer på læseforståelsestesten ved T2 i modsætning til testen ved T1 *ikke* korreleret med fonologisk kodning og ortografisk viden blandt de 54 deltagere. Det tydede på, at afviklingen af læseforståelsestesten ved T2, hvor deltagerne fik hjælp til at afkode ord, virkede efter hensigten, som netop var at minimere afkodningsfærdigheders betydning for læseforståelsesscoren. Afkodningsfærdigheder syntes altså netop at have forskellig betydning for forståelse ved de to testtidspunkter. For at undersøge om dette kunne give anledning til uoverensstemmelse mellem forståelsesscorerne fra T1 og T2, blev der gennemført en partiel korrelationsanalyse, hvor korrelationen mellem læseforståelse ved T1 og T2 blev kontrolleret statistisk for afkodning. Hvis kontrol for afkodning medførte en stærkere sammenhæng mellem læseforståelsesmålene, kunne det indikere, at forskellene i følsomhed for afkodning faktisk bidrog til ustabilitet.

Den partielle korrelation mellem læseforståelsesmålene med afkodning som kontrolvariabel var moderat til stærk, $r = 0,57$, $p < 0,01$ og svarede til korrelationen *uden* kontrol for afkodning ($r = 0,57$). Kontrol for afkodningsfærdigheder ved T1 øgede altså ikke korrelationen mellem læseforståelse ved T1 og T2. Den ikke-perfekte korrelation mellem T1 og T2 læseforståelse syntes dermed at skyldes *noget andet* end afkodningskravene ved T1.

Diskussion

For at vi kan tale om en profil som 'dårlig forstår', og hvis det skal give mening at søge forklaringer på denne profil, så skal profilen være stabil i et vist omfang. Tidligere undersøgelser har fundet, at mønsteret bag profilen er ret stabilt over tid, men der er forskningsresultater, der tyder på, at stabiliteten kan være meget begrænset på tværs af test. En del forskningsresultater tyder på, at forskelle i testindhold og testformat er vigtige faktorer for ustabiliteten. I den aktuelle undersøgelse var det hensigten at undersøge stabiliteten på tværs af en relativt kort periode og under forskellige, men hyppigt sete, *betingelser* og derved bidrage til den generelle undersøgelse af, om 'dårlig forstår' er en stabil profil af en læser. Først blev graden af stabilitet for gruppen af dårlige forståere på tværs af tid og testbetingelse undersøgt, og dernæst blev mulige kilder til eventuel ustabilitet forsøgt indkredset.

Det første spørgsmål lød: 1) *Hvor stor stabilitet er der i klassifikationen af dårlige forståere over tid og over typiske måder at gennemføre læseforståelsestest på (gruppetestning og individuel testning)?* I alt 81% af de deltagere, der blev identificeret som dårlige forståere eller gode forståere ved T1 (gruppetestning), kunne genklassificeres som dårlige eller gode forståere ved T2 (individuel testning). For dårlige forståere specifikt var stabiliteten 59%. Denne stabilitet var lidt lavere end den rapporterede stabilitet over tid hos Catts et al. (2012), men lidt højere end den stabilitet på tværs af forskellige test som Keenan et al. (2014) rapporterede. I hvilken grad den observerede stabilitet kan forventes at afspejle en udvikling over tid eller faktorer forbundet med de to test diskuteres i afsnittet om 'Andre kilder til ustabilitet'.

De næste spørgsmål drejede sig om to mulige kilder til ustabilitet. Det første af disse lød: 2a) *I hvilket omfang kan ustabilitet i klassifikationen skyldes begrænset pålidelighed af læseforståelsestestene?* Estimeringen af korrelationen mellem de to test givet perfekt testpålidelighed viste, at korrelationen mellem de to læseforståelsestest ikke ville være perfekt, selv om testenenes pålidelighed var det. Det syntes altså ikke at være begrænsninger i pålideligheden af læseforståelsestestene, som var hovedårsagen til ustabilitet for profilen 'dårlig forstår'. Andre faktorer må bidrage til ustabiliteten.

Det næste spørgsmål lød: 2b) *Er der indikationer på, at ustabilitet for profilen 'dårlig forstår' kan have at gøre med, at de to betingelser varierer i følsomhed for afkodningsfærdigheder?* De to betingelser varierede i følsomhed for afkodning: T1-testbetingelsen var mere afkodningsafhængig end T2-testbetingelsen. Det kunne hænge sammen med, at T2-testen blev afviklet med afkodningsstøtte. Stillelæsning (som ved T1) har desuden vist sig generelt at medføre højere korrelationer mellem afkodning og læseforståelse end ved højtlesning (som ved T2), og det er ligeledes blevet foreslået, at et tidspres (som ved T1) kan medføre, at læseren ikke kan kompensere for en usikker afkodning, hvilket også ville føre til en højere korrelation mellem afkodning og læseforståelse under tidspres sammenlignet med sammenhængen i en situation uden tidspres. Men der kunne i den aktuelle undersøgelse ikke spores konsekvenser af at kontrollere for betydningen af afkodningsfærdigheder ved T1. Der er altså ikke indikationer på, at ustabiliteten af profilen 'dårlig forstår' kommer af, at de forskellige betingelser varierer i følsomhed for afkodningsfærdigheder. Der er dog et vigtigt forbehold for denne konklusion. Resultatet skal ses i lyset af, at variation i

afkodningsfærdigheder var begrænset i deltagergruppen: De af deltagerne, der var blevet udpeget som dårlige og gode forståere ved første testtidspunkt, var netop relativt stærke afkodere. Den samlede deltagergruppe ved andet testtidspunkt inkluderede dermed kun få deltagere med lave afkodningsscorer (under -1 standardafvigelse).

Andre kilder til ustabilitet

Ustabiliteten på tværs af test kan altså ikke alene forventes at skyldes begrænset testpålidelighed eller forskelle i læseforståelsesmålenes afhængighed af afkodning. Hvilke andre kilder er der da til den observerede grad af ustabilitet? Først og fremmest må regressionseffekten antages at bidrage til ustabilitet. Deltagerne blev udvalgt på baggrund af relativt lave og høje læseforståelsesscorer ved T1, og dermed er undersøgelsesdesignet særligt udsat for regressionseffekt. Regressionseffekten vil netop medføre, at de to grupper af forståere scorer lidt tættere på gennemsnittet ved anden testgang – det vil sige, at det er forventeligt at se, at de nærmer sig hinanden lidt. Det er dog kun gennemsnittet for gruppen af dårlige forståere, der bevæger sig mod gruppens fælles gennemsnit. En medvirkende forklaring på de ret stabile karakteristika ved gruppen af gode forståere kan her være, at de er *mindre* ekstremt udvalgt end de dårlige forståere og derfor også er mindre udsat for regression mod gennemsnittet. Regressionseffektens størrelse er umulig at bestemme i denne undersøgelse, fordi testen af læseforståelse ved de to testtidspunkter ikke var den samme test og ikke gennemført to gange over kort tid. Derfor kan der være mange andre årsager end regressionseffekt til ustabiliteten over tid.

Fra langtidsundersøgelser, der har set på stabilitet over tid (og over meget længere tid, end der her er gået mellem de to test) af bl.a. Catts et al. (2012), fremgår det, at stabiliteten ved testning med samme test/mål over tid er meget høj. Undersøgelsen af Keenan et al. (2014) tyder på, at stabiliteten på tværs af test er meget lavere. Det tyder på, at det hovedsageligt er forskellene på de anvendte test, der er årsag til ustabiliteten i den aktuelle undersøgelse. Det kan for det første være, at de to læseforståelsestest – trods det at de er udviklet med parallelt indhold – ikke stiller helt de samme krav til læserens viden og færdigheder. Selv om hvert parallelt tekst-par på tværs af testene deler genre og omhandler matchende emner, omhandler teksterne netop ikke de *helt samme* emner, og den relevante baggrundsviden for besvarelse af opgaverne i de to test kan dermed variere lidt. På samme måde kan man forestille sig, at de to test trods testudviklernes anstrengelser stiller lidt forskellige krav til komponenter i tekstforståelse som ordforråd og inferensfærdigheder og/eller til generelle forudsætninger som arbejdshukommelse. Sådanne små forskelle kan i teorien bidrage til ustabilitet. Men det er også oplagt, at der må være faktorer, der knytter sig til de meget forskellige betingelser, som testene blev afviklet under, der medvirker til den observerede ustabilitet.

Og hvad kan det så være ved de to betingelser, der kan bidrage til ustabilitet? Et bud er, at de to forskellige betingelser er forbundet med forskellige muligheder for – eller krav om – at udøve selvregulerende adfærd, som beskrevet af Vidal-Abarca et al. (2011). I gruppetesten er det fx muligt at (vælge at) læse spørgsmålene, før man læser teksten. Det er ikke på samme måde muligt at (vælge at) høre spørgsmålene, før man læser teksten i den individuelle test. I den individuelle test skal deltageren i det hele taget ikke strukturere sin opgaveløsning og tage beslutninger i samme grad som under en gruppetest, for der er en anden person, der så at sige strukturerer opgaven og

dikterer forløbet. Forskelle i krav til selvregulerende adfærd kan også omfatte det som Elgart (1978) og Swalm (1972) beskrev som en tvang/ikke tvang til at være opmærksom på alle ord i teksten under henholdsvis højtlesning og stillelæsning, der for svage læsere kunne resultere i bedre forståelse af højt læste tekster end stille læste tekster. Der sås i den aktuelle undersøgelse netop en tendens til, at de dårlige forståere opnåede højere forståelsesscorer ved højtlesning, hvor gennemsnittet for gruppen af gode forståere var stabilt på tværs af testbetingelser.

En anden forskel på de to testbetingelser, der også kan have betydning, er, at deltagerne i gruppetesten besvarede spørgsmål skriftligt, hvor deltagerne i den individuelle test svarede mundtligt. Skriftlig besvarelse stiller i nogen grad krav til skriftlig formuleringsevne. Der er dog i scoringen af besvarelsenerne *ikke* lagt vægt på skriftlig formuleringsevne, ligesom der heller ikke er lagt vægt på den sproglige korrekthed i scoringen af de mundtlige svar fra den individuelle test. Men det at skulle formulere et svar skriftligt kan potentielt kræve flere ressourcer af eleven med svage skriftlige formuleringsevner og dermed påvirke denne elevs mulighed for at demonstrere sin forståelse. Selvregulerende adfærd synes også relevant i denne forbindelse – og i særdeleshed i kombination med en øvre tidsgrænse på testen. Eleven skal afgøre, hvornår det skriftlige svar er tilfredsstillende/godt nok til, at han eller hun vil gå videre til næste opgave eller tekst – og det kan igen få konsekvenser for, hvor langt han eller hun når i testen. Deltagere i læseforståelsestest må forventes at udnytte deres tid forskelligt.

Fremtidige undersøgelser vil kunne undersøge formodningen om, at testbetingelsen spiller en selvstændig rolle for stabiliteten af profilen 'dårlig forstår' (uafhængigt af udvikling over tid) ved at genteste med korte mellemrum. I fremtidige undersøgelser vil man også have mulighed for at isolere de enkelte forskelle på de to testbetingelser, der her er blevet undersøgt samlet, og dermed vurdere betydningen af hvert enkelt aspekt og dernæst eventuelle interaktioner, fx mellem skriftlig/mundtlig besvarelse og tidsgrænse/ingen tidsgrænse. Hvis man har parallelle testmaterialer med meget høj pålidelighed, kan man i nye undersøgelser af stabilitet desuden forsøge at vurdere, hvor stor en rolle regressionseffekten spiller for ustabiliteten, der ses på tværs af testbetingelser: Ved udvælgelse af grupper med lave og høje scorer til gentestning, udvælges der kontrolgrupper efter samme kriterier. Hvor eksperimentgrupperne testes under to forskellige betingelser (med indholdsmæssigt parallelle test), testes kontrolgrupperne under den samme betingelse begge gange. Hvis gentestningen ligger i et par dages forlængelse af den første testning, kan man med god grund udelukke færdighedsudvikling som en årsag til fremgang for dårlige forståere – eller for den sags skyld stagnering/tilbagegang for gode forståere – og en sådan ændring i scorer kan derfor sandsynligvis tilskrives regressionseffekten. For at vurdere betydningen af ændringen i testbetingelsen kan den ændring, der må skyldes regressionseffekten, trækkes fra den forskel, man måtte observere for deltagerne, der blev testet under to forskellige betingelser.

Dårlig forstår eller dårlig læseforståelse under en bestemt betingelse?

Er det rimeligt at tale om "dårlig forstår" som en fast profil for en læser på tværs af situationer, eller er forskellige betingelsers indflydelse så markant, at det ikke giver mening? De lave til moderate grader af stabilitet for dårlige forståere i litteraturen (33-41%) og i denne undersøgelse (59%) kan delvist bruges til at argumentere for, at vi i stedet skal tale om, at en læser har dårlig

tekstforståelse under en bestemt betingelse. Men omvendt viser den korrelation, der ses mellem forståelsesscorer på de to tidspunkter og den grad af stabilitet af profilen 'dårlig forstår', der trods alt er, at der er deltagere, der har lav forståelse i begge betingelser. I denne undersøgelse havde 10 elever, blandt de elever der havde vanskeligheder på gruppetesten (som ikke kunne forklares ved svage afkodningsfærdigheder), også svært ved den individuelle test, der ellers syntes generelt at være lettere end gruppetesten for elever med vanskeligheder. Disse elevers problemer med læseforståelse må forventes at være alt andet en flygtige og vil kræve en omfattende intervention. Resultater fra langtidsundersøgelser tyder som før nævnt også på, at stabiliteten er høj over tid og måske særligt, hvis man afgrænser dårlige forståere ved hjælp af adskillige test. Læseforståelsesproblemer er altså reelle nok, og som hovedregel forsvinder de ikke over tid.

Hvilken betingelse skal man vælge?

Testbetingelsernes mulige betydning for stabiliteten giver også anledning til at spørge, hvilken betingelse man så skal vælge at teste læseforståelse under. Er der fx én af betingelserne, der er lettere/sværere end den anden? Og/eller er der én af de to betingelser, det er mere eller mindre relevant at undersøge læseforståelse under i praksis?

For en del af de deltagere, der ved T1 blev identificeret som dårlige forståere, kan T2-betingelsen se ud til have haft positiv effekt på læseforståelsen, hvilket umiddelbart er i overensstemmelse med resultaterne fra Elgarts (1978) undersøgelse, der tydede på, at en højt-læsningsituation kunne være en fordel for tekstforståelse – og Swalms (1972) undersøgelse, der fandt, at dette kun gjaldt for svage læsere. Men omvendt var der deltagere blandt de elever, der blev identificeret som gode forståere under gruppetestningen, der scorede lavere ved den individuelle test ved T2 – også selv om den gennemsnitlige score i læseforståelse for gode forståere ikke faldt. Det er nærliggende at slutte, at der her er tale om læsere, der har væsensforskellige styrker og svagheder eller præferencer, når det kommer til de betingelser, som sætter rammen for forståelsesopgaven. Man kan dog ikke sige noget generelt om betingelser, der resulterer i fordele og ulemper for elever med enten lave eller høje scorer ud fra denne undersøgelse. Deltagerne var netop udvalgt på baggrund af lave eller høje scorer på testen med stillelæsning og tidsgrænse. Det kan ikke afvises, at et andet mønster vil kunne findes blandt deltagere, der blev udvalgt på baggrund af lave eller høje scorer på den individuelle test højt-læsning. Og igen betyder regressionseffekten, at det er mest sandsynligt, at se dårlige forståere og gode forståere score tættere på gennemsnittet ved anden testgang, uanset hvilken rækkefølge man præsenterer betingelserne i.

Stillelæsning er oftest den måde voksne læser på, og stillelæsning med skriftlig besvarelse, evt. med tidsramme, genfinder man i skolen og fx i eksamenssituationer. Højt-læsning af tekster med krav til forståelsen, der fx sættes i form af efterfølgende forståelsesspørgsmål, finder man også i skolen. Det kan være i lærerstyrede aktiviteter i klassen og i forskellige former for samarbejde, fx gruppearbejde om forståelse og produktion af tekst. Begge testbetingelser kan altså siges at udgøre relevante betingelser for læseforståelse og er dermed relevante ved testning af læseforståelse. Man kunne have forventet, at gruppetesten ved T1 var et mindre pålideligt mål, fordi det ikke med sikkerhed kan vides, om alle deltagere i en gruppetest forsøger at svare på de spørgsmål, de har læst. Med individuelle test har man større garanti for dette. Men testenens interne homogenitet kan

ikke underbygge denne formodning. Koefficienterne for de to test er stort set identiske. Samlet kan man altså ikke sige, at den ene testbetingelse giver et bedre mål/udtryk for læseforståelse end den anden.

Så hvilken testbetingelse skal man vælge? Eller skal man overhovedet vælge? I praksis vil det kunne give værdifuld information om elevens styrker og svagheder, hvis det er muligt at afdække tekstforståelse under forskellige betingelser. I forskning må man gøre sig klart, hvilken betingelse man primært er interesseret i at undersøge forståelse under – eller om man er interesseret i et mål, der dækker flest mulige betingelser for forståelse. I sidstnævnte tilfælde kan man anvende gennemsnit på adskillige test med tilhørende forskellige testbetingelser. Denne fremgangsmåde har den fordel, at den begrænser indflydelse fra målefejl, men medfører også, at man kommer til at udtale sig om en abstrakt gennemsnitlig forståelsesfærdighed på tværs af et (tilfældigt?) udvalg af betingelser. Og hvis man ønsker at sammenligne elevernes tekstforståelse med andre færdighedstest, synes det ligeledes nødvendigt med flere test af disse færdigheder (og i princippet bør disse undersøges under det samme udvalg af testbetingelser som læseforståelse), hvis sammenhængen mellem disse færdigheder og det gennemsnitlige mål for forståelse på tværs af betingelser skal ydes retfærdighed. Multiple test af alle de færdigheder, man er interesseret i, er ressourcekrævende. En anden mulighed synes da også at være at sørge for, at de færdigheder, man tester ud over tekstforståelse, undersøges under den samme testbetingelse, som gælder for læseforståelsestesten – og så holde sig for øje, at resultaterne af denne sammenligning er knyttet til forståelse under denne specifikke betingelse.

Konklusion

Anvendelse af forskellige test med forskelligt indhold og forskellige testformater kan bidrage til betydelig ustabilitet for profilen 'dårlig forstår', men forståelsesvanskeligheder forsvinder som hovedregel ikke, når forståelse testes med de samme materialer over tid. Der er altså grund til at tage forståelsesvanskeligheder alvorligt, og det er rimeligt at tale om en dårlig forstår-profil. Men denne undersøgelse tyder dog på, at det er hensigtsmæssigt at være opmærksom på, om tekstforståelse under visse betingelser udgør særlige problemer for nogle elever. Over en kort periode og på tværs af to forskellige testbetingelser (knyttet til henholdsvis gruppetestning og individuel testning) sås der i den aktuelle undersøgelse en moderat stabilitet for profilen 'dårlig forstår'. På grund af den høje stabilitet over tid, som man har set i tidligere undersøgelser, er det mest sandsynligt, at ustabilitet i denne undersøgelse kan tilskrives forskelle på testbetingelserne (og i nogen grad til testmaterialernes pålidelighed og en regressionseffekt over tid). Den individuelle test syntes at være relativt nemmere end gruppetesten for en del af de dårlige forstaaere (der var blevet afgrænset ved hjælp af gruppetesten). En afgørende forskel på afvikling af typiske gruppetest og individuelle test synes at kunne være de krav, som testbetingelsen stiller til læseren om at udøve selvregulerende adfærd. Selvregulerende adfærd synes i højere grad at være nødvendigt i gruppetest end i individuelle test – fx via kravet om selvstændig stillelæsning og skriftlig besvarelse. Forskelle

mellem læsere her kan muligvis have endnu større konsekvenser for forskelle i læseforståelsesscorer, hvis gruppetesten afvikles med en tidsgrænse.

Kapitel 4: Hvilke aspekter af inferensfærdigheder er vigtige for læseforståelse?

Indledning

En læser er optaget af en tekst om en bestemt type flagermus. I teksten står der bl.a.:

Flagermusen lever hovedsageligt af insekter. For at overleve den lange vinter går flagermusen i hi.

Måske begrundes læseren for sig selv flagermusens behov for at gå i hi om vinteren med fødemangel. Læseren ved nemlig, at insekter ikke findes i rigt mål om vinteren. Læseren har dermed draget en inferens. En inferens er en følgeslutning, der drages med det formål at skabe sammenhæng i forståelsen af teksten. Læseren har i dette eksempel etableret sammenhæng på tværs af tekstens informationer ved at trække på sin baggrundsviden om de ord og forhold, der beskrives i teksten. Ved at drage inferensen tilføjer læseren her information til sin mentale repræsentation af teksten, som ikke er eksplicit beskrevet i teksten. Ikke alle inferenser kræver, at læseren trækker på sin baggrundsviden. Jf. fx følgende korte tekst:

Lottokuponen er i bukserne. Bukserne er i vaskemaskinen.

Her kan læseren drage en inferens om, at lottokuponen er i vaskemaskinen, og det kan læseren i princippet gøre uden at aktivere baggrundsviden om fx lottokuponer, bukser eller vaskemaskiner. Inferensen kan dermed kaldes en *logisk* følgeslutning. Imidlertid er langt de mest almindelige inferenser i tekster, inferenser der kræver aktivering af viden, og jeg fokuserer derfor her og i den følgende empiriske undersøgelse på inferenser, der i én eller anden grad kræver, at læseren aktiverer baggrundsviden.

Færdighed i at drage inferenser har vist sig at kunne forudsige forskelle i børns tekstforståelse. Inferensfærdighedens betydning for læseforståelse kan også påvises, når man kontrollerer for andre sproglige færdigheder og arbejdshukommelse (Cain, Oakhill & Bryant, 2004). I modsætning til en generel forudsætning for tekstforståelse, som fx arbejdshukommelse, er inferensfærdighed desuden en færdighed, der har vist sig at kunne trænes på en måde, så effekten kan aflæses ikke bare i en primær effekt (på inferensfærdigheder), men også i en overføringseffekt til tekstforståelse. Træningsundersøgelser har vist sådanne effekter med elever generelt (Elbro & Buch-Iversen, 2013; Hansen, 1981) og med dårlige forståere (Yuill & Oakhill, 1988; McGee & Johnson, 2003), og meget tyder på, at det også er muligt at forbedre tekstforståelse hos elever med afkodningsvanskeligheder med en systematisk undervisning i at drage inferenser (Daugaard & Elbro, 2013). Den formodning støttes af, at elevers fremgang efter en sådan inferensundervisning

ikke behøver at være afhængig af elevernes afkodningsfærdigheder, ordforråd eller IQ (Elbro & Buch-Iversen, 2013).

Mange klassifikationssystemer til beskrivelse af forskellige inferenstyper er blevet foreslået (fx Graesser, Singer & Trabasso, 1994). Som det vil fremgå nedenfor, finder jeg imidlertid ikke disse klassifikationssystemer særlig klare. Og da der tydeligvis ikke er nogen enighed i læseforskningen om at fokusere på bestemte inferensdistinktioner, betyder det, at de enkelte klassifikationers betydning for læseforståelsen og for dårlige forståeres vanskeligheder er meget sporadisk undersøgt. I stedet for at fokusere på forskellige typer af inferenser, fokuserer jeg i dette kapitel – og i den empiriske undersøgelse, der beskrives senere i kapitlet – på mulige kognitive komponenter i inferensfærdigheder med udgangspunkt i inferenser, der er relevante for praksis (fx i inferensundervisning), og som desuden kan rummes indenfor de forskellige klassifikationssystemer, om end de ikke i alle tilfælde kan placeres entydigt i disse.

Jeg stiller mig for det første kritisk over for den praktiske relevans af en distinktion mellem *automatiske* og *strategiske inferenser*. Ifølge McKoon og Ratcliff (1992) er det kun inferenser, som etablerer sammenhæng mellem de dele af teksten, der processeres samtidigt (holdes samtidigt i arbejdshukommelsen), og/eller beror på tekst-informationer eller viden, som man får meget hurtigt og let adgang til, der bliver draget *under* læsning (såfremt der ikke er et læseformål) – disse inferenser kalder de automatiske inferenser. Har læseren et formål med læsningen, anerkender McKoon og Ratcliff, at også andre (strategiske) inferenser drages under læsningen. Spørgsmålet er dog, hvornår og hvor tit læsere læser uden et formål. Selv forlystelseslæsning og læsning for tidsfordriv synes at have formål (hhv. underholdning og tidsfordriv). Også når læsere deltager i eksperimentelle undersøgelser af læseprocesser synes forskellige mulige formål at presse sig på, selv om disse måske ikke ekspliciteres. Det kan fx være at udføre opgaven bedst muligt ud fra instruktionen eller at få eksperimentet overstået hurtigst muligt. I min undersøgelse af aspekter af inferensfærdighed, der kan være vigtige for læseforståelse, er der fokus på inferensfærdighed i situationer, hvor læseren har et læseformål i form af at skulle løse en eller flere opgaver ud fra teksten, fordi det er langt det mest almindelige, at læsere læser med et formål, og fordi opgaveløsning ud fra tekst er et hyppigt forekommende læseformål for undersøgelsens målgruppe; børn på skolens mellemtrin.

En anden hyppig distinktion er mellem *nødvendige inferenser* (coherence inferences, necessary inferences) og *tillægsinferenser* (elaborative inferences, association-based inferences) (fx Buch-Iversen, 2010; Cain et al., 2001; van den Broek, 1994; O'Brian, Dolores, Myers & Rayner, 1988). Førstnævnte anses for at være nødvendige for tekstens sammenhæng og meningsfuldhed (kohærens), hvorimod tillægsinferenser tilføjer information, der ikke er begrundet i teksten i samme grad. Begge typer inferenser kan drages fra følgende tekst:

Christians far havde samlet en stor bunke tøj på gulvet foran maskinen, da Christian trådte ind af døren. Faren så bebrejdende på Christian og sukkede. Han sagde: "Der er stadig plads". Han åbnede lågen: "Jeg kan tage din trøje med, hvis du skynder dig". Christian gik ind på sit værelse og lukkede døren.

Fra eksempelteksten herover kan man fx drage en nødvendig inferens om den centrale begivenhed i teksten (hvad er det for en aktivitet, der sætter scenen for historien?) og derved skabe sammenhæng i teksten. Inferensen kan begrundes i en række informationer i teksten som *en stor bunke tøj, maskinen, der er stadig plads, åbnede lågen, tage din trøje med*, og der er umiddelbart ikke andre begivenheder, der ville kunne begrundes i samme grad. Hvis læseren ikke drager inferensen 'tøjvask' er det svært at se, hvordan de enkelte dele i teksten, fx farens udsagn, kan give mening. Med udgangspunkt i fx personernes adfærd i ovenstående tekst kan læseren også drage mere usikre inferenser, det vil sige tillægsinferenser. Det kan fx være en inferens om, at Christian er teenager, eller en inferens om, at han gik ind på sit værelse og lukkede døren, fordi han blev sur. Disse inferenser er plausible, og læseren kan henvise til specifikke informationer i teksten, der får ham eller hende til at drage inferensen (*faren så bebrejdende på ham og sukkede, Christian gik ind på sit værelse og lukkede døren*), men der er ikke sikkerhed for inferensen i samme grad som for inferensen om, at den centrale begivenhed er en tøjvask. Kunne Christian fx ikke også være en glad 10-årig fodboldspiller, der går ind på sit værelse og lukker døren for at skifte til en anden trøje?

I min undersøgelse af aspekter af inferensfærdighed i læseforståelse var der fokus på nødvendige inferenser, fordi disse sikrer den basale sammenhæng i forståelsen af teksten, der skal til, for at teksten bliver meningsfuld. Det synes også at være disse inferenser, der må have førsteprioritet i inferensundervisning, der har til formål at forbedre elevernes læseforståelse: Læseren, der drager mange tillægsinferenser (og ikke de nødvendige inferenser), kan nok opnå detaljerede forestillingsbilleder under læsning, men er ikke sikret en sammenhængende forståelse af teksten. Relevansen af forskning i nødvendige inferenser er dermed høj for praksis.

En tredje hyppig distinktion er mellem bagudrettede og fremadrettede inferenser (van den Broek, 1994). Ved bagudrettede inferenser etablerer læseren sammenhæng mellem den aktuelle tekstdel og noget tidligere nævnt i teksten. Ved fremadrettede inferenser forventer læseren en sammenhæng mellem den aktuelle tekstdel og indholdet i den følgende tekst, som fx 'om lidt kommer Christian ud og smider trøjen på gulvet'. Distinktionen mellem bagudrettede og fremadrettede inferenser overlapper delvist med den forrige distinktion mellem nødvendige inferenser og tillægsinferenser, idet fremadrettede inferenser ofte vil være tillægsinferenser. Fremadrettede inferenser kan være ressourcemæssigt uøkonomiske, da de kan vise sig at være fejlagtige, og det kan derfor synes som en god idé at koncentrere sig om andre inferenser end de fremadrettede under læsning. Men selv gode læsere drager dog fremadrettede inferenser, hvis disse er *tilstrækkeligt begrundede* i teksten (Murray, Klin & Myers, 1993).

Distinktionen mellem bagudrettede og fremadrettede inferenser synes ikke at være fyldestgørende for en skelnen mellem inferenserne i ovenstående tekst. Ikke alle inferenserne kan klart siges at have en retning. Hvilken retning har inferensen om en tøjvask for eksempel? Når man stiller spørgsmål for at undersøge, om en læser drager en bestemt inferens, må spørgsmålet desuden også spille en rolle for, hvordan man *beskriver* inferensen ud fra distinktionen mellem bagudrettede og fremadrettede inferenser. Formuleringen af inferenser som fx 'der er tale om et tøjvask-scenarie' og 'om lidt vasker faren tøjet' til teksten ovenover *kunne* således også være affødt af to forskellige spørgsmål som *hvad er det, der foregår?* og *hvad skal der til at ske?* Eller de kunne simpelthen afspejle to måder at formulere den samme inferens på frem for to grundlæggende forskellige inferenser. I min undersøgelse af aspekter af inferensfærdigheder var der derfor fokus på inferenser,

der med rimelighed kan siges at være overbevisende begrundet i teksten frem for fokus på inferensernes retning.

En fjerde hyppigt anvendt distinktion er mellem tekstbaserede og vidensbaserede inferenser (fx McNamara & Magliano, 2009; Kintsch & Rawson, 2005). Denne skelnen kan lede til den (mis)forståelse, at der er mange inferenser med betydning for læseforståelse, der ikke kræver aktivering af viden. Cain og Oakhill (1999) anvendte en distinktion, der afspejler denne klassiske skelnen mellem tekstbaserede og vidensbaserede inferenser. Cain og Oakhill skelnede mellem to slags inferenser (introduceret af Baker & Stein (1978)), nemlig udfyldende inferenser (gap-filling) og brobyggende inferenser (text-connecting). Udfyldende inferenser er ifølge Cain og Oakhill inferenser, hvor information uden for teksten (det vil sige læserens viden) skal flettes med information fra teksten for at udfylde manglende detaljer i teksten. Alle de hidtil gennemgåede inferenser fra eksempelteksten om Christian og hans far kan dermed klassificeres som udfyldende. En brobyggende inferens er derimod ifølge Cain og Oakhill en inferens, hvor information, der er eksplicit beskrevet i to sætninger, integreres. Stein og Baker beskrev brobyggende inferenser som inferenser, der etablerer sammenhæng på tværs af sætninger ('intersentence relationships'), og Cain og Oakhill (2014) beskriver, at den måde, hvorpå der ved brobyggende inferenser etableres sammenhæng mellem sætninger, sker ved en reference til de samme personer/entiteter i disse sætninger.

Hvis læseren ved læsning af eksempelteksten ovenover infererer, at Christians far åbner lågen for at flytte tøjet ind i maskinen, er der i princippet tale om en inferens, der vil kunne kaldes brobyggende i Cain og Oakhills terminologi. Læseren skaber sammenhæng mellem to sætninger, hvori der refereres til samme entitet ved to forskellige ord (*maskinen*, *lågen*). Det synes dog klart, at der er baggrundsviden impliceret i denne inferens. Helt grundlæggende må læseren fx vide, at en maskine (af en bestemt slags) kan have en låge.

Fortolkning af anaforer lever umiddelbart også op til beskrivelsen af brobyggende inferenser, da anaforer indebærer, at et pronomen refererer til en anden person eller entitet i teksten. I eksempelteksten ovenover er det også nødvendigt, at læseren fortolker anaforer for at skabe sammenhæng i teksten. Læseren skal fx fortolke pronominet *han*, som refererende til faren for at skabe sammenhæng mellem de to sætninger:

Faren så bebrejdende på Christian og sukkede. Han sagde: "Der er stadig plads".

Da der er overensstemmelse i køn mellem de to mulige antecedenter *faren* og *Christian*, er det imidlertid ikke helt givet, at det *ikke* kan være Christian, der kommer med udtalelsen i anden sætning, hvis ikke læseren trækker på sin baggrundsviden og sin mentale repræsentation af den forudgående teksts indhold. Selv om faren er i fokus som den handlende aktør i første sætning, må det være baggrundsviden om forældre og børns typiske roller samt den allerede infererede information, at der er tale om en tøjvask etc., der gør faren til den bedst begrundede. Der er endda en tekst-pragmatisk mekanisme, der kunne pege på, at det var Christian, der blev refereret til. *Christian* er den sidstnævnte mulige antecedent, og læsere har tendens til at foretrække tæt reference. Men læserens mentale repræsentation og almene viden giver her belæg for at vælge den lidt fjernere mulighed.

I min undersøgelse af inferensfærdighed i læseforståelse skelner jeg ikke mellem tekstbaserede og vidensbaserede inferenser eller mellem bagudrettede og fremadrettede inferenser. En væsentlig begrundelse er, at disse distinktioner er uklare, og at jeg ønsker at undersøge inferenser, der er relevante for praksis. Reelt er det meget få af de inferenser, som læsere må drage for at få tekster til at hænge sammen, der slet ikke kræver aktivering af viden, og det er ikke altid meningsfuldt at tilskrive en retning til en inferens. Jeg fokuserer på typiske (strategiske) inferenser i tekster, som kan karakteriseres som videnskrævende og som nødvendige i det omfang, at de bidrager til tekstens sammenhæng og meningsfuldhed og kan siges at være *velbegrundede* i teksten. Jeg undersøger så mulige kognitive komponenter, der knytter sig til færdighed i at drage disse inferenser – eller med andre ord: jeg undersøger aspekter ved viden og tekst, som læseren skal kunne anvende for at drage en bestemt inferens. Det er for det første, hvilken *type* viden der skal aktiveres for at drage inferensen, og for det andet, hvorvidt inferensen er sprogligt signaleret i teksten eller ej. Undersøgelsen starter dog med at træde ét skridt tilbage og se generelt på betydningen af den sproglige kode for inferensfærdighedens indflydelse på læseforståelse. Har inferensfærdigheder betydning for læseforståelse, uanset hvilken modalitet de undersøges i (fx sproglig modsat visuel)?

Modalitet

At lytte- og læseforståelse er højt korreleret og forventes at bero på mange af de samme underliggende (sproglige) processer er veletableret viden (fx Gough et al., 1996; Kendeou, Bohn-Gettler, White & van den Broek, 2008; Cain & Oakhill, 1999). Men forståelse kan ifølge Gernsbacher, Varner og Faust (1990) anses for at være generaliseret ikke bare på tværs af talt sprog og skrevet sprog, men også i en vis udstrækning på tværs af sproglig og ikke-sproglig modalitet. Gernsbacher et al. (1990) fandt meget høje korrelationer mellem universitetsstuderendes forståelse af billedhistorier og historier, som deltagerne enten selv læste eller lyttede til – såvel som høje korrelationer mellem de to sproglige mål for forståelse¹⁰. Gernsbacher et al. konkluderede på den baggrund, at forståelse var generaliseret på tværs af sproglig og ikke-sproglig modalitet. De fandt desuden, at konklusionen også gjaldt, når man så specifikt på de opgaver på tværs af historierne, der krævede inferensdragning. De opgaver med billedhistorier, der blev brugt til at teste nonverbal forståelse i undersøgelsen kan imidlertid ikke siges fri for at være at konfunderet med sprog. Svarformatet krævede sproglig processering: Deltagerne skulle svare på mundtligt formulerede spørgsmål (sprogforståelse) og mundtligt formulere deres svar (sprogproduktion). Spørgsmålet melder sig derfor, hvor fast konklusionen om en generel forståelses- og inferensfærdighed på tværs af modalitet er.

¹⁰ De korrelationer, som Gernsbacher et al. anvendte som argument for en generel forståelsesevne, er dog korrigeret for attenuation, og de ukorrigerede korrelationer opgives ikke i artiklen. Det vides således ikke, i hvilken grad den høje korrelationskoefficient for fx sammenhæng mellem forståelse af billedhistorier og forståelse af læste tekster ($r = 0,72$) afspejler en oprindeligt høj koefficient, eller at der er forsøgt taget højde for upålidelige testmaterialer.

Resultater fra en række undersøgelser af børns inferensdragning fra nonverbalt materiale indikerer lige som undersøgelsen af Gernsbacher og kolleger, at inferensdragning *er* generaliseret på tværs af modalitet, men også i disse undersøgelser stiller de nonverbale opgaver i udstrakt grad krav om sproglig processering. Bishop og Adams (1992) testede fx børn med SLI i alderen 8 til 12 (n = 61) og kontrolbørn i alderen 5 til 12 år (n = 50) bl.a. for at se, om børnene med sprogforstyrrelser ville have forståelsesvanskeligheder, hvis stimulimaterialet (historier) blev forelagt nonverbalt i form af billedserier i stedet for verbalt (lytteforståelse). Deltagerne besvarede efter hver historie mundtlige åbne spørgsmål til historien, der enten var inferentielle eller bogstavelige.

Bishop og Adams fandt ingen effekt af den modalitet, stimulusmaterialet blev præsenteret i: Børn med SLI scorede generelt lavere end kontrolbørn på deres egen alder og på niveau med børn, der var 2-3 år yngre end dem selv, og inferensopgaver var generelt sværere end opgaver med bogstavelig information. Da modalitet ikke havde en effekt på scorerne (hverken for SLI-børn eller kontrolbørn) kan det være nærliggende at slutte, at der generelt ikke er forskel på børns evne til at drage inferenser fra verbalt og nonverbalt materiale, og at færdigheden derfor kan være generaliseret. Men det kan ikke udelukkes, at deltageres scorer på tværs af tekst og billedopgaver i undersøgelsen af Bishop og Adams var på niveau, fordi begge forståelsestest reelt var verbale. Selv om stimuli blev præsenteret nonverbalt i billedopgaven, stillede opgaven direkte krav til deltagerens sprog. Deltageren skulle forstå den sprogligt udtrykte formulering af opgaven i form af et spørgsmål (sprogforståelse), og deltageren skulle formulere sit svar verbalt (sprogproduktion). Bishop og Adams noterede sig da også, at nogle af børnenes svar i flere tilfælde indikerede, at de ikke havde forstået det spørgsmål, de skulle besvare. Der er altså væsentlige begrænsninger for, hvad man kan konkludere om børns nonverbale inferenser på baggrund af denne undersøgelse.

Yussen, Rembold & Mazor (1989) undersøgte et aspekt af forståelse, der kan siges at have inferentiel karakter, nemlig hvor gode børn er til at udlede pointen (the main idea) i historier bestående af enten billedserier eller korte tekster, og denne undersøgelse er dermed også relevant for spørgsmålet om inferensfærdighed på tværs af modalitet. Yussen et al. testede børn i 2., 5. og 8. klasse (hhv. n = 17, n = 16 og n = 15) for at undersøge udviklingsmæssige forskelle i evnen til at udlede pointer fra simple historier og sammenlignede også de enkelte gruppers præstationer på opgaverne med billeder og tekst. Deltagerne så og læste en række korte historier. For hver historie løste de en opgave ved at rangordne fire sætninger efter, hvor godt de repræsenterede historiens pointe.

Ældre børn var generelt bedre til at udlede pointer fra historierne, men Yussen et al. (1989) fandt ingen forskel på, hvor gode deltagerne var til at udlede pointer i hhv. billedhistorier og tekst, det vil sige på tværs af modalitet. I udviklingen af materialerne var opgaverne dog blevet udformet og matchet, så voksne deltagere præsterede lige godt (i loftet) på tekster og billeder. En alternativ fortolkning af Yussen et al.'s resultat kan således være, at det ikke kunne påvises, at børnene havde relativt sværere ved at udføre én af disse opgavetyper, *end voksne havde*. Det må være en væsentlig begrænsning for undersøgelsens konklusion om, at det ikke er sværere for børn at udlede pointen fra den ene eller anden slags historie. En anden væsentlig begrænsning er, at Yussen et al. (1989) ikke undersøgte sammenhæng i færdighed på tværs af de to modaliteter på individniveau. Det kunne de have gjort ved en korrelationsanalyse. I stedet nøjedes de med at konstatere, at hver

gruppe på et bestemt klassetrin opnåede samme gennemsnitlige niveau i de to opgavetyper. Hertil kommer, som ved de foregående undersøgelser, forbeholdet for, at opgaven med billeder i høj grad var konfunderet med sprog. I denne undersøgelse var det tilfældet, fordi svarformatet i billedopgaven introducerede tekst i form af sætninger, der skulle læses og vurderes (sprogforståelse). Undersøgelsen giver heller ikke belæg for en sammenhæng mellem inferenser fra billedhistorierne og læseforståelse. Yussen et al. (1989) formodede, at børns evne til at udlede pointer fra billedhistorier (og tekster) var relateret til deres generelle forståelse, men undersøgte det ikke.

Kendeou et al. (2008) undersøgte børns inferensfærdigheder og forståelse af historier i to kohorter. Deltagerne var ved undersøgelsens start hhv. 4 og 6 år ($n = 116$ og $n = 116$). Hver kohorte blev testet to gange med to års mellemrum. Undersøgelsen omfatter dermed data fra børn, da de enten var hhv. 4 og 6 år eller 6 og 8 år. Sammenhængen mellem inferenser og forståelse blev undersøgt på tværs af TV-historier og tekster, som børnene enten lyttede til eller selv læste (sidstnævnte kun for 8-årige deltagere).

TV-historierne, der blev anvendt, var audio-visuelle historier. Det vil sige, at stimulusmaterialet, der blev brugt til at teste inferenser, ud over at være visuelt også inkluderede talt sprog – ligesom de historier, børn ser på TV, som oftest gør. Der er altså ikke tale om en undersøgelse af inferensfærdighed på tværs af sprogligt og ikke-sprogligt stimulusmateriale, men på tværs af medier. Men undersøgelsen er alligevel relevant for spørgsmålet om sammenhæng mellem inferenser og læseforståelse, fordi den i modsætning til de forrige undersøgelser inddrager et mål for læseforståelse og dermed kan give et indtryk af, om inferensdragning fra ikke-tekst kan have betydning for læseforståelse.

Inferensmålene for alle de tre undersøgte medier blev baseret dels på deltagernes mundtlige besvarelse af åbne inferenskrævende spørgsmål stillet af testeren, og dels på deltagernes mundtlige genfortælling af historierne. Man registrerede hvor mange inferenser, der blev udtrykt i disse genfortællinger. Kendeou et al. udledte også tre generelle forståelsesmål fra hvert medie, som de derefter kombinerede til ét mål for forståelse i det pågældende medie. Et af disse mål for forståelse i et medie var det samlede antal begivenheder i historien (story events), som børnene rapporterede i deres genfortælling divideret med antallet af begivenheder i historien. Et andet var børnenes opmærksomhed på historiens kausale forløb (sensitivity to causal structure), som bestod af antallet af de begivenheder i historien, som børnene gengav, der blev vurderet som centrale for historiens kausale forløb (story structure), divideret med antallet af disse begivenheder i historien. Det tredje generelle forståelsesmål bestod af antal rigtige svar på mundtligt stillede spørgsmål til historien divideret med antal spørgsmål¹¹.

Kendeou et al. (2008) fandt signifikante svage til moderate korrelationer mellem alle samtidige inferensmål og forståelsesmål ved både 4, 6, og 8 år. Således korrelerede inferenser fra tv-historier, oplæst tekst og selvlæst tekst alle med læseforståelse ved 8 år (hhv. $r = 0,36$, $r = 0,30$ og $r = 0,30$), og forståelsesmålene korrelerede ligeledes indbyrdes (lytteforståelse og forståelse af tv-historier: r

¹¹ Det fremgår ikke klart, hvorvidt dette tredje generelle forståelsesmål inkluderede de inferensspørgsmål, der også indgik i inferensmålet for samme medie. Hvis det er tilfældet, kan en del af korrelationen mellem det udledte generelle forståelsesmål og det udledte inferensmål i hvert enkelt medie samt korrelationen på tværs af medier potentielt skyldes delvist sammenfald i de items, der indgår i variablene.

= 0,59; lytteforståelse og læseforståelse: $r = 0,59$; forståelse af TV-historier og læseforståelse: $r = 0,49$). Kendeou et al. (2008) fandt også, at inferenser sammenlagt på tværs af de tre medier kunne forklare variation i et kombineret forståelsesmål (tv+lytte+læse), selv om der blev kontrolleret for betydningen af fonologisk opmærksomhed, bogstavkendskab, afkodning og ordforråd. Kendeou et al. (2008) konkluderede, at inferensfærdighed, såvel som forståelse, er generaliseret på tværs af medier. Det kan dog ikke udelukkes, at det har været en medvirkende faktor til sammenhængen mellem færdigheder på tværs af medier, at testformatet i alle tilfælde krævede besvarelse af mundtlige spørgsmål (sprogforståelse og sprogproduktion) og genfortælling af historier (sprogproduktion). Det synes rimeligt, at børns verbale formulering af et hændelsesforløb må korrelere med deres verbale formulering af et andet hændelsesforløb.

Selv i undersøgelser af inferensfærdigheder, hvor stimulusmaterialet til inferenser i nonverbal modalitet har været visuelt materiale i form af billedhistorier uden ledsagende tekst, synes opgaverne, der stilles, ikke at kunne løses uden omfattende sproglig processering. Svarformaterne kræver, at deltageren skal svare verbalt på mundtlige spørgsmål til stimulusmaterialet, genfortælle en historie eller tage stilling til, hvor godt forskellige sætninger matcher pointen i billedhistorien. Man kan sige, at testformatet i den ikke-sproglige opgave på den måde alligevel tvinger deltageren til sproglig processering – og spørgsmålet er så, hvor reel forskellen i modalitet er på en sådan inferensopgave og en inferensopgave, hvor både stimulusmateriale og svarmåden er sproglig. Det synes altså ikke belyst, i hvilken grad der er sammenhæng mellem inferensdragning fra nonverbalt materiale – i en opgave hvor der vel at mærke ikke direkte stilles krav om sproglig processering – og inferensdragning fra tekst. Endvidere er der indikationer på, at inferensdragning fra ikke-tekst kan have forklaringsværdi for læseforståelse, men det er ikke undersøgt, om færdighed i at drage inferenser fra nonverbalt materiale kan have betydning for børns læseforståelse.

Videnstype

Videns betydning for inferensfærdigheder er ikke blot et spørgsmål om at have eller ikke at have den nødvendige viden for en inferens. Barnes og Haefele-Kalvaitis (1996) skelnede mellem, at viden kan være til *rådighed* i varierende grad (availability), og at viden kan være mere eller mindre *tilgængelig* (accessibility). Viden er til rådighed, hvis informationen findes i hukommelsen og kan hentes frem under i hvert fald *nogle* omstændigheder. Tilgængelighed beskriver derimod, hvor hurtigt viden, der er til rådighed, kan hentes frem, og/eller i hvor mange kontekster den kan hentes frem.

I Barnes og Haefele-Kalvaitis undersøgelse indlærte deltagere på 6 til 15 år en vidensbase med karakteristiske træk ved ting, dyr og personer på en fiktiv planet. Dernæst hørte de en historie, der foregik på planeten, og besvarede efterfølgende inferenskrævende spørgsmål til historien. De inferenser, der blev testet, var afhængige af den viden, som alle deltagerne netop havde tillært sig. Det var ikke alle svar på opgaverne, der indgik i de efterfølgende analyser: Det var kun svar på spørgsmål, hvor deltageren efter opgavebesvarelsen kunne genkalde sig den relevante viden (svare på et spørgsmål om den relevante viden). Dermed blev inferensmålet et udtryk for færdighed i at

drage inferenser i de tilfælde, hvor den nødvendige viden med sikkerhed var til rådighed for deltagerne.

Barnes og Haefele-Kalvaitis (eksperiment 1) fandt, at aldersforskelle i inferensdragning ikke blev udlignet, når man sikrede, at alle børn havde den viden til rådighed, der var nødvendig for inferenserne. Cain, Oakhill, Barnes og Bryant (2001) fandt et tilsvarende resultat ved sammenligning af gode og dårlige forståere med samme materialer og fremgangsmåde: Der var signifikant forskel på dårlige og gode forståeres inferensfærdigheder, selv om det blev sikret, at alle deltagere havde den nødvendige viden for inferenserne til rådighed. Forskelle i inferensfærdigheder syntes altså ikke blot at være et spørgsmål om at have eller ikke at have den relevante viden.

Barnes og Haefele-Kalvaitis (eksperiment 2) indhentede et mål for graden af tilgængelighed (accessibility) af den viden, som deltagerne havde til rådighed, og som var nødvendig for inferenserne. Tilgængelighed blev målt ved den tid, det tog deltagerne at påbegynde deres svar på de spørgsmål, der afdækkede, om de havde den nødvendige viden til rådighed. Barnes og Haefele-Kalvaitis fandt, at det var dobbelt så sandsynligt, at viden, der var let tilgængelig, blev anvendt til inferensdragning, som viden, der var mindre tilgængelig (som det tog længere tid at aktivere). Videns tilgængelighed syntes altså at spille en rolle for inferensfærdighederne.

I en række undersøgelser har Shears og kolleger (2004, 2007, 2008) opponeret imod, at det meste forskning i inferenser ikke har skelet til det vidensdomæne, der skal trækkes på ved de inferenser, der sammenlignes. Shears, Miller, Ball, Hawkins, Griggs & Varner (2007) undersøgte, hvor kognitivt krævende det var for voksne at drage inferenser på baggrund af viden om fysiske årsagsforhold (physical knowledge) og viden, der relaterer personers handlinger med formål (planning knowledge). Det gjorde de i en testsituation, hvor deltagerne simultant med den egentlige opgave skulle huske ordlister af forskellige længder.

Deltagerne læste tekster, der lagde op til bestemte inferenser, og matchende kontroltekster, der ikke lagde op til disse inferenser. Inferensfærdighed blev målt indirekte på to måder. Man så for det første på deltagernes falsk-positive genkendelse af ordbetydninger, der var centrale i de pågældende inferenser, men ikke fandtes i teksten overfor genkendelse af mindre centrale ordbetydninger, der fandtes i teksten eller ordbetydninger, der var helt irrelevante. Til sætningen '*Dorothy poured the bucket of water on the fire. The fire went out*' var ordbetydningerne således: *extinguish, put / bucket, water / shelf, plate*. Man så desuden på deltagernes besvarelse af ja/nej-spørgsmål til teksten som: *Is fire extinguished by water?*.

Som eksempel på stimulusmateriale, der skulle eliciteren en inferens på baggrund af viden om fysiske årsagsforhold, gav Shears et al. (2007) det netop nævnte sætningspar:

*Dorothy poured the bucket of water on the fire. The fire went out*¹².

Som eksempel på stimulusmateriale, der skulle eliciteren en inferens på baggrund af viden om, hvordan personers handlinger er relateret til mål, giver Shears et al. følgende sætningspar:

*Malcolm realized Valentine's Day was tomorrow. He went to the candy shop*¹³.

¹² Ord, der forventedes at genkendes (falsk-positivt) på baggrund af en inferens baseret på den relevante viden var *extinguish* og *put*, og det efterfølgende spørgsmål lød: *Does water extinguish fire?*

Shears et al. (2007) fandt, at den ekstra kognitive belastning i form af den simultane opgave med at huske ordlister først og fremmest gik ud over inferensdragning i betingelsen med viden, der relaterede personers handlinger og formål (og kun når ordlisterne var lange). Dette resultat tolkede Shears et al. som en indikation på, at inferenser der kræver viden, som relaterer personers handlinger og mål er kognitivt mere krævende end inferenser, der kræver viden om fysiske årsagsforhold¹⁴. Undersøgelsen af Shears et al. (2007) giver altså et bud på en distinktion mellem videntyper, der kan være afgørende for inferensfærdigheder.

Spørgsmålet er dog, om ikke opgaverne hos Shears et al. (2007) adskiller sig på andre parametre end blot det ene, om den viden, der er nødvendig for inferenserne, er viden om fysiske årsagsforhold eller angår personers formål med handlinger. En høj grad af bekendthed kan forventes, hvad angår vands funktion som ildslukker. Heroverfor synes viden om, at (nogle) folk køber slik (eventuelt til andre) i anledning af Valentinsdag at være af en anderledes specialiseret karakter. Desuden kan sætningerne i det første eksempel meget vel tænkes at elicite en inferens om Dorothy's *formål* med sin handling (hun vil slukke ilden). Shears et al. (2007) kommenterede ikke på disse skævheder mellem de to sætningspar, og de opgav ikke yderligere eksempler på sætningspar i de to betingelser, der kunne indikere, om sådanne skævheder også kunne kendetegne resten af stimulusmaterialet. Det kan derfor ikke udelukkes, at betingelsen med viden om fysiske årsagsforhold *generelt* forudsætter viden, der kan forventes i højere grad at være til rådighed og tilgængelig end viden om personers formål med handlinger (som at købe slik på Valentinsdag), og at denne skævhed kan bidrage til resultatet. Det kan heller ikke udelukkes, at betingelsen med viden

¹³ Ord, der forventedes at genkendes (falsk-positivt) på baggrund af en inferens baseret på den relevante viden, var *buy* og *girlfriend*, og det efterfølgende spørgsmål lød: *Do people buy candy on special days?*

¹⁴ For at bygge bro mellem egne resultater og resultater fra en undersøgelse af Calvo og Castillo (2001) beskriver Shears et al. (2007), hvordan man kan forestille sig en interaktion mellem vidensdomæne og inferensers retning (fremad- eller bagudrettede inferenser). Shears et al. opfatter inferenserne i deres egne opgaver som bagudrettede og drager samtidig paralleller mellem de vidensdomæner, de selv undersøgte, og de vidensdomæner, som de mener, indgik i Calvo og Castillos testmaterialer. Calvo og Castillo (2001) undersøgte ængstelige og mindre ængstelige personers tilbøjelighed til at drage prædiktive (fremadrettede) inferenser fra tekster, der var konstrueret enten med eller uden en potentiel fare. Shears et al. sammenligner deres egne opgaver i betingelsen 'viden der relaterer handlinger og formål' med de opgaver og den viden, der skal trækkes på i de af Calvo og Castillos stimulitekster, der er konstrueret med en potentiel fare, og de sammenligner deres egne opgaver i betingelsen 'viden om fysiske årsagsforhold' med de opgaver og den viden, der skal trækkes på i de af Calvo og Castillos stimulitekster, der ikke er konstrueret med en potentiel fare. Shears et al. skriver, at Calvo og Castillo fandt, at prædiktive inferenser fra tekster med potentiel fare krævede *mindre* tid at drage end inferenser fra tekster uden potentiel fare. Shears et al. opfatter dette resultat som værende i modstrid med deres eget resultat (gældende for bagudrettede inferenser), og det synes at udløse deres hypotese om en interaktion mellem vidensdomæne og inferensens retning. Imidlertid er det uklart, hvor fra i Calvo og Castillos seks eksperimenter Shears et al. henter dette resultat, at fremadrettede inferenser fra tekster med potentiel fare generelt krævede mindre tid at drage. Der synes *ikke* at være en hovedeffekt af stimultype (tekster med potentiel fare vs. tekster uden potentiel fare). Manuel Calvo (email til HTD, 24. april, 2014) skriver: "Upon re-reading (in a rather "oblique" fashion) my article after so many years, I wouldn't say there was a main effect of threat vs. non-threat content. Rather, there was an interaction between threat content and anxiety. This means that there was no advantage of threat or non-threat inferences in general or in absolute terms, but that the advantage was specific (high anxiety, for threat inferences; low anxiety, for non-threat inferences). So I wouldn't agree with the Shears et al.'s interpretation." Shears et al.'s forsøg på at bygge bro mellem de to undersøgelser med en hypotese om interaktion mellem vidensdomæne og inferensensretning synes altså at bero på en misforståelse. Manuel Calvos resultater understøtter ikke en sådan hypotese.

om fysiske årsagsforhold er konfunderet med viden om, hvordan personers handlinger er relateret til formål. Man *kan* derfor forestille sig, at der faktisk var *to* slags viden, der understøttede inferenserne i betingelsen med viden om fysiske årsagsforhold – og at det kunne være en forklaring på, at inferensdragning i denne betingelse var mere robust over for den ekstra kognitive belastning i form af den simultane opgave med at huske ordlister.

Barnes og Häfele-Kalvaitis (1996) skelnede som nævnt mellem *mere eller mindre tilgængelig baggrundsviden*, og Shears et al. (2007) gjorde opmærksom på, at der kan være indsigter at hente i at fokusere på, hvilket vidensdomæne der skal aktiveres i forbindelse med dragning af en inferens. Et alternativ til distinktionen mellem viden om fysiske årsagssammenhænge og viden om personers formål med handlinger, der inkluderer en skelnen mellem mere eller mindre tilgængelig viden, er en distinktion mellem

- personlige erfaringer som er tillært på førstehånd, og som kan formodes at være umiddelbart tilgængelige for læseren
- viden om verden, ofte tillært andenhånds som fx emnespecifik eller akademisk viden, og som i højere grad skal mobiliseres af læseren.

Denne distinktion matcher et stykke ad vejen én af mange forskelle på narrativer og fagtekster. Narrativets forløb følger ofte karakterer og disses handlinger, mål, tanker og følelser. Personlige erfaringer med alment menneskelige forhold kan være et vigtigt grundlag for inferenser med dette indhold.

Fagtekster indbefatter sjældent, at man følger karakterers tanker og følelser indgående, men drejer sig ofte om ”fjernere” emner, hvor den relevante baggrundsviden i højere grad er tillært andenhånds og er emnespecifik. Best, Floyd & McNamara (2008) fandt da også, at viden om verden (i form af en test af akademisk viden om biologi, fysik, historie, geografi, statsledelse (government), økonomi, kunst, musik og litteratur) var særlig vigtig for forståelse af fagtekster i sammenligning med narrativer.

Men der er dog langt fra fuldstændig overensstemmelse mellem overordnet teksttype og den videnstype, der er nødvendig for inferenser fra en tekst. Som det ses af nedenstående to teksteksempler fra den læseforståelsestest, der er anvendt i den aktuelle undersøgelse, kan begge videntyper fx være nødvendige for inferenser fra narrativer.

For to år siden overhørte jeg, at min mor bestilte nogle dyre sko. Min fantastiske lillebror fik dem til sin fødselsdag. Jeg fik kun en skjorte.

Og hvad så? Fodbold er kedeligt.[...]

Personlige erfaringer med at blive slemt skuffet eller forbigået og reagere derpå kan danne grundlag for en inferens om, at fortælleren havde håbet, at skoene var til ham eller hende selv. Sådanne personlige erfaringer er ikke direkte relevante i det følgende eksempel:

Min nabo var vældigt fornøjet med den store jagthund, han havde købt. Den gøede fjendtligt og knurrede af alle, der nærmede sig hans ejendom. [...]

Her kan aktivering viden om verden, nærmere bestemt en viden om hundes mulige funktioner og at folk generelt gerne vil holde uvedkommende væk fra deres ejendomme, danne grundlag for en inferens om naboens formål med at købe sig en hund/en grund til, at naboen var fornøjet med en fjendtlig hund.

Undersøgelsen af Barnes og Haefele-Kalvaitis (1996) tyder på, at tilgængeligheden af den relevante viden generelt har en betydning for sandsynligheden for, at læseren drager en bestemt inferens. Inferenser fra forskellige tekster kan forventes at være afhængige af forskellige typer af viden, og forskellige typer af relevant viden kan netop forventes at være tilgængelig i forskellig grad: For nogle inferenser (ofte i narrativer) kan den relevante baggrundsviden beskrives som førstehånds erfaringer om almenmenneskelige forhold, som må formodes altid at være umiddelbart tilgængelig. For andre inferenser (ofte i fagtekster) er det i høj grad nødvendigt at aktivere andenhånds tillært viden om verden. Det er derfor muligt, at den videnstype, der danner grundlag for inferenser fra den specifikke tekst, spiller en rolle for inferensfærdighedens betydning for forskelle i læseforståelse.

Sproglige signaler om inferens i teksten

Cain og Oakhill (1999) skelner som nævnt mellem to typer af inferenser, der blev introduceret af Baker & Stein (1978): Brobyggende inferenser hvor information, der er eksplicit beskrevet i to sætninger, skal integreres (text-connecting inferences), og udfyldende inferenser hvor information uden for teksten (det vil sige læserens viden) skal inkorporeres med information fra teksten for at udfylde manglende detaljer (gap-filling inferences). Som vist tidligere er dette skel dog ikke rigtig klart, idet næsten alle inferenser med betydning for læseforståelse i én eller anden grad er baseret på viden om verden. Men for mange af de inferenser, der i Cain og Oakhills terminologi kaldes brobyggende, er der dog en anden fællesnævner, der typisk *ikke* karakteriserer de inferenser, der kan kaldes udfyldende. Der er et leksikalt tekstbånd, der kan tjene som signal om inferensen: I eksemplet herunder, som er frit oversat efter eksempelt tekst i Cain og Oakhill (1999), refererer 'appelsinjuicen' til 'en læskedrik' (eller rettere sagt til det samme som ordet 'læskedrik'), og denne relation er afgørende for at inferere, at appelsinjuicen kom fra rygsækken.

Debbie tog en eftermiddag på tur med sin ven Michael. Da de nåede frem, var de meget tørstige. Michael fandt en læskedrik i sin rygsæk, og den delte de. Appelsinjuicen var meget forfriskende. Debbie tog sit badetøj på, men vandet var for koldt, så de byggede sandslotte i stedet.

Den semantiske relation mellem de to ord (læskedrik, appelsinjuicen) er her en relation mellem over- og underbegreb. Relationen kunne fx også være udtrykt ved brug af det samme ord i de to sætninger (læskedrik, læskedrikken) eller synonyme (appelsinjuice, orangesaften). Som det ses, er henvisningen delvist grammatisk kodet, idet en bestemthedsform (appelsinjuicen) dels indikerer, at læseren allerede er stødt på, eller i forvejen kender til, den omtalte appelsinjuice.

Leksikale tekstbånd er et af flere slags tekstbånd, der sprogligt udtrykker sammenhængen i og på

tværs af sætninger. Konnektiver fungerer også som kohæsive tekstbånd. Konnektiver udtrykker, foruden at der er en sammenhæng mellem to sætninger, også at der er en bestemt form for semantisk relation mellem indholdet i sætningerne. Det kan fx være, at sammenhængen består i en årsagsrelation (fordi, derfor, så), en tidsrækkefølge (bagefter, før), et modsætningsforhold (men, derimod) eller at der er tale om en tilføjelse (desuden) til noget allerede nævnt. Sådanne kohæsive tekstbånd kan også fungere som en støtte eller et signal til læseren om at drage en bestemt inferens. Det kan ses i en let omskrevet version af sidste helsætning i Cain og Oakhills eksempel:

Debbie tog badetøj på, men vandet var koldt.

Her kan en velbegrundet inferens være, at Debbie ikke kom ud at bade (trods det, at hun havde et ønske eller hensigt om dette – eftersom hun tog badetøj på), og at grunden var, at vandet var (ubehageligt) koldt. Konjunktionen *men* signalerer nemlig, at der er et modsætningsforhold eller en uoverensstemmelse mellem informationerne i de to sætninger.

En tredje form for tekstbånd (reference) finder man, hvor et pronomen henviser til et substantiv (antecedent) i teksten og dermed bærer substantivets betydning. Det er fx tilfældet i Cain og Oakhills eksempeltekst, hvor *de* henviser til Debbie og Michael, og hvor *den* henviser til den omtalte læskedrik. Netop etablering af reference ud fra et pronomen anser Perfetti (1993) som én af de få inferentielle processer, der må gennemføres obligatorisk og automatisk under læsning. Perfetti sætter spørgsmålstegn ved den opfattelse, at læseren per se drager alle de mulige inferenser fra en tekst under læsning, og videre om der findes en systematik i inferensdragning, som det kan lade sig gøre at afdække. Men nogle inferensers grammatiske grundlag i teksten – i form af et pronomen eller en bestemthedsendelse, som kendetegner leksikale tekstbånd – anser Perfetti altså som det, der adskiller dem afgørende fra andre inferenser. Det grammatisk kodede signal i teksten igangsætter ifølge Perfetti en søgning efter en matchende antecedent i hukommelsen, og processen sker hurtigt og automatisk og er afskærmet for andre kognitive processer. Sanders og kolleger beskriver på lignende vis, hvordan sproglige signaler om sammenhæng i teksten som pronominer og konnektiver indeholder en 'processeringsinstruks' til læseren (Sanders & Pander Maat, 1996; van Silfhout, Evers-Vermeul & Sanders, 2015). Inferenser, der i Cain og Oakhills terminologi kan beskrives som udfyldende, synes som hovedregel ikke at være sprogligt signaleret på denne måde. Når læseren, der læser teksten om Debbie og Michael fx drager en inferens om, at Debbie og Michael var taget til stranden, sker dette på baggrund af læserens kendskab til de ting og aktiviteter, der beskrives i teksten (badetøj, vandet var for koldt, byggede sandslotte), og altså uden at der er en egentlig processeringsinstruks om inferensen til stede i teksten. En inferens om, at Debbie og Michael er børn, er fx også en inferens, der kan drages på baggrund af den aktuelle tekst, og som heller ikke er sprogligt signaleret, men som kan tage udgangspunkt i de beskrevne personers adfærd og viden om den genre, som teksten tilsyneladende hører til.

Sproglige signaler om sammenhæng (kohæsive tekstbånd) er ofte, men ikke altid til stede i teksten, når der er brug for, at læseren drager en inferens, for at teksten opleves som sammenhængende og meningsfuld (kohærent). Kohæsive tekstbånd er bl.a. leksikale tekstbånd og konnektiver.

Spørgsmålet er nu, i hvilken grad færdighed i at drage signalerede og usignalerede inferenser er relateret indbyrdes og til forskelle i børns læseforståelse.

Formål

Hensigten med den følgende empiriske undersøgelse var at belyse tre forskellige aspekter ved inferensfærdighed og betydningen af disse aspekter for læseforståelse. Det første aspekt var modalitet. Det var for det første et formål at undersøge, om man kan tale om en generel inferensfærdighed på tværs af modalitet, og dernæst at undersøge, om modalitet er vigtig for relationen mellem inferensfærdighed og læseforståelse. Følgende forskningsspørgsmål blev stillet:

- 1a) *Er inferensfærdighed fælles for verbal og nonverbal modalitet?*
- 1b) *Spiller modalitet en rolle for inferensfærdigheds betydning for læseforståelse?*

Det andet aspekt var videnstype. Det var her formålet at undersøge, om typen af den baggrundsviden, det er relevant at aktivere i forbindelse med inferensdragning fra en given tekst, er afgørende for relationen mellem inferensfærdighed og læseforståelse. Følgende forskningsspørgsmål blev stillet:

- 2) *Kan typen af den viden, der skal aktiveres ved inferensdragning fra en tekst, være afgørende for relationen mellem inferensfærdighed og læseforståelse?*

Det tredje aspekt var signalering. Formålet var her at undersøge, hvorvidt distinktionen mellem signalerede og ikke signalerede inferenser har en reel betydning. Følgende forskningsspørgsmål blev stillet:

- 3a) *I hvilken grad afspejler færdighed i at drage inferenser med og uden sprogligt signal den samme inferensfærdighed?*
- 3b) *Har både færdighed i at drage inferenser med og uden sprogligt signal en betydning for læseforståelse?*
- 3c) *Har færdighed i at drage inferenser hhv. med og uden sprogligt signal i teksten forskellige roller at spille i læseforståelse?*

Metode

Design

Undersøgelsen var en tværsnitundersøgelse, hvor færdigheder hos et udsnit af elever på 6. klassetrin blev afdækket med individuelle test.

Deltagere

Deltagerne var de 54 6.-klasseelever på forskellige læseforståelsesniveauer, som er beskrevet i kapitel 3.

Materialer

Læseforståelse

Læseforståelse blev undersøgt med den individuelle læseforståelsestest, som blev beskrevet i kapitel 3. For hvert niveau i sværhedsgrad er der både en fagtekst og en narrativ tekst i materialet. Til den aktuelle undersøgelse blev der foruden en samlet sumscore (46 opgaver) også opgjort separate sumscorer for fagtekster og narrativer. En eksperimentel opgørelse af teksterne, hvor den relevante baggrundsviden var hhv. erfaringsbaseret og emnespecifik, blev desuden foretaget og resulterede i to sumscorer, der var baseret på hhv. 28 og 18 opgaver¹⁵.

Nonverbale og verbale inferenser fra narrativer

Nonverbal inferensfærdighed blev undersøgt ved hjælp af opgaver baseret på en samling af opgaver konstrueret af Safarti, Brunet og Hardy-Baylé (2003). Hver opgave bestod af en billedserie (tre billeder). Efter at have studeret billedserien var deltagerens opgave at vælge ét af tre billeder, der kunne være det næste i billedserien. Opgaverne har bl.a. været anvendt i undersøgelser af evnen til at tilskrive intentioner til personer hos voksne patienter med skizofreni (Safarti, Hardy-Baylé, Besche & Widlöcher, 1997). Opgaverne blev i den aktuelle undersøgelse modificeret sådan, at ingen af distraktorerne var aparte i deres udtryk. Fx bestod en distraktor i opgave 6 af et billede af en mand, der ror en båd med en kost. Dette billede blev erstattet med et billede af en mand, der ror en båd med årer (men i en retning, der ikke var i overensstemmelse med den forudgående historie).

¹⁵ Tekster, der syntes overvejende at kræve mobilisering af emnespecifik viden, inkluderede to narrativer og én fagtekst (tekst 1,2,3,4 og 6 i bilagssamlingen). Tekster, der syntes overvejende at kræve involvering af erfaringer/viden om almenmenneskelige forhold, inkluderede to narrativer og tre fagtekster (tekst 5,7, og 8 i bilagssamlingen).

Eksempler på billedserier og valgmuligheder findes i bilag 1. Alle opgaver findes desuden i den selvstændige bilagssamling (bilag C).

En test af *verbal* inferensfærdighed blev udviklet efter samme skabelon som testen af nonverbal inferensfærdighed. Her læste deltageren først et udsnit af en historie (tre sætninger), og skulle dernæst vælge én af tre sætninger, der kunne være den næste sætning i historien. Eksempler på historier og valgmuligheder findes i bilag 2. Alle opgaver findes desuden i den selvstændige bilagssamling (bilag D).

Den nonverbale test og et udkast til 10 opgaver til den verbale test af inferensfærdighed blev piloteret med elever på 5. og 6. klassetrin og blev derefter justeret. Pilotafprøvningen viste (som forventet), at opgaverne var relativt nemme for eleverne, og det blev derfor prioriteret at indhente et hastighedsmål i undersøgelsen, sådan at der kunne udregnes en effektivitetsscore som supplement til opgørelsen af antal rigtige. Opgaverne blev derfor i hovedundersøgelsen afviklet på en bærbar computer ved hjælp af programmet E-prime. I de nonverbale opgaver blev hver billedserie vist i 5 sekunder, før valgmulighederne kom frem på skærmen (den aktuelle billedserie forblev på skærmen, indtil deltageren havde truffet sit valg). Valgmulighederne var markeret som A, B og C, og deltageren tilkendegav sit valg ved at trykke på én af tre knapper (A, B, C) på en tilsluttet responsboks, der kan optage reaktionstider med ned til millisekunders præcision. I de verbale opgaver, hvor deltageren skulle (stille)læse korte historier, blev eksponeringstiden af historierne justeret, så den svarede til den tid, som en sikker læser anvendte til at læse sætningerne (højt) og yderligere 1 sekund. Når de tre valgmuligheder derefter kom til syne, kunne deltageren tilkendegive sit valg (den aktuelle historie forblev på skærmen, indtil deltageren havde truffet sit valg). Den endelige test af nonverbal inferensfærdighed bestod af 26 opgaver, og den endelige test af verbal inferensfærdighed bestod af 24 opgaver. Før gennemførelsen af testene prøvede deltageren 2 øveopgaver og havde mulighed for at stille spørgsmål til disse og den generelle instruktion.

Inferens fra fagtekster

En test af inferensfærdighed fra fagtekster blev udviklet til undersøgelsen. Testen havde et konventionelt design: Deltagerne læste korte tekster højt og besvarede derefter mundtligt inferenskrævende spørgsmål til teksten, som testlederen stillede. Afviklingen af testen blev lydoptaget. Foruden de inferenskrævende spørgsmål indgik der også 10 spørgsmål til information, der kunne hentes direkte i teksten, og hvor inferensdragning altså ikke var nødvendig. Besvarelsen af disse spørgsmål behandles ikke yderligere i denne undersøgelse, men bliver anvendt i undersøgelsen i kapitel 6.

Teksterne tog udgangspunkt i tekster og emner fra fagbøger til 5.-6. klassetrin. Der blev udviklet tekster, der nødvendiggjorde inferenser, og kravene til ordkendskab blev samtidig begrænset, og komplekse sætningsopbygninger blev undgået. De inferenskrævende spørgsmål kunne deles i to kategorier: inferenser, der var sprogligt signalerede ved enten leksikal kohæsion eller et konnektiv, og inferenser, der ikke var sproglig signalerede. Opgaverne med de to kategorier af inferenser signaleret/ikke signaleret blev *ikke* matchet på andre parametre (som fx tekstmængde eller krav til ordforråd). Derfor kan forskelle i resultater mellem de to typer komme af andet en tilstedeværelse eller fravær af signaler om inferenskrav i teksterne. Den sproglige signalering kan til dels siges at

være 'bundet' til inferenser, der ofte beskrives som forskellige i *type* fra de inferenser, der ikke har et sprogligt signal i teksten, og også indbyrdes kan de signalerede inferenser beskrives som forskellige typer af inferenser.

Tekster og spørgsmål blev pilotafprøvet med elever på 5. og 6. klassetrin og videreudviklet af to omgange. Den endelige test indeholdt 22 inferenskrævende spørgsmål (12 opgaver med sprogligt signalerede inferenser og 12 opgaver med ikke-signalerede inferenser). Før eleven gennemførte testen, prøvede eleven at læse en tekst højt og besvare spørgsmål til denne øvetekst og havde mulighed for at stille spørgsmål til instruktionen. Elevernes besvarelser blev scoret af én person ud fra optagelser af testsessionen og udskrifter af dialogen mellem deltager og testleder. Hvis en deltagers besvarelse af et spørgsmål blev anset for korrekt, blev der tildelt 2 point. Hvis en deltagers besvarelse blev anset for at være delvist eller muligvis korrekt, stillede testleder et uddybende spørgsmål efter skabelonen: *Kan du sige lidt mere om det?* Hvis deltagerens uddybende svar blev anset for korrekt, blev der tildelt 1 point. Hvis det uddybende svar ikke blev anset for korrekt, blev der ikke tildelt point. Eksempler på tekster og spørgsmål findes i bilag 3. Alle tekster og spørgsmål, samt eksempler på svar, findes desuden i testlederens spørgeguide, der er gengivet i den selvstændige bilagssamling (bilag E).

Fremgangsmåde

Fremgangsmåden ved testningen er beskrevet i kapitel 3 under fremgangsmåden ved T2. Alle test blev afviklet som individuelle test.

Resultater

Inferensmålenes pålidelighed. Analyser af den interne homogenitet i de eksperimentelle inferenstest viste, at et mindre antal items korrelerede negativt med den eller de skalaer, de indgik i. Items med item-total korrelationer på $-0,05$ og derunder blev ekskluderet, idet der i alle tilfælde kunne identificeres egenskaber ved disse items, der kunne forklare, hvorfor det enkelte item potentielt kunne være upålideligt.

I inferenser fra fagtekster blev item 7 og 12 ekskluderet, og desuden ekskluderedes item 4 og 8, da de adskilte sig fra de andre items i delskalaen med inferenser med signal, som de var en del af. Fx blev item 7 anset for at være upålideligt, da flere elever – både elever med høje og lave scorer – svarede: "Fordi det er sundt" til spørgsmålet: *Hvorfor anbefales det at spise godt med havregryn, spinat og grønkål?* Dette svar blev i overensstemmelse med spørgeguide og scoringsmanual scoret som forkert, da den inferens, som teksten var konstrueret til at elicitere, var inferensen om, at havregryn, spinat og grønkål, lige som mejeriprodukter, indeholder kalk, og at man derfor anbefaler at spise rigeligt af disse grøntsager for at styrke knoglerne. Men besvarelsen er dog helt korrekt ud fra relevant tilegnet viden om verden, hvorfor dette item efterfølgende blev regnet for upålideligt.

I verbale inferenser fra narrativer blev item 2,5,9 og 19 ekskluderet, og i nonverbale inferenser fra narrativer blev item 12 og 24 ekskluderet. Fx blev item 12 i den nonverbale inferenstest ekskluderet, fordi en elev under testningen gjorde opmærksom på, at en vigtig detalje (revner i et vinglas) i denne opgave var svær at se. I opgaven er det afgørende, at man kan se revnerne i vinglasset, for at man overhovedet har mulighed for at inferere, at tjeneren hidkaldes for at udskifte glasset og ikke for at fylde glasset. Alle de 30 elever, der svarede forkert på dette item, havde faktisk valgt den mulighed, at tjeneren skænkede vin i glasset. Det tydede på, at tegningen ikke var tydelig nok, og at dette item derfor var upålideligt.

Enkelte af de resterende items var meget lette. Det vil sige, at 1-2 af deltagerne havde fejl på disse items, og det blev derfor undersøgt, om der var indikationer på, at variationen på disse items først og fremmest skyldtes tilfældigheder. Der var tale om ét item i det samlede mål for inferenser fra fagtekster (og delskalaen med inferenser uden signal), tre items i verbale inferenser fra narrativer, og to items i nonverbale inferenser. Forsøgsvis udelukkelse af disse items resulterede dog ikke i betydningsfulde ændringer i den interne homogenitet for skalaerne. For det samlede mål for inferenser fra fagtekster, og for delskalaen med inferenser uden signal, var Cronbach's alfa den samme med og uden det lette item (hhv. 0,74 og 0,71). For verbale inferenser fra narrativer ændredes Cronbach's alfa fra 0,66 til 0,61 ved udelukkelse af de tre lette items. For nonverbale inferenser fra narrativer ændredes Cronbach's alfa fra 0,71 til 0,69 ved udelukkelse af de to lette items. Der var altså en konsistent kerne af items i skalaerne, og der var ikke indikationer på, at de lette items først og fremmest bidrog med variation, der beroede på tilfældigheder, sådan at udelukkelse af de lette items ville forbedre pålideligheden. De lette items blev bibeholdt i skalaerne.

Behandling af reaktionstider. For verbale og nonverbale inferenser fra narrativer blev der opgjort hastighedsmål til brug for udregning af effektivitetsmål. I få tilfælde betød afbrydelser og lignende under testafviklingen, at reaktionstider for enkelte deltagere på enkelte items ikke var retvisende. Disse reaktionstider blev i første omgang slettet og senere (efter udskiftning af ekstreme reaktionstider) erstattet med den værdi, der svarede til den percentil, deltageren generelt lå på, men justeret for sværhedsgraden af det pågældende item. Dette skete syv gange i tilfældet med nonverbale inferenser og otte gange i tilfældet med verbale inferenser fra narrativer.

Hvor en deltagers reaktionstid på et item lå mere end 2 standardafvigelser over deltagerens middelværdi, blev denne erstattet med værdien, der svarede til 2 standardafvigelser over middelværdien for den enkelte deltager. I testen med verbale inferenser blev der udskiftet 60 sådanne ekstreme reaktionstider, og i testen med nonverbale inferenser blev der udskiftet 62 ekstreme reaktionstider.

Outliers. De udledte effektivitetsscorer og sumscorer for verbale og nonverbale inferenser fra narrativer blev vurderet som funktion af scorerne for det mere konventionelt udformede inferensmål (inferenser fra fagtekster). Outliers blev defineret som deltagere med residualer svarende til ± 3 standardafvigelser. Disse få deltagere (én deltager i testen af verbale inferenser fra narrativer og to deltagere i testen af nonverbale inferenser fra narrativer) er derfor taget ud på de pågældende mål og indgår ikke i resultater med disse mål.

Deskriptiv statistik for alle læseforståelses- og inferensmål fremgår af tabel 1.

Tabel 1. Deskriptiv statistik for læseforståelses- og inferensmål

	Max score	Middel-værdi	Standard-afvigelse	Spænd	α
Læseforståelse					
Samlet	46	36,8	4,5	25-46	0,73
Fagtekst	23	17,6	2,7	10-23	0,58
Narrativ	23	19,0	2,3	13-23	0,52
Emnespecifik viden	28	23,4	2,6	15-28	0,56
Erfaringsbaseret viden	18	13,4	2,6	7-18	0,61
Inferensfærdighed					
Fagtekst					
Samlet	36	25,9	6,0	11-36	0,74
Med signal	16	12,8	2,8	5-16	0,51
Uden signal	20	13,2	4,2	4-20	0,71
Narrativ					
Verbal					
Præcision	20	17,6	1,7	13-20	0,66
Effektivitet		8,2	3,2	3,1-17,6	
Nonverbal					
Præcision	24	20,7	2,3	14-24	0,71
Effektivitet		12,1	4,6	4,5-27,5	

Note. For verbale inferenser i narrativer er n = 53. For nonverbale inferenser i narrativer er n = 52. For øvrige mål er n = 54.

Alle scorefordelinger for målene i tabel 1 havde acceptable værdier for skævhed (inden for $\pm 1,5$). Der var dog en tendens til loftseffekt på målet for inferenser med signal. Skalaernes pålidelighed var generelt acceptabel. For delskalaerne for læseforståelse og for inferenser med signal var pålideligheden dog moderat, hvilket der tages højde for i fortolkningen af resultater, der inkluderer disse mål.

Der var en vis tendens til loftseffekt, og dermed muligvis en begrænset spredning, på målene for antal korrekt besvarede verbale og nonverbale inferenser fra narrativer. Fx gik spændet for verbale inferenser fra 13 til 20 rigtige ud af 20 opgaver, middelværdien var 17,6, og standardafvigelsen var 1,7. Disse test kan derfor have en lidt reduceret følsomhed. De følgende analyser udnyttede derfor effektivitetsmål fra disse test, dvs. antal rigtige pr. minut.

Tabel 2 viser korrelationer mellem målene for læseforståelse og inferensfærdigheder.

Tabel 2. Korrelationer mellem læseforståelses- og inferensmål

	1	2	3	4	5	6	7	8	9
Læseforståelse									
1. Samlet									
2. Fagtekst	0,92**								
3. Narrativ	0,89**	0,64**							
4. Emnespecifik viden	0,88**	0,88**	0,70**						
5. Erfaringsbaseret viden	0,88**	0,74**	0,86**	0,55**					
Inferenser									
Fagtekst									
6. Samlet	0,58**	0,50**	0,56**	0,46**	0,56**				
7. Med signal	0,50**	0,42**	0,49**	0,36**	0,52**	0,77**			
8. Uden signal	0,49**	0,43**	0,46**	0,41**	0,45**	0,91**	0,43**		
Narrativ									
9. Verbal (effektivitet)	0,27*	0,21	0,30**	0,08	0,40**	0,39**	0,23	0,40**	
10. Nonverbal (effektivitet)	0,14	0,12	0,13	0,02	0,22	0,23	0,14	0,24	0,65**

Note. ** $p < 0,01$, * $p < 0,05$. For korrelationen mellem verbale og nonverbale inferenser fra narrativ er $n = 51$. For andre korrelationer med verbale inferenser fra narrativ er $n = 53$. For andre korrelationer med nonverbale inferenser fra narrativ er $n = 52$. For øvrige korrelationer er $n = 54$.

Modalitet

Det første spørgsmål (1a) angik betydningen af modalitet. Det var, hvorvidt inferensfærdigheder er uafhængige af modalitet, dvs. om inferensfærdigheder er nært forbundne på tværs af fx verbal og nonverbal modalitet. Af tabel 2 ses, at verbale og nonverbale inferenser fra narrativer korrelerede stærkt indbyrdes ($r = 0,65$, $p < 0,01$). De to inferenstest havde parallelt indhold og testformat, og de adskilte sig derfor kun på modalitet (verbal vs. non-verbal). Den stærke korrelation tyder dermed på, at inferensfærdighed i vid udstrækning er fælles for verbal og nonverbal modalitet. Verbale inferenser fra narrativer korrelerede desuden moderat og signifikant med målet for inferenser fra fagtekster ($r = 0,39$, $p < 0,01$), et mere konventionelt udformet mål for inferensdragning, hvis indhold og testformat dog adskilte sig væsentligt fra testene af verbale og nonverbale inferenser fra narrativer. Korrelationen mellem nonverbale inferenser fra narrativer og inferenser fra fagtekster var her ikke signifikant ($r = 0,23$, $p = 0,10$)¹⁶.

Det næste spørgsmål (1b) var, om modalitet spiller en rolle for inferensfærdigheds betydning for læseforståelse. Sammenhængen mellem de to parallelle inferensmål, der adskilte sig på modalitet, og læseforståelse fremgår af tabel 2. Verbale inferenser fra narrativer og det samlede mål for læseforståelse korrelerede svagt, men signifikant ($r = 0,27$, $p < 0,05$). Der var derimod ikke signifikant korrelation mellem nonverbale inferenser fra narrativer og læseforståelse ($r = 0,14$, $p > 0,05$)¹⁷.

For at afdække bidraget fra verbal inferensfærdighed til læseforståelse uafhængigt af variationen, der deltes med nonverbal inferensfærdighed, blev der gennemført en hierarkisk regressionsanalyse med det samlede læseforståelsesmål som afhængig variabel. Resultaterne fremgår af tabel 3.

Tabel 3. Resultater fra hierarkisk regressionsanalyse med samlet læseforståelse som afhængig variabel og nonverbale og verbale inferenser som uafhængige variable

Trin	Mål	R^2	ΔR^2	F ændring	β
1	Nonverbal	0,02	0,02	1,17	-0,09
2	Verbal	0,10	0,08	4,32*	0,38*

Note. * $p < 0,05$, ** $p < 0,01$. N = 51.

Verbal inferensfærdighed forklarede fortsat signifikant variation (8%) i læseforståelse ($p < 0,05$), efter der var blevet taget højde for den variation i læseforståelse, der var fælles med nonverbal inferensfærdighed. Så selv om inferensfærdigheder i høj grad korrelerede på tværs af modalitet, så var det *verbal* inferensfærdighed, der havde særlig betydning for læseforståelse: Verbal

¹⁶ Antal korrekte opgaver i testen med nonverbale inferenser korrelerede dog signifikant med inferenser fra fagtekster ($r = 0,40$, $p < 0,01$). Korrelationen mellem antal korrekte *verbale* inferenser fra narrativer og inferenser fra fagtekster nåede dog ikke tilsvarende signifikans ($r = 0,22$, $p = 0,12$, to-halet). Dette skyldes formentlig, at målet for antal korrekte på grund af loftseffekt og begrænset spredning ikke var tilstrækkeligt følsomt.

¹⁷ Hverken antal korrekte verbale eller nonverbale inferenser fra narrativer korrelerede signifikant med det samlede læseforståelsesmål. For målet for korrekte verbale inferenser fra narrativer forventes begrænset følsomhed at være årsag hertil.

inferensfærdighed havde signifikant – og selvstændig – betydning for læseforståelse. Det samme var ikke tilfældet for nonverbal inferensfærdighed.

Videnstype

Spørgsmålet angående videnstypens betydning (2) var, om den type af viden, der skal aktiveres ved inferensdragning fra en given tekst, kan være afgørende for relationen mellem inferensfærdighed og læseforståelse. Betydningen af videnstypen blev undersøgt fra to vinkler: Én hvor læseforståelsesmålet blev opdelt i overensstemmelse med den prædefinerede inddeling i fagtekster og narrativer, og én hvor teksterne forsøgsvis inddeltes i tekster, hvor de nødvendige inferenser overvejende krævede enten emnespecifik eller erfaringsbaseret viden (svarende til viden, der krævede enten mobilisering, eller som var spontant tilgængelig).

De relevante korrelationer fremgår af tabel 2. *Inferenser fra narrativer* henviser i det følgende til målet for verbale inferenser fra narrativer. Først og fremmest ses det, at inferenser fra fagtekster var signifikant korreleret både med forståelse af fagtekster ($r = 0,50, p < 0,01$) og med forståelse af narrativer ($r = 0,56, p < 0,01$). Inferenser fra narrativer var også signifikant svagt til moderat korreleret med forståelse af narrativer ($r = 0,30, p < 0,05$), men var derimod ikke signifikant korreleret med forståelse af fagtekster ($r = 0,21, p = 0,14$). Korrelationen mellem inferenser fra narrativer og forståelse af narrativer var dog ikke signifikant forskellig fra korrelationen mellem inferenser fra narrativer og forståelse af fagtekster ($t = 0,73, df = 50, p > 0,05$).

Ved den alternative opdeling af teksterne i læseforståelsestesten (tekster med vidensbaserede inferenser overfor tekster med erfaringsbaserede inferenser) ses samme mønster, men med større kontrast i forklaringsværdien for inferenser fra narrativer: Inferenser fra fagtekster korrelerer signifikant både med tekster, hvor inferenserne kræver emnespecifik viden ($r = 0,46, p < 0,01$), og med tekster med erfaringsbaserede inferenser ($r = 0,56, p < 0,01$). Derimod var korrelationen mellem inferenser fra narrativer og forståelse med inferenser baseret på emnespecifik viden ikke signifikant ($r = 0,08, p = 0,57$), mens korrelationen mellem inferenser fra narrativer og forståelse af tekster med erfaringsbaserede inferenser var moderat og signifikant ($r = 0,40, p < 0,01$). Denne korrelation var signifikant større end korrelationen mellem inferenser fra narrativer og forståelse af tekster med inferenser baseret på emnespecifik viden ($t = 2,58, df = 50, p < 0,05$).

Den *type af viden*, der skal aktiveres ved inferensdragning fra teksten, syntes således at have afgørende betydning for relationen mellem inferensfærdighed og læseforståelse. Videnstypens betydning forekom delvist indfanget af det traditionelle skel mellem fagtekster og narrativer, men den alternative opdeling af teksterne, baseret på en analyse af den type viden, der overvejende forventedes at være vigtig for inferenserne indikeret i teksten, syntes bedre at indfange betydningen af arten af den viden, der er nødvendig for inferensen.

Sproglig signalering

Betydningen af sproglig signalering af inferenser var det næste forskningsspørgsmål (3a): I hvilken grad er der sammenhæng mellem færdighed i at drage inferenser med og uden sprogligt signal? Af tabel 2 fremgår det, at færdigheder i at drage signalerede og ikke signalerede inferenser fra fagtekster korrelerede moderat indbyrdes ($r = 0,43, p < 0,01$). Korrelationen tyder på, at skalaerne delvist, men langt fra fuldstændigt, var udtryk for den samme inferensfærdighed. Den delvise uafhængighed understøttedes desuden af, at de signalerede og ikke-signalerede inferenser fra fagtekster korrelerede i lidt forskellig grad med verbale inferenser fra narrativer. Ikke-signalerede inferenser fra fagtekster korrelerede moderat og signifikant med verbale inferenser fra narrativer, der netop heller ikke var signaleret i teksterne ($r = 0,40, p < 0,01$). Signalerede inferenser fra fagtekster korrelerede derimod kun svagt med inferenser fra narrativer, og korrelationen nåede ikke signifikans ($r = 0,23, p = 0,09$). Forskellen på disse to korrelationer var dog ikke signifikant ($t = 1,13, df = 50, p > 0,05$).

Det næste spørgsmål (3b) var, om både færdighed i at drage inferenser med og uden sprogligt signal har en betydning for læseforståelse. Af tabel 2 fremgår det, at både færdighed i at drage inferenser med signal og færdighed i drage inferenser uden signal var moderat og signifikant korreleret med det samlede mål for læseforståelse (med signal: $r = 0,50, p < 0,01$; uden signal: $r = 0,49, p < 0,01$). Der var endvidere ikke noget, der tydede på, at enten inferenser med eller uden signal skulle have en forholdsmæssig *særlig* betydning for forståelse af enten fagtekster, narrativer eller tekster med inferenser baseret på emnespecifik viden eller erfaring. Korrelationerne med inferenser med og uden signal var konsekvent af samme styrke (moderate) for forståelse målt på hver af disse skalaer. Både færdighed i at drage inferenser med signal og færdighed i at drage inferenser uden signal synes således at have en betydning for læseforståelse.

Det sidste spørgsmål (3c) var, om færdighed i at drage inferenser med signal og færdighed i at drage inferenser uden signal har *forskellige* roller at spille i læseforståelse. For at belyse dette spørgsmål blev der gennemført to hierarkiske regressionsanalyser med det samlede læseforståelsesmål som afhængig variabel og inferenser med og uden signal i skiftevis 1. og 2. trin i analyserne. Resultaterne fremgår af tabel 4.

Tabel 4. Resultater fra hierarkisk regressionsanalyse med samlet læseforståelse som afhængig variabel og inferenser med og uden signal som afhængige variable

Trin	Mål	R^2	ΔR^2	F ændring	β
1	Inferenser uden sprogligt signal	0,24	0,23	16,29**	
2	Inferenser med sprogligt signal	0,34	0,10	8,03**	
1	Inferenser med sprogligt signal	0,25	0,26	17,45**	0,36**
2	Inferenser uden sprogligt signal	0,34	0,09	7,04**	0,33*

* $p < 0,05$, ** $p < 0,01$.

Både inferenser med signal og inferenser uden signal bidrog med unik variation til læseforståelse. Inferenser med signal forklarede 10% af variationen i læseforståelse efter kontrol for inferenser uden signal, og inferenser uden signal forklarede 9% af variationen efter kontrol for inferenser med signal. De to mål var dog også fælles om en del variation i læseforståelse. Tilsammen forklarede inferenser med og uden signal 34% af variationen i læseforståelse.

Diskussion

Det første aspekt ved inferensdragning, der blev undersøgt, var modalitet.

Det første spørgsmål angående modalitet lød:

(1a) *Er inferensfærdighed fælles for verbal og nonverbal modalitet?* Resultaterne viste, at inferensfærdighed i høj grad er fælles for verbal og nonverbal modalitet. Der var en stærk sammenhæng mellem scorerne på testene af verbale og nonverbale inferenser fra narrativer. Da disse test, på nær hvad angår modalitet, grundlæggende er parallelle inferenstest, tyder den væsentlige fælles variation i disse scorer på, at der er en inferensfærdighed, som er i hvert fald delvis uafhængig af modalitet.

Resultatet er i overensstemmelse med antagelsen fra tidligere undersøgelser om eksistensen af en generaliseret inferensfærdighed (Gernsbacher et al., 1990; Bishop & Adams, 1992; Yussen et al., 1989; Kendeou et al., 2008), og bidrager desuden til en styrkelse af denne antagelse, idet der ikke i den nonverbale inferenstest blev introduceret et tydeligt confound fra sproglige færdigheder, som det har været tilfældet i tidligere undersøgelser. Det kan naturligvis ikke fuldstændigt afvises, at sproglige færdigheder kan bruges som en støtte i løsningen af de nonverbale opgaver, men udformningen af testen sikrede som noget nyt, at deltageren ikke blev afkrævet sprogforståelse og/eller sprogproduktion for at svare på en opgave.

Konklusionen om en inferensfærdighed, der er delvist uafhængig af modalitet, understøttes også af, at målene for både verbale og non-verbale inferensfærdigheder korrelerede med den mere konventionelt konstruerede test af inferensfærdigheder, nemlig testen af inferensfærdigheder fra

fagtekster. Dette belæg er dog noget svagere end den direkte korrelation mellem de parallelle verbale og ikke-verbale inferenstest, bl.a. fordi kun præcisionsmålet og ikke effektivitetsmålet fra den nonverbale inferenstest korrelerede med inferensfærdigheder i fagtekster.

De relativt lave korrelationer mellem narrativ inferensfærdighed og de øvrige mål i undersøgelsen diskuteres nedenfor (under spørgsmål 1b).

Det næste spørgsmål om modalitet lød:

(1b) *Spiller modalitet en rolle for inferensfærdigheds betydning for læseforståelse?* Resultaterne viste, at det stort set kun er den specifikt sproglige inferensfærdighed, der har betydning for læseforståelse. Verbale inferenser fra narrativer korrelerede signifikant med læseforståelse, hvorimod der ikke var tilsvarende konsistente korrelationer mellem nonverbale inferenser og læseforståelse. Verbale inferenser fra narrativer bidrog desuden selvstændigt til læseforståelse efter kontrol for den variation i læseforståelse, der var fælles med nonverbale inferenser.

Et nærliggende spørgsmål er nu, hvorfor der er denne sammenhæng mellem *verbal* inferensdragning og læseforståelse, når der ikke er sammenhæng mellem *nonverbal* inferensdragning og læseforståelse. En oplagt mulighed er, at både verbal inferensdragning og læseforståelse trækker på sproglige ressourcer, fx deltagerens ordforråd. Denne mulighed bliver undersøgt nærmere i kapitel 5, hvor emnet er sammenhænge mellem inferensfærdigheder, ordforråd og arbejds hukommelse i læseforståelse.

Det kan umiddelbart undre, at verbale inferenser fra narrativer ikke var en *stærkere* prædikator af læseforståelse (0,27), særligt i sammenligning med korrelationen mellem inferens fra fagtekster og læseforståelse (0,58). Korrelationen mellem verbale inferenser fra narrativer og det samlede læseforståelsesmål (bestående af både fagtekster og narrativer) synes da også at være lav ved sammenligning med korrelationer mellem inferensfærdigheder fra narrativer og læseforståelse fra andre studier (fx 0,61 blandt 10-årige hos Oakhill & Cain, 2012). Men der er dog andre publicerede eksempler på tilsvarende lave korrelationer: Hos Kendeou et al. (2008) var korrelationen mellem inferenser fra narrativer og læseforståelse (kun narrativ) hos 8-årige børn fx 0,30. En del af forklaringen på den relativt lave korrelation i den aktuelle undersøgelse kan ligge i, at der i tilfældet med verbale inferenser fra narrativer er tale om, at et *effektivitetsmål* for inferensfærdighed sammenholdes med et *præcisionsmål* for læseforståelse. Der var netop hverken noget direkte eller indirekte krav om hastighed i testen af læseforståelse, og den forskellige vægtning af præcision og hastighed i de to mål kan således bidrage til at begrænse sammenhængen mellem præstationerne på de to test¹⁸.

Forskelle i testenenes formater kan også være en del af forklaringen: I testen af verbale inferenser fra narrativer havde opgaverne lukkede svarmuligheder. Dermed var der ikke nogen sikring af, om deltageren aktivt *konstruerede* inferensen – det *kunne* være nok først og fremmest at *genkende* det

¹⁸ Denne antagelse støttes af, at læseforståelsesmålet ved T1 (kapitel 2) korrelerede højere ($r = 0,48$, $p < 0,01$) med effektivitetsmålet for inferensfærdighed, end forståelsesmålet ved T2 gjorde ($r = 0,27$). Forståelsesmålet ved T1 havde netop en tidsgrænse og var således også et effektivitetsmål. Korrelationen med dette læseforståelsesmål var for nonverbale inferenser (ved T2) signifikant ($r = 0,29$, $p < 0,05$). I en supplerende hierarkisk regressionsanalyse forklarede de to effektivitetsmål til sammen 30% af variationen i scorerne på læseforståelsestesten med tidsgrænse. Den supplerende analyse med brug af læseforståelsesmålet fra T1 styrker dog undersøgelsens overordnede konklusion, idet kun det verbale effektivitetsmål for inferensfærdighed selvstændigt kunne forklare variation i scorerne på læseforståelsestesten med tidsgrænse (21,4 % efter kontrol for nonverbal inferensfærdighed).

korrekte svar. At genkende den korrekte inferens kan forventes at være mindre krævende end aktivt at drage den. De tre valgmuligheder i denne test kan også tænkes at have en indflydelse, fordi deltageren i princippet kan anvende information fra distraktorerne til at udelukke svarmuligheder ("det er i hvert fald ikke dét der, der er det rigtige") og på den måde udlede, hvilket svar der *er* det rigtige. I modsætning til testen af verbale inferenser fra narrativer nødvendiggør læseforståelsestestens åbne spørgsmål, at deltageren selv *konstruerer* en forståelse, og herunder aktivt drager de nødvendige inferenser – uden støtte i en eventuel udelukkelsesstrategi.

Korrelationerne mellem narrativ inferensfærdighed og inferensfærdighed fra fagtekst var også relativt lave. Her kan de samme forhold som ovenfor beskrevet (effektivitetsmål over for præcisionsmål og forskelle i testformat) have medvirket til de lave korrelationer. Desuden adskiller testene af inferensfærdighed fra narrativer og inferensfærdighed fra fagtekst sig jo netop på den teksttype, som inferenserne skal drages fra, og dermed adskiller de sig også i deres typiske indhold: Den viden, der skulle aktiveres i hhv. inferenser fra narrativer og inferenser fra fagtekster, kan i høj grad forventes at være forskellig. Måske er inferensfærdigheder i læsning kun i mindre grad et spørgsmål om en *almen* kognitiv færdighed og i højere grad et spørgsmål om, hvilket *indhold* læseren skal drage inferenser med. Det vil sige, at inferensfærdigheder i læsning kan være særlig afhængige af, *hvilket vidensgrundlag* inferensen foregår på.

Spørgsmålet om betydningen af forskelle i videntyper var netop det andet forskningsspørgsmål.

Spørgsmålet lød:

(2) *Kan typen af den viden, der skal aktiveres ved inferensdragning fra en tekst, være afgørende for relationen mellem inferensfærdighed og læseforståelse?* Resultaterne tydede på, at svaret er positivt, dvs. at videntypen *kan* være afgørende for relationen mellem inferensfærdighed og læseforståelse. Der var således en specifik sammenhæng mellem færdigheder i at drage inferenser i narrativer og læseforståelse med tekster, der krævede *erfaringsbaseret* viden – hvorimod der ikke var sammenhæng med læseforståelse med tekster, der krævede *emnespecifik* viden. Og denne særlige betydning af erfaringsbaseret viden kunne forklare, hvorfor inferensfærdigheder med narrativer kun korrelerede signifikant med læseforståelse med narrativer – ikke med fagtekster.

For færdigheder i at drage inferenser fra fagtekster var der derimod en sammenhæng med forståelse med både emnespecifik og erfaringsbaseret viden. Det, som læseren med gode færdigheder i at drage inferenser i *fagtekster*, kan eller gør, ser altså ud til at være vigtigt for forståelse af tekster, uanset typen af viden der skal aktiveres i forbindelse med inferenserne i disse tekster.

Et vigtigt forbehold i fortolkningen af disse resultater er dog den moderate pålidelighed, der kendetegnede alle delskalaerne for forståelse (mellem 0,52 og 0,61). Dette er særligt relevant ved sammenhængen mellem inferenser fra narrativer og forståelse af fagtekster. Det er således muligt, at korrelationen (0,21) kunne have været stærkere og nået signifikansniveauet med et mere pålideligt forståelsesmål. Det samme er i princippet tilfældet med korrelationen med læseforståelse med tekster der krævede emnespecifik viden (0,08), hvilket dog må anses for usandsynligt.

Et andet forbehold kan være de tidligere beskrevne forskelle på de to test af inferensfærdigheder. Det er dog ikke klart, hvordan disse forskelle skulle kunne redegøre for mønsteret i sammenhænge med forståelsesmålene.

Shears et al. (2007) spekulerede i en mulig interaktion mellem vidensdomæne og inferensernes retning. De to anvendte test af inferensfærdigheder kan ikke siges entydigt at adskille sig på inferensernes retning, men der er en overvægt af inferenser i testen med inferenser fra fagtekster, der *kan* beskrives som bagudrettede. Heroverfor *kan* inferenserne i testen af verbale inferenser fra narrativer beskrives som fremadrettede – læseren forudsiger med sit svar en konsekvens af forløbet. Men det må fremhæves, at inferenserne er *tydeligt begrundede* i historierne, og man kan også hævde, at læseren for at træffe sit valg om, hvad der vil ske, ofte først må udlede en persons oplevelse af et problem i situationen eller en anden årsag til personens beskrevne handlinger, og at det er sådanne inferenser, der dernæst danner grundlag for læserens valg af den ene svarmulighed frem for den anden. Som fx i: *Hun har svært ved at holde sig vågen. Hun skynder sig hjem. Hun tager et lagen frem fra en skuffe.* Her kan det rigtige svar (hun reder sin seng) vælges på baggrund af en inferens om, at hovedpersonen ønsker at sove, og at det er det, der motiverer hendes handlinger. Dermed bliver inferensens retning mere uklar. Selv om der kan argumenteres for en overvægt af bagudrettede inferenser i den ene test, og for at inferenserne i den anden test primært er fremadrettede, er det dog som med de andre forskelle på de to inferenstest ikke klart, hvordan forskelle i inferensernes retning skulle kunne redegøre for, hvorfor fx inferenser fra narrativer forudsiger forståelse af tekster med erfaringsbaseret viden, men ikke forståelse af tekster med emnespecifik viden. Endeligt viste den fortolkning af Calvo og Castillos (2001) resultater, som Shears et al. (2007) byggede deres spekulationer om en interaktion mellem vidensdomæne og inferensers retning på, sig da også at være en misforståelse.

Det tredje aspekt ved inferensdragning, der blev undersøgt, var sproglig signalering. Det første spørgsmål angående signalering lød:

(3a) *I hvilken grad afspejler færdighed i at drage inferenser med og uden sprogligt signal den samme inferensfærdighed?* Resultaterne tydede på, at færdighed i at drage inferenser med og uden signal delvist er udtryk for den samme færdighed. Korrelationsanalysen viste en moderat sammenhæng mellem de to slags inferenser fra fagtekster. Desuden var der helt det samme mønster i korrelationerne mellem henholdsvis inferenser med og uden signal og de øvrige mål i undersøgelsen.

Det næste spørgsmål lød:

(3b) *Har både færdighed i at drage inferenser med og uden sprogligt signal en betydning for læseforståelse?* Resultaterne viste, at både færdighed i at drage inferenser med signal og færdigheder i at drage inferenser uden signal synes at have en betydning for læseforståelse. Begge slags inferenser var moderat korreleret med læseforståelse.

Det sidste spørgsmål angående signalering lød:

(3c) *Har færdighed i at drage inferenser hhv. med og uden sprogligt signal i teksten forskellige roller at spille i læseforståelse?* Trods de udprægede ligheder var der resultater, som tydede på, at inferenser med signal og inferenser uden signal til dels spiller forskellige roller i læseforståelse. De to inferensskalaer bidrog nemlig til dels selvstændigt til læseforståelse. På grund af den begrænsede pålidelighed af målet for inferenser med signal, er det et påfaldende resultat, at disse inferenser forudsagde variation i læseforståelse ud over den variation, der blev forudsagt af inferenser uden signal (det mere pålidelige mål). Derimod kan det være en følge af den begrænsede pålidelighed af

målet for inferenser med signal, at færdighed i at drage inferenser uden signal forklarede variation i læseforståelse ud over, hvad der blev forklaret af færdighed i at drage inferenser med signal.

Distinktionen mellem inferenser med og uden signal synes altså at være relevant for læseforståelse. Tilstedeværelsen af det Sanders og Pander Maat (1996) kaldte en processeringsinstruks i form af et sprogligt signal til inferens viste sig i den aktuelle undersøgelse at kunne have en betydning. Færdighed i at drage inferenser med signal har en selvstændig betydning for læseforståelse, og det er sandsynligt, at også færdighed i at drage inferenser uden signal har en selvstændig betydning.

Konklusion

De undersøgte aspekter af inferensfærdighed har alle betydning for læseforståelse. Modalitet er for det første afgørende. Det er den *sproglige* inferensfærdighed, der er vigtig i læseforståelse, ikke den generaliserede inferensfærdighed (selv om den formentlig findes). Der er desuden indikationer på, at forskellige slags inferenser (med og uden signal) delvist har selvstændig betydning for læseforståelse. I særdeleshed er resultaterne stærke for den selvstændige betydning af inferenser, der er signaleret i teksten. Men også videnstypen, som inferenserne i en tekst er afhængig af, er afgørende for inferensfærdigheds betydning for læseforståelse.

Resultaterne synes at kunne nuancere vores viden om relationen mellem inferensfærdighed og læseforståelse – og måske den gængse opfattelse af inferensfærdighedens betydning i forskellige teksttyper. Teksttypen synes ikke i sig selv at være afgørende for, i hvilken grad inferensfærdighed har betydning for forståelse af tekster, for den alternative opdeling af teksterne i læseforståelsestesten, der netop gik på tværs af de prædefinerede genrer, resulterede i et endnu klarere mønster i resultaterne. Mønsteret tyder på, at det afgørende er, hvilken slags viden, der skal aktiveres og udnyttes i inferenserne. Nærmere bestemt, om der er tale om en viden, der er tillært (andenhåndsviden), som relativt sjældent mobiliseres, og som kan være mere eller mindre omstændelig at aktivere, eller om den relevante viden i bund og grund er en personlig erfaring om almenmenneskelige forhold, som ikke kræver decideret mobilisering.

Man kan dog indvende, at der ikke i denne undersøgelse er nogen garanti for, at deltagerne rent faktisk besad den emnespecifikke viden, der var nødvendig i inferenserne. Forskellene kunne derfor være opstået som konsekvens af forskelle i vidensniveauer frem for forskelle i, hvor nemt deltagerne kunne aktivere den nødvendige viden. Imidlertid var der i udviklingen af opgaverne i testen med inferenser fra fagtekster fokus på at holde kravene til den nødvendige viden nede, og således kun kræve viden, som måtte forventes at være almen kendt, også for 12-årige børn. Eksempler på den viden, der er nødvendig i disse opgaver er fx: at træernes blade bliver gule og røde om efteråret, at knogler er hårde, og at gær får brød til at hæve. Gennemsyn af teksterne i læseforståelsestesten giver heller ikke umiddelbart grund til at forvente, at resultaterne kan være opstået udelukkende på grund af forskelle i vidensniveau. Den relevante viden kan fx være, at tyve

som regel vil flygte fra åstedet for ikke at blive opdaget, og at en hund kan fungere som vagthund og holde tyve væk.

At det indhold, som læseren skal drage inferenser med, har en betydning, udelukker ikke at kognitive færdigheder også kan have en betydning. Det er nu oplagt at spørge, i hvilken grad kognitive færdigheder som fx arbejdshukommelsesressourcer spiller ind. At kun den verbale inferensfærdighed (og ikke den nonverbale inferensfærdighed) har betydning for læseforståelse, og at inferenser med sprogligt signal bidrager selvstændigt til læseforståelse, kalder også på en undersøgelse af grundlæggende sproglige ressourcers betydning for inferensfærdigheder. Relationen mellem inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse er emnet i kapitel 5.

Kapitel 5: Samspillet mellem inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse

Indledning

Verbal inferensfærdighed snarere end nonverbal inferensfærdighed synes at have en særlig rolle at spille i læseforståelse (kapitel 4). Men dermed er det ikke klart, om inferensfærdigheder er direkte forbundet med læseforståelse eller blot et udtryk for andre verbale færdigheder af betydning for læseforståelse. For eksempel er ordforråd af notorisk betydning for læseforståelse. Også verbal arbejdshukommelse har vist sig som en robust prædikator af læseforståelse. Samtidig anses inferensfærdigheder netop for afhængige af arbejdshukommelsesressourcer. Man kunne derfor forestille sig, at inferensfærdighedens rolle i læseforståelse kunne tilskrives viden om ords betydning i kombination med en verbal arbejdshukommelseskapacitet, det vil sige en generel evne til at processere verbal information. Nyere forskning viser imidlertid, at der også er andre mulige relationer mellem især ordforråd og inferensfærdigheder i læseforståelse, der må overvejes.

Ordforråd og læseforståelse

At der er en sammenhæng mellem ordforråd og forståelse af tekst virker umiddelbart ligetil: En tekst består af ord, og hvis man ikke kender betydningen af ét eller flere af de enkelte ord, er der risiko for, at man ikke forstår teksten (Nation, 2009). Korrelationen mellem ordforråd og læseforståelse er da også moderat til stærk (fx Oakhill & Cain, 2012), og ordforråd forklarer unik variation i læseforståelse efter kontrol for variation associeret med alder, nonverbal IQ og afkodningsfærdigheder (Ricketts et al, 2007). Et basalt kendskab til ordbetydningerne i en tekst er dog ikke nødvendigvis fyldestgørende for god forståelse. Forståelse synes særligt at være knyttet til, hvor meget man ved om ordene og deres relationer til andre ord (Ouellette, 2006; Cain & Oakhill, 2014)¹⁹. Dette aspekt af ordforråd kaldes ofte for ordforrådets dybde, og det holdes op over for

¹⁹ I en undersøgelse af Tannenbaum, Torgesen og Wagner (2006) fandt man en stærkere sammenhæng mellem bredde og læseforståelse end mellem dybde (i kombination med mål for mobilisering, der dog også til en vis grad testede dybde) og læseforståelse. Imidlertid var én af de to test, der blev anvendt som udtryk for ordforrådets bredde, en test med definitionsopgaver. Definitionsopgaver regnes af andre forfattere netop for at være opgaver, der kan bruges til at undersøge *dybden* af ordforrådet – idet de giver information om, *hvor godt* en deltager kender et ords betydning. Tannenbaum et al. anvendte da også en graderet pointgivning i opgørelsen af scorer på dette mål, sådan at der blev differentieret mellem *kvaliteten* af orddefinitionerne. Dette mål må derfor forventes i betydelig grad at afspejle ordforrådets dybde. Den afhængige variabel hos Tannenbaum et al. var desuden læseforståelse uden kontrol for afkodningsfærdigheder, hvilket også i høj grad kan have medvirket til det resultat, at bredde havde en særlig betydning for læseforståelse. Ouellette (2006) viste netop, at hvor ordforrådets dybde synes at have en særlig betydning for læseforståelse, synes ordforrådets bredde at have en særlig betydning for ordafkodningsfærdigheder.

ordforrådets bredde, det vil sige antallet af ord man kender til. Viden om ords betydning skal desuden også kunne mobiliseres relativt effektivt. Forståelsesprocessen kan risikere at bryde sammen, hvis det tager læseren lang tid at tilgå de enkelte ordbetydninger. Den tid, en læser er om at bedømme, hvorvidt to ord er synonyme, og den tid, det tager læseren at tage stilling til et muligt overbegreb til et ord, har da også vist sig at have selvstændig betydning for læseforståelse ud over afkodning, generel responstid og viden om synonyme og overbegreber (Oakhill, Cain, MacCarthy & Nightingale, 2013). Resultaterne kan tolkes i overensstemmelse med *the lexical quality hypothesis* (Perfetti, 2007), der sætter leksikale repræsentationer i centrum for forståelsesprocessen. Hypotesen går ud på, at ord er lagret som mentale repræsentationer i leksikon. Repræsentationerne indeholder både fonologiske, ortografiske og semantiske komponenter, og læseforståelse er afhængig af kvaliteten af disse leksikale repræsentationer. God læseforståelse kræver, at læseren har leksikale repræsentationer af høj kvalitet, der kan tilgås hurtigt og uden kognitiv anstrengelse.

Kendskab og adgang til de enkelte ordbetydninger er altså en forudsætning for læseforståelse, men forståelse af teksten virker også tilbage på ordforrådet. Læserens (eventuelt ufuldstændige) forståelse af teksten udgør en meningsfuld kontekst, der kan gøre læseren i stand til at tildele betydning til ukendte ord i teksten (Nation, 2005) og derved lære sig aspekter af nye ord. Børn med god forståelse er under nogle omstændigheder således bedre til at slutte sig til nye ords betydning fra konteksten end børn med dårlig læseforståelse (Cain, Oakhill & Lemmon, 2004). Konteksten er dog ikke kun vigtig, når det drejer sig om at tildele betydning til ukendte ord under læsning. Eftersom de fleste ord har flere betydninger, er det konteksten omkring et ord, der afgør, hvilken betydning af ordet der er relevant og præcis i den aktuelle situation, og forståelse af konteksten bliver dermed også en støtte i indkredsning af kendte ords relevante og præcise betydning i teksten (Perfetti, 1985). Sammenlign fx betydningen af ordet *anlæg* i følgende sætninger: *Hans anlæg udviklede sig tidligt* og *hans anlæg stod i kælderen*. De to kontekster, som sætningerne udgør, medfører, at der næppe kan være tale om samme betydning af ordet *anlæg* i de to sætninger. I første sætning er der sandsynligvis tale om anlæg i betydningen *evne*, hvor der i den anden sætning nok nærmere er tale om et teknisk system af en slags, fx et stereoanlæg. Det sidste kunne læseren tillade sig at slutte på baggrund af en efterfølgende sætning som *høj musik var en saga blot*.

Den gensidige påvirkning mellem ordforråd og læseforståelse er påvist mere direkte i langtidsundersøgelser, idet ordforråd og læseforståelse forudsiger hinanden over tid (Tomesen & Aarnoutse, 1998), og selv når der tages højde for den autoregressive effekt af tidlig læseforståelse eller tidligt ordforråd (Wagner, Muse, & Tannenbaum, 2007).

Samspillet mellem ordforråd og læseforståelse kompliceres formentlig yderligere af påvirkning fra andre faktorer. Der er således forskellige hypoteser om fælles bagvedliggende årsagsvariable, det vil sige faktorer, der kan tænkes at være årsag til *både* ordforråd og læseforståelse, og som dermed bidrager til, at disse to korrelerer. Blandt sådanne foreslåede mulige faktorer, der potentielt både kan drive udvikling af ordforråd og læseforståelse, finder man IQ og generel viden (Anderson & Freebody, 1981), metasproglig opmærksomhed (Nagy, 2007) og inferensfærdigheder, der sætter en person i stand til at lære fra konteksten (Sternberg & Powell, 1983).

Arbejdshukommelse og læseforståelse

Arbejdshukommelse er en separat og begrænset hukommelseskapacitet, der muliggør midlertidig lagring og samtidig processering af information, og arbejdshukommelse er blevet forbundet med læseforståelse i en lang række af undersøgelser (fx Daneman & Carpenter, 1980; Cain et al., 2004). Nogle arbejdshukommelsesmål kan da også forklare variation i børns læseforståelse ud over den variation, der kan forklares af de to veletablerede prædiktorer, ordlæsning og ordforråd (Seigneuric, Ehrlich, Oakhill & Yuill, 2000; Oakhill, Yuill & Garnham, 2011). Den form for arbejdshukommelse, der bedst forudsiger læseforståelse, er arbejdshukommelse for verbalt kodet symbolsk information, det vil sige ord og sætninger. Visuo-spatial arbejdshukommelse har fx langt ringere forudsigelsesværdi for læseforståelse (Oakhill et al., 2011; Seigneuric et al., 2000; Caretti, Borella, Cornoldi & De Beni, 2009).

Den præcise karakter af bidraget fra arbejdshukommelse til læseforståelse er ikke fuldt ud kortlagt, men de begrænsede arbejdshukommelsesressourcer forventes at udgøre en flaskehals for en lang række processer i forståelse (Perfetti, Landi & Oakhill, 2005). Præstationer på arbejdshukommelsesmål kan forklare variation i forståelse af lange og komplekse sætninger, hvor læseren netop forventes at have brug for at opbevare ordlyden af sætningen midlertidigt for at kunne genanalysere den undervejs i forståelsen (Gathercole & Baddeley, 1993). Arbejdshukommelse forventes endvidere at have betydning for forståelse på tværs af sætninger og tekstafsnit ved at udgøre en afgørende mental opbevaringsplads, hvor flere informationer kan optræde samtidigt, hvorved det dermed bliver muligt at integrere dem (Kintsch & Rawson, 2005). Det kan fx være nyligt læste informationer, afgørende detaljer i teksten, der blev læst på et langt tidligere tidspunkt, eller viden der hentes frem fra langtidshukommelsen og opbevares midlertidigt, alt imens nye informationer læses (Kintsch & Rawson, 2005; Cain et al., 2004).

Vægtningen af krav til midlertidig lagring og processering i arbejdshukommelsesopgaver er afgørende for, i hvilken grad målet forudsiger læseforståelse. Nogle simple arbejdshukommelsesopgaver kræver primært midlertidig lagring (korttidshukommelse). Deltageren skal typisk huske rækker af præsenterede tal eller andre ord for en kort periode. Heroverfor kræver deciderede arbejdshukommelsesopgaver, at deltageren samtidig foretager en bearbejdning af information, enten af de ord, der lagres, eller af anden information. Det kan være ved, at deltageren skal holde et stigende antal ord i hukommelsen, samtidig med at nye ord læses, eller ved, at deltageren skal foretage en vurdering eller kategorisering af de ord, der lagres midlertidigt (Cantor, Engle & Hamilton, 1991). Cain (2006) fandt, at et mål for korttidshukommelse ikke diskriminerede mellem gode og dårlige forståere, hvorimod arbejdshukommesscorer var relateret til forståerstatus. Dette resultat er i overensstemmelse med resultater fra en lang række undersøgelser, der tilsammen viser, at det især er opgaver, der stiller krav til processeringskomponenten i arbejdshukommelse, der er relateret til læseforståelse. Metastudier viser, at dette både er tilfældet, når det gælder forståelse blandt læsere generelt (Daneman & Merikle, 1996), og når det kommer til at diskriminere mellem gode og dårlige forståere (Caretti et al., 2009).

Inferensfærdigheder og ordforråd i læseforståelse

Betydningen af inferensfærdigheder og ordforråd i læseforståelse er ofte blevet undersøgt separat. I en del tidlige undersøgelser af inferensfærdigheder i læseforståelse har man fx holdt ordforråd konstant ved at sammenligne grupper af gode og dårlige forståere, som blev matchet på ordforrådsmål (fx Cain & Oakhill, 1999). Omvendt har man i undersøgelser, der lægger vægt på dårlige forståeres semantiske vanskeligheder på ordniveau, ikke inddraget mål for inferensfærdigheder (fx Nation et al., 2004).

The lexical quality hypothesis sætter, som tidligere nævnt, ord i form af deres leksikale repræsentationer i leksikon i centrum for læseforståelse. I forlængelse af denne hypotese argumenterer Perfetti og Stafura (2014) for, at ord-til-tekst integration spiller en afgørende rolle i forståelse af sætninger og på tværs af nabosætninger. Parafrase-effekten er et eksempel på en sådan ord-til-tekst integration. Når voksne læsere i ERP studier læser sætningspar som det følgende *While Cathy was riding her bike in the park, dark clouds began to gather, and it started to storm. The rain ruined her beautiful sweater*, har Perfetti og kolleger observeret, at pulsen N400 ved læsning af ordet *rain* i den anden sætning er reduceret i forhold til pulsen i en kontrolbetingelse, hvor den forudgående sætning ikke specificerer begivenheden *storm* (Perfetti, Yang & Schmalhofer, 2008). Perfetti og Stafura tolker reduktionen i N400 som tegn på, at integrationen af ordet *rain* lettes af den tidligere nævnte begivenhed *storm*. De hypotiserer, at denne integration bygger på ubesværet aktivering af en velspecificeret leksikal repræsentation i hukommelsen for ordet *rain*. Den kontekst-relevante betydning af *rain* giver genklang med *storm* i hukommelsen og integreres i læserens mentale repræsentation af teksten. Perfetti og Stafura beskriver også anaforer og leksikalt afhængige brobyggende inferenser som ord-til-tekst-integrationsprocesser, og de antager, at ord-til-tekst integration hos voksne erfarne læsere hovedsageligt består i ubesværede, relativt passive og måske endda automatiske hukommelsesdrevne processer. Bevidst konstruktion, hvor læseren aktivt trækker på sin forståelse af teksten, kan dog blive nødvendig, hvis fx forståelsen bryder sammen.

Li og Kirby (2014) testede aspekter ved både ordforråd og udvalgte højere kognitive færdigheder i læsning, bl.a. inferensfærdigheder, hos 13-årige skoleelever. Deltagernes første sprog var kinesisk, men de modtog en del af deres undervisning i skolen på engelsk, og undersøgelsen fokuserede på deres engelske læseforståelsesfærdigheder. Elever, der ud fra deres afkodningsfærdigheder (på engelsk), nonverbale færdigheder og alder kunne karakteriseres som enten overraskende dårlige forståere, forventeligt gennemsnitlige forståere eller overraskende gode forståere, blev sammenlignet. Det blev sikret, at forskelle i engelsk læseforståelse ikke var relateret til forskelle i afkodningsfærdigheder eller læseforståelse på kinesisk. Li og Kirby fandt, at de overraskende dårlige forståere var signifikant dårligere end gruppen af forventeligt gennemsnitlige forståere på to ud af tre ordforrådsmål (et mål for bredde og et mål for dybde). De overraskende gode forståere adskilte sig til gengæld fra de forventeligt gennemsnitlige forståere ved at være signifikant bedre på tre ud af fire mål for højere kognitive færdigheder i læsning, heriblandt inferensfærdigheder.

Li og Kirbys undersøgelse tager udgangspunkt i læseforståelse under særlige betingelser, i og med at den afdækker komponenters bidrag til læseforståelse på et andet sprog end modersmålet, og tilmed et fremmedsprog, der lå fjernt fra deltagernes modersmål, men resultaterne fra undersøgelsen

kan alligevel indikere et vist hierarki eller en relation mellem komponenter som inferensfærdigheder og ordforråd i læseforståelse generelt. Resultaterne tyder på, at ordforråd ikke i sig selv kan udløse god forståelse, og at inferensfærdigheder måske først kan bidrage til god læseforståelse, når/hvis et ordforråd af et vist omfang og kvalitet er til stede.

Cain og Oakhill (2014) undersøgte sammenhængen mellem forskellige ordforrådsmål og færdighed i at drage to typer af inferenser samt bogstavelig forståelse hos 10-11-årige børn i en tværsnitsundersøgelse. Cain og Oakhill fandt for det første, at ordforråd var en vigtig prædikator for børnenes inferensfærdigheder. Det er i overensstemmelse med adskillige andre undersøgelser, der ligeledes har fundet en sammenhæng mellem ordforråd og inferensfærdigheder (fx Singer, Andrusiak, Reisdorf & Black, 1992; Dixon, LeFevre & Twilley, 1988). De fandt dernæst, at ordforråd ikke forudsagde hukommelse for bogstavelig information fra teksten – et mål der blev udledt fra de samme tekster, som blev brugt til at undersøge inferensfærdigheder. Det er her en vigtig pointe, at ordmaterialet i teksterne var holdt simpelt. Havde et mere varieret og/eller avanceret ordmateriale været anvendt, måtte man forvente en korrelation mellem ordforråd og hukommelse for bogstavelig information i overensstemmelse med de mange undersøgelser, der finder sammenhæng mellem ordforråd og generel læseforståelse. Men det, at svar på spørgsmål til information, der skulle infereres, men ikke svar på spørgsmål til bogstavelig information under de her givne betingelser, korrelerede med ordforråd, kan tolkes som en indikation på, at der kan være en *særlig* relation mellem ordforråd og inferensfærdighed i læseforståelse.

Cain og Oakhill skelnede mellem lokale inferenser (local cohesion inferences) og globale inferenser (global coherence inferences). Der er ifølge Cain og Oakhill tale om lokale inferenser, når betydningen af to sætninger integreres på baggrund af en kobling af to relaterede ord, som i fx: *He fetched a glass of orange juice. The drink was very refreshing.* Her forventes viden om ord og/eller generel viden, i dette tilfælde viden om hvad appelsinjuice er for en størrelse, at være central for integrationen af de to sætninger. Globale inferenser tilvejebringer ifølge Cain og Oakhill derimod information, der ikke er nævnt eksplicit i teksten. Det er fx, når læseren slutter sig til, hvor historien foregår ud fra ord som *sandslotte, padler i vandet og mole*. Globale inferenser synes altså også at trække på ordforråd og generel viden – og måske i endnu højere grad end lokale inferenser, fordi flere semantiske relationer umiddelbart er implicerede. Cain og Oakhills resultater kan tyde på dette. Når der blev kontrolleret for alder og afkodningsfærdigheder, kunne ordforråd – både i form af bredde og dybde – forklare variation i lokale inferenser (hhv. 6% og 5%), men hverken ordforrådets bredde eller dybde forklarede unik variation i konkurrence med hinanden. Tilsammen forklarede de to typer ordforrådsmål 7% af variationen i disse inferenser. Begge typer ordforrådsmål kunne forklare variation i globale inferenser, men kun dybde forklarede unik variation. Tilsammen forklarede de to typer ordforrådsmål 25% af variationen i denne type inferensfærdighed.

Lokale inferenser var signifikant korreleret med bogstavelig forståelse ($r = 0,38$). Når der blev kontrolleret for betydningen af bogstavelig forståelse, bidrog ordforråd ikke længere signifikant til lokale inferenser. Bogstavelig forståelse kunne derimod ikke forklare signifikant variation i globale inferenser ($r = 0,20$), og dybde vedblev med at bidrage med unik variation efter kontrol for bogstavelig forståelse. Det kan umiddelbart undre, hvorfor ordforråd ikke forklarede variation i lokale inferenser, når der kontrolleres for bogstavelig forståelse. Dette resultat er, som Cain og

Oakhill også påpeger, overraskende, da der i disse opgaver sættes krav til at identificere en forbindelse mellem et ord og et kategoriexemplar eller et synonym. En mulig årsag til ordforrådets begrænsede betydning for lokale inferenser i undersøgelsen synes dog at kunne være, at ordmaterialet i inferenstesten var holdt simpelt med målgruppen for øje. Det relativt simple ordmateriale kan måske have begrænset forskelle i præcision i genkendelse af ordenes betydning og dermed også have begrænset den betydning, disse forskelle kunne have for lokale inferenser. Netop ved lokale inferenser syntest ordforrådets bredde (altså det at man har kendskab til ordet i én eller anden grad) jo at være mindst lige så god en prædiktør som ordforrådets dybde (graden af kendskab til ordet). Ifølge Cain og Oakhill kan det have været afgørende, at der ikke blev indhentet mål for hastighed. Hastighed, hvormed et ord kan mobiliseres, har netop vist sig at bidrage selvstændigt til læseforståelse (jf. Oakhill et al., 2013; Richter, Isberner, Naumann & Neeb, 2013). Det er muligt, at relationen mellem ordforråd og inferenser i høj grad afhænger af hastigheden, hvormed læseren aktiverer ordene. Calvo (2004), som blandt andet sammenlignede betydningen af leksikal adgang og præcision i viden om ord for prædiktive inferenser i et eye-tracking forsøg med voksne, fandt dog imidlertid, at den tid, det tog deltagerne at aktivere ords betydning, det vil sige få leksikal adgang, var associeret med forståelse generelt, hvorimod præcision i viden om ords betydning bidrog unikt til inferensdragning.

Cromley og Azevedo (2007) undersøgte samspillet mellem fem faktorer i læseforståelse: ordforråd, inferensfærdigheder, baggrundsviden, brug af strategier (som summary og question generation) og afkodning i læseforståelse. Deltagerne var i gennemsnit 14 år. Ordforråd blev afdækket med en test, hvor deltagerne matchede skrevne ord, der kunne være synonymer. Inferensfærdigheder blev undersøgt med en multiple choice-test. Med hensyn til relationen mellem ordforråd og inferensfærdighed forventede Cromley og Azevedo (2007), at både ordforråd og inferensfærdigheder ville have en selvstændig effekt på læseforståelse, og at ordforråd desuden kunne have en indirekte effekt på læseforståelse gennem inferensfærdigheder. De tilskrev en sådan forventet effekt, at betydningen af enkelte ord i en tekst kan være centrale for nogle af de inferenser, der er nødvendige for at opnå en fyldestgørende forståelse af teksten. Hvis betydningen af et sådan centralt ord er ukendt for læseren, kan læseren ikke drage den nødvendige inferens og dermed heller ikke opnå forståelse. Eksperimentelle studier, blandt andet en effektundersøgelse af Medo og Ryder (1993), bestyrkede desuden Cromley og Azevedo i en forventning om en forbindelse fra ordforråd til inferensfærdigheder. I Medo og Ryders undersøgelse fik deltagerne i en eksperimentgruppe undervisning i specifikke ordbetydninger og disse ords relationer til andre ord. Undervisningen forbedrede deltagerne forståelse af kausale forhold i en tekst med de aktuelle ord i forhold til en kontrolgruppe, og nogle af de items, hvorpå effekten blev målt, kunne siges at være inferentielle. Cromley og Azevedo (2007) fandt effekt af både ordforråd og inferensfærdighed på læseforståelse. En ændring på 1 standardafvigelse i ordforråd svarede til en ændring på 0,37 standardafvigelse i læseforståelse uafhængigt af de andre variable. En ændring på 1 standardafvigelse i inferensfærdighed svarede til 0,19 standardafvigelse i læseforståelse uafhængigt af de andre variable. Der var som forventet en signifikant, men lille, indirekte effekt af ordforråd via inferensfærdigheder. Den indirekte effekt svarede til en ændring på 0,04 standardafvigelse i læseforståelse ved en ændring på 1 standardafvigelse i ordforråd.

Inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse

Cromley og Azevedo (2007) inddrog ikke arbejdshukommelse i undersøgelsen af relationen mellem inferensfærdigheder og ordforråd. Arbejdshukommelse er dog en potentiel interfererende variabel i afdækningen af relationen mellem inferensfærdigheder og ordforråd, da arbejdshukommelse, som tidligere beskrevet, er en etableret prædikator af læseforståelse, og der endvidere er fundet sammenhæng mellem arbejdshukommelse og inferensfærdigheder i en lang række undersøgelser (fx Singer et al., 1992; Calvo, 2004). Teoretisk synes der desuden at være konsensus om, at inferensfærdigheder til en vis grad er *afhængige* af arbejdshukommelsesressourcer (Cain, 2006; Daneman & Merikle, 1996). Cain (2006) beskriver, hvordan højere kognitive færdigheder i læsning som inferensfærdigheder (og monitorering) indebærer, at noget information lagres midlertidigt, samtidig med at anden information processeres, og dermed stiller krav til arbejdshukommelse. Det kan være betydningsindholdet af en sætning, der holdes (tilgængeligt/aktivt) i hukommelsen, samtidig med at den næste sætning læses eller samtidig med, at der søges efter relevant viden i langtidshukommelsen. Ifølge Singer et al. (1992) understøtter arbejdshukommelse færdighed i at drage brobyggende inferenser på tværs af sætninger dels ved at holde information fra den foregående sætning tilgængeligt og dels ved at være involveret i søgning efter og indsættelse af afgørende information fra langtidshukommelsen, der tidligere er blevet nævnt i teksten.

Eftersom arbejdshukommelse forudsiger læseforståelse og også anses for at være afgørende for inferensfærdigheder, kan man spørge, om inferensfærdigheders betydning for læseforståelse helt kan føres tilbage til arbejdshukommelsesressourcer. Arbejdshukommelse synes dog ikke i sig selv at kunne forklare bidrag fra højere kognitive færdigheder i læsning (inferensfærdigheder og monitorering) til børns læseforståelse fuldt ud (Cain, Bryant & Oakhill, 2004).

Pérez, Paolieri, Macizo & Bajo (2014) undersøgte arbejdshukommelsens betydning for hhv. bogstavelig forståelse og tre forskellige slags inferenser. De fandt for det første, at arbejdshukommelsesniveau ikke var associeret med bogstavelig forståelse. Dette understøtter altså antagelsen om en særlig relation mellem arbejdshukommelse og inferensfærdigheder, men er dog i konflikt med tidligere undersøgelsers resultater (fx Masson & Miller, 1983). Pérez et al. (2014) fandt, at arbejdshukommelse syntes at have en særlig betydning for to af de tre undersøgte inferenstyper: inferenser der forbandt et udsagn med en årsag eller forklaring, og inferenser, der foregreb en konsekvens af et udsagn. Nærmere bestemt forudsagde arbejdshukommelse præcision i færdighed i at drage den første type inferenser og hastighed i dragningen af den anden type inferenser. Pérez et al. (2014) fandt derimod ikke en sammenhæng mellem arbejdshukommelsesniveau og færdighed i at drage inferenser, som de betegnede som associative i natur. Det kunne fx være inferenser, hvor læseren på baggrund af et ords betydning og sin baggrundsviden skulle inferere fx et sted. Pérez et al. (2014) fremhævede, at forbindelsen mellem arbejdshukommelse og inferensfærdigheder synes særligt at gælde inferenser, hvor forskellige informationer fra teksten netop skal opbevares og kombineres. Ved inferenser, der i bund og grund er associationer, er der ikke behov for at forbinde forskellige informationer fra teksten på samme måde, og dermed er der heller ikke på samme måde krav til arbejdshukommelsesressourcer.

Det er her værd at bemærke, at der til Pérez et al.'s definition af associative inferenser hører inferenser, der i høj grad er baseret på viden om ordbetydning. To nyere undersøgelser viser, at

ordforråd under nogle omstændigheder kan mediere indflydelse fra arbejdshukommelse til børns inferensfærdigheder. Ordforråd har vist sig helt at kunne mediere indflydelse fra arbejdshukommelse til de inferenser, som børn drager fra historier, som de lytter til (Currie & Cain, 2015), og ordforråd ser desuden ud til i nogen grad at kunne mediere indflydelse fra verbale arbejdshukommelsesmål til en række forskellige læseforståelsesmål, herunder mål for inferentiell forståelse hos børn (Chrysochoou, Bablekou & Tsigilis, 2011). En mulig tolkning af disse resultater kan være, at arbejdshukommelse er involveret i mobilisering af ordbetydninger fra langtidshukommelsen, som læseren endnu ikke har velspecificerede leksikale repræsentationer for, og som derfor ikke spontant lader sig aktivere, og at det er denne funktion, der primært udgør arbejdshukommelsens rolle i inferensfærdigheder. Shankweiler og van Dyke (2013) går skridtet videre og argumenterer for en langt mere begrænset indflydelse fra arbejdshukommelse til forståelse generelt, end hidtil antaget. De tager afsæt i nyere teorier om hukommelse, der problematiserer en separat hukommelsesfunktion for midlertidig lagring af information, og i det faktum, at variationen i antallet af informationer, der kan holdes i arbejdshukommelsen er begrænset, og dermed ikke virker som en sandsynlig årsag til store individuelle forskelle i forståelse. Shankweiler og van Dyke (2013) peger i stedet på leksikale repræsentationers kvalitet og distinkthed i langtidshukommelsen som en faktor, der er afgørende for, om information lagret i (langtids)hukommelsen (gen)aktiveres af de relevante clues i teksten.

Afdækning af bidrag fra hhv. ordforråd og verbal arbejdshukommelse til læseforståelse såvel som til inferensfærdigheder kompliceres grundlæggende af, at verbale arbejdshukommelsestest indebærer sproglig processering, og der er dermed risiko for, at arbejdshukommelsestest reelt tester komponenter i sprogforståelse som ordforråd og inferensfærdigheder.

Opgaver, der inkluderer opbevaring og manipulation af ord, vil altid i nogen grad kunne risikere indirekte at teste aspekter ved ordforråd. Nation, Adams, Bowyer-Crane og Snowling (1999) har fx vist, at nogle børns dårlige præstationer på test af verbal arbejdshukommelse formentlig kan forklares af semantiske vanskeligheder, der hæmmer evnen til at lagre verbal information i korttidshukommelsen. I nogle opgavetyper er kravene til semantisk viden desuden mere udtalt end andre. Den opgavetype, der hos Seigneuric et al. (2000) var den stærkeste prädiktor af læseforståelse, må fx anses for i høj grad at involvere aktivering af semantisk viden. Deltageren skulle tage stilling til, hvilket af fire ord der ikke var relateret til de andre, og deltageren skulle dernæst huske dette ord, mens han eller hun tog stilling til fire nye ord. Opgaven krævede altså, at deltageren tænkte over ligheder og forskelle mellem de fire ordbetydninger. Forskelle i deltagernes ordforråd (ordforrådets dybde afdækket ved kategorividen samt hastighed, hvormed ordbetydningerne kan aktiveres) kan derfor være medvirkende til den stærke sammenhæng mellem dette mål og læseforståelse.

I andre verbale arbejdshukommelsestest kan man endvidere forestille sig, at også inferensfærdigheder testes indirekte. I en anden opgave hos Seigneuric et al. (2000) og Oakhill et al. (2011) skulle deltageren fx færdiggøre sætninger på baggrund af konteksten, som de ufærdige sætninger udgjorde, og dette ord skulle deltageren så huske, mens han eller hun læste/hørte en ny sætning, som også skulle færdiggøres på baggrund af konteksten. Opgaveløsningen inkluderede dermed en række simple inferenser. Oakhill et al. (2011) beskriver, hvordan man ved pilotering af materialerne sikrede, at stort set alle børn kunne færdiggøre sætningerne uden fejl. Det udelukker

dog ikke, at børn, der har haft mere eller mindre let ved at drage disse simple inferenser, potentielt kunne have mere eller mindre let ved at genkalde sig de infererede ord ved opgavens afslutning af netop den grund.

Få undersøgelser gør noget ved det potentielle problem med indflydelse fra sprogforståelseskompontener til verbale arbejdshukommelsesmål, der fx gør det vanskeligt at vurdere det reelle bidrag fra arbejdshukommelse til læseforståelse i sammenligning med fx ordforråd og inferensfærdigheder. Én måde at tackle problemet på er statistisk kontrol som hos Oakhill et al. (2011), der kontrollerede statistisk for generelle sproglige færdigheder, når bidraget fra verbal arbejdshukommelse til læseforståelse vurderes. En anden, og noget simplere, tilgang kunne være at tage udgangspunkt i arbejdshukommelsesmål, der stiller ingen eller minimale krav til sprogforståelse. Det vil være tilgangen til at mindske sammenblanding af mål for verbal arbejdshukommelse og andre verbale mål i den nærværende undersøgelse.

Opsamling

Inferensfærdigheder er én af de færdigheder, der kan forklare forskelle i læseforståelse (kapitel 1 og 4). Ordforråd og arbejdshukommelse kan også forklare forskelle i læseforståelse, men undersøgelser, hvor den selvstændige betydning af inferensfærdighed (og ordforråd og arbejdshukommelse) for læseforståelse undersøges, er imidlertid sjældne.

Nyere forskning (Cain & Oakhill, 2014; Cromley & Azevedo, 2007; Li & Kirby, 2014) tyder på, at der er et *særligt* samspil mellem ordforråd og inferensfærdigheder i læseforståelse, men en mulig rolle for arbejdshukommelsesressourcer i dette samspil er ikke blevet undersøgt. Det synes derfor at være en mulighed, at den umiddelbare betydning af samspillet mellem ordforråd og inferensfærdigheder for læseforståelse faktisk afspejler forskelle i verbale arbejdshukommelsesressourcer, det vil sige evne til at processere verbal information.

Formål

Hensigten med den aktuelle empiriske undersøgelse var at undersøge samspillet mellem inferensfærdigheder og verbale færdigheder i form af ordforråd og verbal arbejdshukommelse i læseforståelse.

Et første formål var at afdække, i hvilken grad inferensfærdigheder, ordforråd og verbal arbejdshukommelse har betydning for læseforståelse. Det første forskningsspørgsmål lød:

1. I hvilken grad spiller inferensfærdigheder, ordforråd og verbal arbejdshukommelse hver især en rolle for læseforståelse?

Det næste formål var at undersøge, hvorvidt sammenhæng mellem inferensfærdigheder og læseforståelse kan tilskrives viden om ord i kombination med verbal arbejdshukommelse. Det andet forskningsspørgsmål lød:

2. Har inferensfærdigheder en selvstændig betydning for læseforståelse?

Det tredje formål var at undersøge den indirekte effekt af ordforråd via inferensfærdigheder nærmere. Der blev stillet to spørgsmål:

3a. I hvilken grad medierer inferensfærdigheder betydningen af ordforråd for læseforståelse?

3b. I hvilken grad kan en mediationseffekt forklares ved evne til at processere verbal information?

Metode

Design

Deltagernes færdigheder blev undersøgt med samtidige test som beskrevet i kapitel 3 (testning ved T2) og kapitel 4.

Deltagere

Deltagerne ($n = 54$) er beskrevet i kapitel 3.

Materialer

Læseforståelse

Målet for læseforståelse var scorene på den individuelle test af læseforståelse, som blev beskrevet i kapitel 3 (læseforståelse ved T2). Som beskrevet tidligere blev denne test afviklet med hjælp til afkodningen af eventuelle svære ord, og resultaterne på testen korrelerer da heller ikke med et sammensat mål for afkodning baseret på test af fonologisk kodning og kendskab til ords stavemåder.

Inferensfærdigheder

Målet for inferensfærdigheder var scorene på testen af inferenser fra fagtekster, som blev beskrevet i kapitel 4.

Ordforråd

Receptivt ordforråd blev undersøgt med en dansk oversættelse af Peabody Picture Vocabulary Test (PPVT) med en opgaverækkefølge tilpasset til dansk af Jeanette Bremer Nielsen (2008). I hver opgave blev fire billeder forelagt deltageren. Testleder læste et ord højt, og deltageren udpegede dét af de fire billeder, der matchede det hørte ord bedst. I opgave 79 hørte deltageren fx ordet ”forbløffelse” og kunne vælge mellem billeder af følgende: 1. En lille pige der græder ved siden af et glas, der tilsyneladende lige er gået itu. 2. En smilende pige, der klapper en hund. 3. En mand der med overrasket mine ser på et bundt pengesedler, som han holder i den ene hånd, og som han tilsyneladende netop har taget ud af en kuvert, som han holder i den anden hånd. 4. En kvinde iklædt forklæde og med en gulsvaber ved sin side siddende i en lænestol og med lukkede øjne. Før testopgaverne prøvede deltageren to øveopgaver. Der blev tildelt ét point for hver rigtig udpegning i testopgaverne.

Der er oprindeligt 150 testopgaver i materialet, men testen blev i denne undersøgelse påbegyndt ved den 30. opgave, da det forventedes, at alle deltagere kendte de 29 nemmeste ord. Disse første 29 opgaver tæller derfor *ikke* med i sumscorerne. Den maximalt mulige score var derfor 121. Testen blev afbrudt efter 6 fejl inden for 8 på hinanden følgende opgaver.

Arbejdshukommelse

Verbal arbejdshukommelse blev undersøgt med en test af baglæns talspændvidde fra CELF 4 (Semel, E., Wiig, E.H. & Secord, W., 2013). Testlederen læste en talrækkefølge højt (et tal pr. sekund), og deltageren skulle så huske og sige disse tal højt i *omvendt* rækkefølge. Hvis testleder læste: ”3 – 5”, var den rigtige respons altså: ”5 – 3”. Før testopgaverne prøvede deltageren en øveopgave. Antallet af tal i testopgaverne øgedes gradvist fra 2 til maksimalt 8. Der var to opgaver med rækkefølger, der inkluderede 2 tal, 3 tal, osv. Der blev tildelt ét point for hver rigtig respons. Testen blev afbrudt, når deltageren opnåede 0 point i de to på hinanden følgende opgaver med talrækkefølger af samme længde.

Fremgangsmåde

Fremgangsmåden var den fremgangsmåde, der blev beskrevet i kapitel 3 (for testning ved T2) og kapitel 4.

Resultater

Deskriptiv statistik for mål for læseforståelse og de tre prædiktorer, inferensfærdighed, ordforråd og arbejdshukommelse fremgår af tabel 1.

Tabel 1. Deskriptiv statistik for læseforståelse, inferensfærdigheder, ordforråd og arbejdshukommelse

	Max score	Middelværdi	Standardafvigelse	Spænd	α
Læseforståelse	46	36,8	4,5	25-46	0,73
Inferensfærdigheder	36	25,9	6,0	11-36	0,74
Ordforråd	121	74,8	13,3	57-108	0,72
Arbejdshukommelse	14	5,0	1,4	2-8	0,71

Note. For arbejdshukommelse er $n = 53$. For øvrige mål er $n = 54$. For arbejdshukommelse er Cronbach's alfa fra testmanualen (aldersgruppe: 12;00-12;11) gengivet. For øvrige test er Cronbach's alfa for den aktuelle elevgruppe beregnet. For ordforrådstesten er Cronbach's alfa beregnet for en majoritet af items (64%), der blev prøvet af en majoritet af deltagere (67%).

Alle scorefordelinger for målene i tabel 1 havde acceptable værdier for skævhed (inden for ± 1).

Det første spørgsmål var, i hvilken grad hver af de tre prædiktorer, inferensfærdigheder, ordforråd og arbejdshukommelse, spiller en rolle for læseforståelse. Tabel 2 viser korrelationer mellem alle mål.

Tabel 2. Korrelationer mellem læseforståelse, inferensfærdighed, ordforråd og arbejdshukommelse

	Læseforståelse	Inferensfærdighed	Ordforråd
Læseforståelse			
Inferensfærdigheder	0,58**		
Ordforråd	0,41**	0,50**	
Arbejdshukommelse	0,41**	0,32*	0,25

Note. ** $p < 0,01$, * $p < 0,05$. For korrelationer med arbejdshukommelse er $n = 53$. For øvrige korrelationer er $n = 54$.

Af tabel 2 fremgår det, at alle tre prædiktorer var moderat og signifikant korreleret med læseforståelse.

Det andet spørgsmål var, om inferensfærdigheder har en selvstændig betydning for læseforståelse. For at afdække om inferensfærdigheder bidrog til læseforståelse, når der kontrolleredes for bidraget fra ordforråd og arbejdshukommelse, blev der gennemført en hierarkisk regressionsanalyse. Resultaterne fremgår af tabel 3.

Tabel 3. Resultater fra hierarkisk regressionsanalyse med samlet læseforståelse som afhængig variabel og ordforråd samt arbejdshukommelse (trin 1) og inferensfærdigheder (trin 2) som uafhængige variable

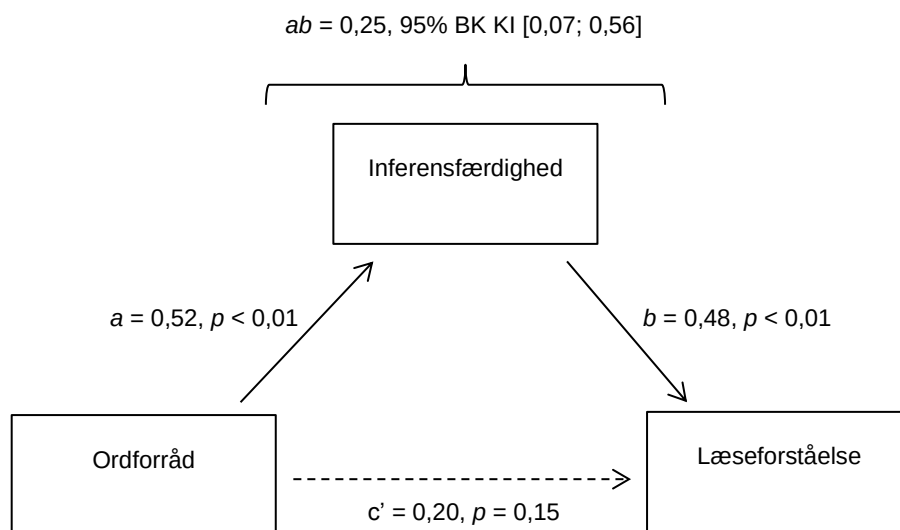
Trin	Mål	R^2	ΔR^2	F ændring	β
1	Ordforråd				0,17
	Arbejdshukommelse	0,29	0,29	10,34**	0,23*
2	Inferensfærdigheder	0,41	0,12	10,02**	0,42**

* $p = 0,049$, ** $p < 0,01$. $N = 53$.

Ordforråd og verbal arbejdshukommelse forklarede tilsammen 29% af variationen i læseforståelsen. Herudover forklarede inferensfærdigheder 12% af variationen i læseforståelse.

De næste spørgsmål angik graden af mediation, det vil sige den effekt, som ordforråd kan have på læseforståelse via inferensfærdigheder, og i hvilken grad denne mediation kan forklares ved verbal processingsevne. For at undersøge den indirekte effekt af ordforråd gennem inferensfærdigheder blev der gennemført to statistiske mediationsanalyser ved hjælp af tilføjelsesprogrammet *Process* skrevet til SPSS af Hayes (2013). Variablene blev forud for disse analyser standardiseret på scorerne fra de 53 deltagere, der havde prøvet alle tre test for at lette sammenligning af resultaterne. For at bestemme signifikansniveauet for de indirekte (medierede) effekter i disse analyser, blev der konstrueret konfidensintervaller for de indirekte effekter baseret på 5000 bootstrap samples. Den indirekte effekt af en variabel gennem en mediator anses for at være signifikant på 0,05 niveauet, hvis det bias-korrigerede 95% konfidensinterval ikke inkluderer nul (Hayes, 2013).

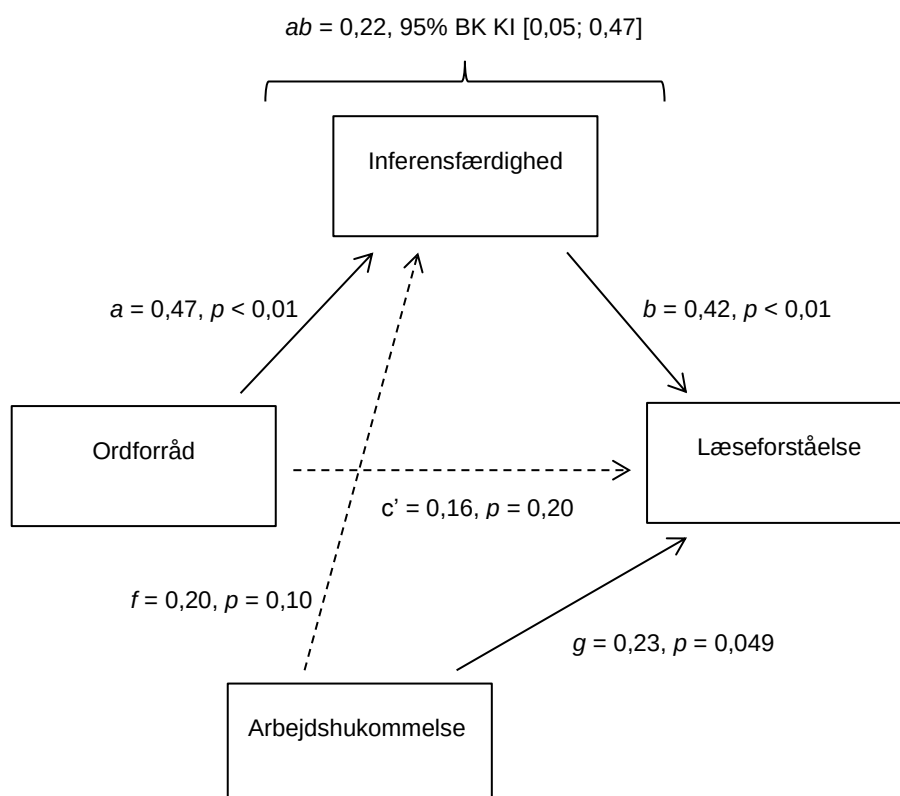
Graden af mediation blev undersøgt i en simpel mediationsanalyse med læseforståelse som afhængig variabel, inferensfærdigheder som mediator og ordforråd som uafhængig variabel. Resultaterne af analysen fremgår af figur 1.



Figur 1. Resultater fra simpel mediationsanalyse.

Den indirekte effekt af en variabel gennem en mediatorvariabel beregnes som produktet af a og b . Der var en signifikant indirekte effekt (ab) af ordforråd på læseforståelse gennem inferensfærdigheder. Størrelsen af den indirekte effekt var betydelig: En forskel i ordforråd på én standardafvigelse svarede til en forskel på 0,25 standardafvigelse gennem inferensfærdigheder i læseforståelse. Den direkte effekt (c') af en uafhængig variabel på den afhængige variabel er i mediationsanalysen kontrolleret for mediatorvariablens effekt (b) på den afhængige variabel. Den direkte effekt af ordforråd på læseforståelse kontrolleret for effekten af inferensfærdigheder var ikke signifikant.

I hvilken grad verbal arbejdshukommelse kan forklare den indirekte effekt af ordforråd gennem inferensfærdigheder til læseforståelse, blev undersøgt ved at kontrollere for effekterne af arbejdshukommelse på hhv. inferensfærdigheder og læseforståelse. Resultaterne af analysen med arbejdshukommelse som kovariat fremgår af figur 2.



Figur 2. Resultater fra mediationsanalyse med arbejdshukommelse som kovariat.

Den indirekte effekt af ordforråd gennem inferensfærdigheder var fortsat signifikant, når der kontrolleredes for effekt af arbejdshukommelse på hhv. inferensfærdigheder (f) og læseforståelse (g). En forskel i ordforråd på én standardafvigelse svarede til en forskel på 0,22 standardafvigelse i læseforståelse via inferensfærdigheder ved et konstant arbejdshukommelsesniveau. Analysen viste også, at effekten af arbejdshukommelse på inferensfærdigheder ikke var signifikant, når der kontrolleredes for ordforråd, og at effekten af arbejdshukommelse på læseforståelse var signifikant efter kontrol for inferensfærdigheder og ordforråd.

Diskussion

Det var hensigten at undersøge samspillet mellem inferensfærdigheder, ordforråd og arbejdshukommelse i læseforståelse. Det indledende forskningsspørgsmål lød: (1) *I hvilken grad spiller inferensfærdigheder, ordforråd og verbal arbejdshukommelse hver især en rolle for læseforståelse?* Resultaterne viste, at både inferensfærdigheder, ordforråd og verbal arbejdshukommelse har betydning for læseforståelse. Der var en moderat sammenhæng mellem

scorerne på test af læseforståelse og test af inferensfærdigheder, ordforråd og verbal arbejdshukommelse. De moderate sammenhænge mellem læseforståelse og ordforråd og mellem læseforståelse og inferensfærdigheder harmonerer med de moderate til stærke sammenhænge i undersøgelser med deltagere i tilgrænsende aldersgrupper, fx 11-årige deltagere hos Oakhill og Cain (2012). Mange undersøgelser af betydningen af verbal arbejdshukommelse for læseforståelse har som nævnt anvendt semantisk krævende opgaver, hvilket kan skabe usikkerhed om, hvorvidt disse arbejdshukommelsesmål først og fremmest bidrager til læseforståelse på grund af testenes krav til ordkendskab og/eller andre sprogforståelseskompener. Imidlertid indikerer resultaterne fra denne undersøgelse, at verbal arbejdshukommelse har en betydning for læseforståelse, der ikke blot kan tilskrives, at en opgave måtte sætte krav til sprogforståelse, eftersom informationen, der skulle lagres og processeres i de anvendte opgaver, var tallene fra 1 til 9 og dermed var tidligt indlært frekvente ord uden multiple betydninger. Resultater fra mediationsanalysen understøtter desuden denne konklusion. Arbejdshukommelse havde betydning for læseforståelse, selv når der statistisk kontrolleredes for betydningen af inferensfærdigheder og ordforråd for læseforståelse.

Det andet forskningsspørgsmål lød: (2) *Har inferensfærdigheder en selvstændig betydning for læseforståelse?* Resultaterne viser, at inferensfærdigheder synes at have en selvstændig betydning for læseforståelse. Betydningen af inferensfærdigheder kan således ikke forklares fuldt ud ved forskelle i viden om ord i kombination med en generel evne til at processere verbal information, som de her blev målt. Ordforråd og arbejdshukommelse havde en betydelig fælles varians med inferensfærdigheder i læseforståelse, men inferensfærdigheder bidrog signifikant til læseforståelse *ud over* ordforråd og arbejdshukommelse. Undersøgelsen bekræfter dermed resultater fra undersøgelser af bl.a. Cain et al. (2004) og Li og Kirby (2014), der tyder på, at inferensfærdighed er en selvstændig komponent i læseforståelse. Når verbale inferensfærdigheder forudsiger læseforståelse, hvor nonverbale inferensfærdigheder ikke gør (kapitel 4), kan det dermed formentlig ikke blot tilskrives, at verbale inferensfærdighedsmål trækker på grundlæggende verbale færdigheder som ordforråd.

En alternativ forklaring på inferensfærdighedsmålets evne til at forklare variation i læseforståelse, der skal adresseres, er muligheden for, at inferensfærdighedsmålet reelt er et udtryk for læseforståelse, og at det er derfor, at målet er så stærk en prædiktør for læseforståelse i sammenligning med ordforråd. Imidlertid adskiller inferensfærdighedsmålet og læseforståelsesmålet sig i en analyse af målenes indholdsvaliditet. Læseforståelsesmålet inkluderer spørgsmål, der er rettet mod flere aspekter af læseforståelse end komponenten inferensfærdighed. Inferensfærdighedsmålet indeholdt kun inferenskrævende opgaver, og i udviklingen af materialet blev krav til baggrundsviden forsøgt begrænset, ligesom potentielt svære ord blev undgået. Hvis de to test grundlæggende skulle måle den samme underliggende færdighed (læseforståelse), burde man desuden forvente en om ikke perfekt så i hvert fald meget stærk korrelation mellem de to mål (Cromley & Azevedo, 2007). Den moderate korrelation mellem inferensfærdigheder og læseforståelse i den aktuelle undersøgelse er som nævnt i overensstemmelse med resultater fra andre undersøgelser (fx Oakhill & Cain, 2012).

Det første af to spørgsmål angående den specifikke relation mellem ordforråd og inferensfærdigheder lød: (3a) *I hvilken grad medierer inferensfærdigheder betydningen af*

ordforråd for læseforståelse? Resultaterne tyder på, at inferensfærdigheder helt medierer betydningen af ordforråd for læseforståelse. Der var en signifikant indirekte effekt af ordforråd til læseforståelse gennem inferensfærdigheder. Til sammenligning var den direkte effekt af ordforråd ikke signifikant. Der var således tale om fuld mediation.

Set i lyset af ordforråds mere end veletablerede grundlæggende rolle i læseforståelse var det overraskende, at der ikke var en signifikant direkte effekt af ordforråd på læseforståelse. Cromley & Azevedo (2007) fandt da også en signifikant direkte effekt af et ordforrådsmål, der dog var væsensforskelligt fra det her anvendte, idet Cromley og Azevedos mål inkluderede læsning. Det overraskende resultat i den aktuelle undersøgelse kan ikke umiddelbart tilskrives problemer med den anvendte ordforrådstests reliabilitet, idet testen korrelerede med læseforståelse inden for det forventede spænd, og den interne homogenitet også var tilfredsstillende. Men det kan have været en faktor, at ordforrådsmålet primært var en test af ordforrådets bredde²⁰. Undersøgelser peger netop på, at især ordforrådets dybde og eventuelt den hastighed, hvormed ordbetydninger kan aktiveres, har en særlig grad af betydning for læseforståelse (Ouellette, 2006; Oakhill et al., 2013). Test, der er følsomme over for disse aspekter af ordforråd, bør derfor inddrages i opfølgende undersøgelser. Sådanne undersøgelser bør også inkludere flere deltagere for at reducere risikoen for at forkaste hypotesen om en direkte effekt, hvor denne faktisk findes.

Den indirekte effekt af ordforråd via inferensfærdigheder var stor, ikke mindst i sammenligning med den effekt som Cromley og Azevedo (2007) fandt. En reduceret følsomhed i Cromley og Azevedos inferensfærdighedsmål kan muligvis forklare noget af forskellen i størrelsen på de indirekte effekter i de to undersøgelser. Der indgik for det første kun 8 opgaver i det samlede inferensfærdighedsmål, som Cromley og Azevedo anvendte i deres analyser, og testen havde desuden et multiple choice format, hvormed det ikke var nødvendigt for deltagerne selv at konstruere inferensen i samme grad som i den her anvendte inferensfærdighedstest med åbne spørgsmål.

I vurderingen af både den ikke-signifikante direkte effekt af ordforråd og den signifikante og betydelige indirekte effekt af ordforråd på læseforståelse gennem inferensfærdigheder er det væsentligt at pointere, at inferensfærdigheder ud over at trække på viden om ords betydninger og relationer, også kan forventes at spille en rolle for betydningen af ordforrådets indflydelse på læseforståelse, hver gang læseren på baggrund af sin foreløbige forståelse af teksten udvælger den relevante og præcise betydning af et bestemt ord i teksten. Indflydelse fra ordforråd til læseforståelse gennem inferensfærdigheder kan altså være aktuel i langt flere tilfælde under læsning af en tekst, end der hvor læseren helt tydeligt må drage inferenser for at skabe sammenhæng på tværs af sætninger og afsnit.

Mediationsanalysen indikerer en bestemt kausal rækkefølge mellem ordforråd og inferensfærdigheder i læseforståelsesprocessen, eftersom ordforråd kunne forklare variation i

²⁰ Der indgår dog et vist aspekt af vurdering af ordbetydningers indbyrdes relationer, idet opgaverne løses ved at identificere den betydning (et billede) blandt fire mulige, der *bedst* matcher det hørt ord. De fire mulige betydninger er ofte til en vis grad betydningsmæssigt relateret, og løsning af en opgave består derfor ikke nødvendigvis blot i en alt-eller-intet sammenligning af hver af de fire skitserede billeders indhold med det ord, der høres, men kan også støttes af en sammenligning af de fire billeders indhold og kendskab til de ordbetydninger, der vil kunne matche de forskellige billeder.

læseforståelse via inferensfærdigheder. Man kan dog ikke reelt sige noget med sikkerhed om årsagsforhold på baggrund af tværsnitsdata. Dertil er der brug for andre undersøgelsesdesigns. En måde at undersøge hypotesen om en særlig betydning af ordforråd og forskellige aspekter af ordforråd for inferensfærdigheder kunne være i et eksperimentelt design, hvori man havde kontrol med omfanget af viden om de ord, der forventedes at være centrale for bestemte inferenser, som derefter blev testet. Ved at introducere helt nye ord og bede deltagere indlære viden om disse ord og deres relationer i forskelligt omfang forud for inferensopgaverne ville undersøgelsen i særdeleshed kunne give indtryk af, i hvilken grad ordforrådets dybde spiller en afgørende rolle for inferensfærdigheder i sammenligning med ordforrådets bredde.

Det var ikke muligt i den aktuelle undersøgelse at afprøve hypoteser om en differentieret betydning af ordforråd for forskellige typer af inferenser. For at opnå en mere detaljeret forståelse af relationen mellem ordforråd og inferensfærdigheder vil det i fremtidige undersøgelser være oplagt ikke kun at inddrage flere forskellige aspekter af ordforråd, men også differentiere mellem inferenstyper, der forventes at trække på viden om ord i forskellig grad, som hos Cain og Oakhill (2014).

Det næste spørgsmål rettede sig mod betydningen af arbejdshukommelse for relationen mellem ordforråd og inferensfærdigheder i læseforståelse og lød: (3b) *I hvilken grad kan en mediationseffekt forklares ved evne til at processere verbal information?* Resultaterne viste, at forskelle i lagring og processering af verbal information som her målt ikke kan forklare mediationseffekten. Selv efter kontrol for verbal arbejdshukommelse var den indirekte effekt af ordforråd på læseforståelse via inferensfærdigheder signifikant og af betydelig størrelse. Det er dog et vigtigt forbehold, at teksterne i både inferens- og læseforståelsestestene var tilgængelige for deltagerne under besvarelsen af spørgsmål. Denne procedure kan have virket som en udlignende faktor for betydningen af forskelle i arbejdshukommelse. Man kan forestille sig, at arbejdshukommelse ville være betydeligt vigtigere for inferensfærdigheder i læseforståelse, hvis teksten ikke var til rådighed under opgavebesvarelsen, og at arbejdshukommelse dermed ville kunne redegøre for en større del af den indirekte effekt af ordforråd på læseforståelse via inferensfærdigheder under sådanne betingelser. Et andet forbehold er, at afstanden mellem de informationer, der skulle forbindes i inferenstesten, var begrænset og under alle omstændigheder ikke blev manipuleret eksperimentelt. Krav til arbejdshukommelse forventes netop at øges med afstanden mellem disse informationer. Undersøgelser af forskellige ordforråds- og arbejdshukommelsesmåls betydning for inferensfærdigheder viser som nævnt, at ordforråd enten helt eller delvist kan mediere indflydelsen fra arbejdshukommelse til inferensfærdigheder (Currie & Cain, 2015; Chrysochoou et al., 2011). Men heller ikke i disse undersøgelser skelner man mellem eller sammenligner inferenser, der skal drages over længere og kortere afstande i tekst. Udvikling af inferenstest, hvor afstanden mellem ordbetydninger, der forventes at være afgørende for inferenserne, varieres systematisk, kan gøre det muligt at afdække, om det fx særligt er i betingelser med kort afstand mellem de afgørende ordbetydninger, at ordforråd medierer arbejdshukommelsens betydning for inferensfærdigheder.

Konklusion

Inferensfærdigheder, ordforråd og arbejdshukommelse var alle associeret med læseforståelse, og inferensfærdigheder var desuden associeret med ordforråd og arbejdshukommelse.

Inferensfærdigheders betydning for læseforståelse kunne ikke forklares fuldt ud af ordforråd og arbejdshukommelse. Undersøgelsen bekræfter dermed andre undersøgelser, der tyder på, at inferensfærdigheder er en selvstændig komponent i læseforståelse.

Inferensfærdigheder medierede desuden helt bidraget fra ordforråd til læseforståelse. Det tyder på, at viden om ords betydning ikke er tilstrækkelig for læseforståelse. Effekten kunne dertil ikke tilskrives forskelle i verbal processering. Resultaterne tyder dermed på, at ordforråd er en vigtig forudsætning for forståelse, men at det samtidig er afgørende at kunne gøre brug af denne viden ved at drage inferenser for at opnå forståelse af en tekst. Med andre ord: Ordforråd understøtter inferensfærdigheder, og inferensfærdigheder understøtter god læseforståelse.

Kapitel 6: Er detaljefokus en hæmsko for inferensfærdigheder?

Indledning

Ordforråd er en vigtig forudsætning for inferensfærdigheder (kapitel 5), men der må også kunne være andre årsager til, at nogle læsere drager flere inferenser end andre.

Som kandidat til en bagvedliggende årsag til individuelle forskelle i inferensfærdigheder finder man i læseforskningen idéen om, at læsere har forskellige forventninger til, hvor godt deres forståelse af tekster skal hænge sammen. Antagelsen er, at forventninger om sammenhæng sætter læseren i gang med at drage inferenser. I autismeforskningen finder man idéen om en tendens til at fokusere på detaljer, der er blevet kædet sammen med ringe inferensfærdigheder. I den aktuelle undersøgelse er det idéen at afprøve, om der er indikationer på, at lave forventninger om sammenhæng kan henføres til en tendens til at fokusere på detaljer.

Forventning om sammenhæng

Van den Broek og kolleger har introduceret idéen om, at læseren har forventninger til, hvor godt deres forståelse af tekster skal hænge sammen (reader's standard for coherence) (van den Broek & Kremer, 2000; van den Broek, Ridsen, & Husebye-Hartmann, 1995). Ifølge hypotesen opererer læseren med et sæt af forventninger eller kriterier for, hvornår forståelsen af en tekst er tilfredsstillende sammenhængende. Det er disse forventninger, der sammen med de rammer og begrænsninger (constraints), som er i teksten, afgør omfanget af inferensdragning under læsning. Man kan sige, at læseren gør, hvad der skal til for at opnå den grad af sammenhæng, som læserens egne forventninger dikterer, i det materiale som den aktuelle tekst udgør. Forventningerne er i princippet i spil for hver sætning eller information, læseren bearbejder. Når tilfredsstillende sammenhæng mellem en information og tidligere tekst og/eller baggrundsviden er etableret, kan læseren bevæge sig videre i teksten. Van den Broek og kolleger beskriver især læserens forventninger om kausal og referentiel sammenhæng, men hævder, at forventninger om sammenhæng kan være rettet mod flere forskellige dimensioner, og at bestemte teksttyper kan tænkes at fordre bestemte typer af forventninger. Hvordan eller om en læsers forventninger om sammenhæng lader sig afdække uafhængigt af læserens forståelse af tekst er ikke klart, men van den Broek og kolleger mener, at der er individuelle forskelle blandt læsere med hensyn til hvor høje forventninger om sammenhæng, de opererer med. Denne idé om individuelle forskelle i forventninger om sammenhæng er blevet taget op i forskningen, der har fokuseret på forskelle mellem gode og dårlige forståere. Oakhill og Cain (2007) har således foreslået, at dårlige forståere er kendetegnet ved lave forventninger om sammenhæng, og at disse lave forventninger kan forklare,

hvorfor dårlige forståere drager færre inferenser end gode forståere: De stræber ikke efter sammenhæng i det læste i samme grad som gode forståere.

En lang række af faktorer kan, ifølge van den Broek et al. (1995), influere på en læsers aktuelle forventninger for sammenhæng i en bestemt tekst og på, om læseren er i stand til at indfri sine forventninger. På tværs af tekster kan selv gode læsers forventninger variere som funktion af baggrundsviden om tekstens emne. Læseren med stor baggrundsviden om tekstens emne er i stand til at have og indfri høje forventninger om sammenhæng i sin forståelse af teksten. Grad af motivation og et bestemt læseformål kan desuden påvirke en læsers aktuelle forventninger om sammenhæng, og ydre faktorer, som fx den tid, der gives til at læse en tekst, kan influere på, om læseren kan indfri sine forventninger.

Men på trods af disse mange faktorer, der kan føre til intra-individuelle forskelle i læsers forventninger om sammenhæng, synes der altså at være en antagelse om, at et grundlæggende subjektivt udgangspunkt eller en grundindstilling i den enkeltes forventninger kan spille en rolle for læserens inferensdragnet på tværs af tekster og situationer. Denne tanke afspejles også i den litteratur, der foreslår, at vanskelighederne hos dårlige forståere, der opfattes som relativt konstante på tværs af situationer, kan skyldes lave forventninger om sammenhæng. Spørgsmålet er som nævnt her, om lave forventninger kan være sammenfaldende med en tendens til at fokusere på detaljer.

Tendens til at fokusere på detaljer

Idéen om, at nogle personer kan karakteriseres ved en tendens til at fokusere på detaljer, tager sit udspring i en anden idé, nemlig idéen om at nogle personer ikke er i stand til at etablere overordnet sammenhæng i forskellige stimuli i samme grad som andre. Uta Frith, som introducerede denne idé i 1980'erne, var interesseret i årsager til autisme og foreslog, at en forstyrrelse i et centralt kognitivt system, der normalt opretter overordnet sammenhæng, kunne være en årsag til autisme. Hun kaldte denne forstyrrelse 'weak central coherence'. I sin oprindelige hypotese beskrev Frith, hvordan forstyrrelsen medfører, at personer med autisme ikke stræber efter sammenhæng, og denne manglende stræben efter sammenhæng blev igen kædet sammen med studier, der viste, at personer med autisme generelt ligger lavt på test af inferensfærdigheder (Frith, 2003). Antagelsen om en svag evne til at etablere sammenhæng og en forbindelse til ringe inferensfærdigheder er siden blevet meget udbredt og har været central i en række undersøgelser af personer med autisme (fx Jolliffe & Baron-Cohen, 1999), og den inddrages ofte i diskussionen af resultater, der belyser læseforståelse og/eller inferensfærdigheder hos personer med autisme (fx Norbury & Nation, 2011; Wahlberg & Magliano, 2004).

Happé og Frith (2006)²¹ modificerede dog den oprindelige hypotese som konsekvens af en række nyere forskningsresultater. I stedet for en forstyrrelse i evnen til at etablere sammenhæng, der endvidere er årsag til autisme, beskrev Happé og Frith (2006) en detaljeorienteret kognitiv stil – en tendens til at fokusere på detaljer – som ofte vil kunne iagttages hos personer med autisme. Denne modifikation imødekommer bl.a. resultater fra undersøgelser, der finder, at personer med autisme

²¹ I den forbindelse benytter Happé og Frith betegnelsen 'weak coherence' i stedet for 'weak central coherence'.

kan finde og opnå sammenhæng i stimuli i forskellige situationer, fx når det eksplicit fremgår af opgaveformuleringen, at formålet er at opnå sammenhæng eller helhedsforståelse. Happé og Frith formulerede endvidere en arbejdshypotese, hvor en spontan tendens til at fokusere på detaljer kontrasterer med en spontan tendens til at fokusere på helhed og sammenhæng (strong coherence) (Booth & Happé, 2010; Happé, 1999; Happé & Frith, 2006). De beskrev, hvordan en spontan tendens til at fokusere på detaljer og en spontan tendens til at fokusere på sammenhæng og helhed kan være polerne på et kontinuum, hvorpå populationen af personer med autisme og populationen af personer uden autisme er distribueret i to normalfordelinger. Fordelingen for personer med autisme forventes at ligge tættest på polen *spontan fokus på detaljer*, men de to normalfordelinger overlapper i nogen grad. Ifølge Happé og Frith (2006) er det ikke sådan, at personer med en detaljeorienteret kognitiv stil ikke *kan* etablere helhed og sammenhæng i stimuli (fx ved at drage inferenser), men de har en *spontan* tendens til at fokusere på detaljer frem for helhed og sammenhæng²². Ligeså synes det at være muligt, at en person med en spontan tendens til at fokusere på helhed og sammenhæng strategisk kan fokusere på detaljer og dertil muligvis være god til det. Arbejdshypotesen om en normalfordeling mellem detalje- og helhedsfokus imødekommer forskningsresultater, der har udfordret den oprindelige hypotese. En distribution langs et kontinuum i autismepopulationen og et samtidigt overlap med en distribution for populationen uden autisme kan potentielt forklare, hvorfor man i en del undersøgelser ikke finder, at grupper af personer med autisme har stærkere tendens til at fokusere på detaljer end de kontrolgrupper, de sammenlignes med. Undersøgelser, der inddrager personer uden autismediagnoser, giver netop belæg for at forvente, at en tendens til at fokusere på detaljer kan være distribueret blandt personer uden autisme (Booth & Happé, 2010; Happé, Briskman, & Frith, 2001).

En række test er, siden Frith første gang beskrev sin hypotese, blevet anvendt til at undersøge først en svag evne til at etablere sammenhæng og siden en detaljeorienteret kognitiv stil. Når sproglige test anvendes, er der dog reelt tale om test, der enten direkte eller indirekte kan siges at teste inferensfærdigheder. For at undersøge om personer med autisme stræber efter sammenhæng i overensstemmelse med hypotesen om en svag evne til at etablere sammenhæng, anvendte Jolliffe og Baron-Cohen (1999) fx tre forskellige test, der alle kan siges at kræve inferensdragning. Det var ud over en inferenstest, hvor deltageren skal vælge en sætning, der kan få to andre sætninger til at hænge meningsfuldt sammen, en test med højtlesning af homografer, hvor udtalen af en ordform med flere mulige udtaler kan infereres på baggrund af ordformens meningskontekst, og det var en test med flertydige sætninger, hvor den aktuelle betydning af en sætning kan udvælges (infereres) på baggrund af konteksten. Booth & Happé (2010) anvendte en produktiv test, hvor deltageren skal færdiggøre sætninger, og hvor inferensfærdigheder ligeledes kan være central for deltagerens færdighed i at producere meningsfulde afslutninger på sætningerne, der er i overensstemmelse med konteksten (sætningen). Imidlertid er der også blevet anvendt nonverbale test til at undersøge, om en person kan karakteriseres ved en svag evne til at etablere sammenhæng/en tendens til at fokusere

²² Happé og Booth (2008) genintroducerer dog ideen om en læderet evne til at etablere overordnet sammenhæng hos nogle personer med autisme. I studier hvor man har forsøgt at undersøge detaljefokus og evnen til at etablere helhed og sammenhæng hos personer med autisme med særskilte opgaver, ser Happé og Booth således indikationer på, at markant detaljefokus og reduceret evne til at etablere helhed og sammenhæng begge (og uafhængigt af hinanden) kan forekomme hos personer med autisme.

på detaljer. Hverken Friths oprindelige idé om en forstyrrelse i et centralt kognitivt system eller idéen om en detaljeorienteret kognitiv stil drager da heller skel mellem modaliteterne. En tendens til at fokusere på detaljer synes at antages at være en generel kognitiv stil. Én af de først anvendte og meget populære nonverbale test i autismeforskningen er *The Embedded Figures Test* (Happé & Frith, 2006; White & Saldaña, 2011).

The Embedded Figures Test

The Embedded Figures Test (EFT) findes i forskellige udgaver, men fælles for dem alle er, at deltageren i en række opgaver skal finde en lille figur (der beskrives som en detalje) inden i en større figur (der beskrives som en helhed). Hvis deltageren hurtigt finder de små figurer inden i de større figurer, tages det til udtryk for en tendens til at fokusere på detaljer frem for sammenhæng og helhed (Happé & Frith, 2006). Det betyder, at man bør forvente en negativ sammenhæng mellem scorer på testen og inferensfærdigheder, hvis detaljefokus er relateret til relativt dårlige inferensfærdigheder.

EFT blev udviklet af Herman Witkin og kolleger og blev oprindeligt anvendt som et redskab til vurdering af en persons kognitive stil. Den kognitive stil, som testen skulle afdække, var en såkaldt 'field independent' stil. En person med denne kognitive stil blev beskrevet som analytisk i sin perception og tænkning og som mindre påvirkelig af sine omgivelser ('the field'), end mennesker som var karakteriseret ved en 'field dependent' stil (Witkin, Moore, Goodenough, & Cox, 1977).

Happé og Frith (2006) beskriver forskellen på tolkning af scorer på EFT indenfor hhv. FI/FD-forskningen og deres egen brug, hvor testen tages til udtryk for et fokus på detaljer frem for helhed og sammenhæng: Ifølge Witkin's teori er selv FI-personer spontant indstillet eller fokuseret på helhed (den overordnede struktur), men FI-personer forventes at klare sig godt på EFT på grund af deres evne til at *modstå eller abstrahere fra* helheden. Happé og Frith forventer derimod, at personer klarer testen godt, fordi de ikke spontant retter opmærksomheden mod helheden/den store figur, men først ser figurens enkeltdele i form af detaljer.

Hvad angår den eventuelle brug af EFT som en (negativ) indikator på forventning om sammenhæng der kan forklare, at nogle læsere drager færre inferenser end andre, er det interessant, at nogle forskere i FI/FD-traditionen har fundet en positiv sammenhæng mellem scorer på EFT og mål for læseforståelse (Bruininks, 1969; Drane, Halpin, Halpin, von Eschenbach & Worden, 1989), og andre FI/FD-forskere har formuleret forventninger om, at scorer på EFT vil være specifikt relaterede til inferensfærdigheder – vel at mærke *gode* inferensfærdigheder (Davey, 1990; Pitts & Thompson, 1984). Pitts og Thompson begrundede deres forventning om en sammenhæng med inferensfærdigheder med, at EFT-scorer forventes at reflektere, dels hvor godt deltageren kan finde væsentlige elementer i stimuli (features) *på trods* af et globalt mønster, og dels i hvilken grad deltageren er i stand til at reagere på flertydig stimuli på en analytisk måde. Hverken Pitts og Thompsons resultater eller Daveys resultater giver dog empirisk belæg for en specifik sammenhæng mellem EFT-scorer og inferensfærdigheder. Davey (1990) fandt, at EFT-scorer var associeret med inferentiel forståelse, når forståelsesmålet indebar åbne spørgsmål, og det samtidig ikke var muligt for deltagere at genlæse teksten før spørgsmålene besvares, men Davey undersøgte ikke, om denne

sammenhæng adskilte sig fra en relation mellem EFT-scorer og et ikke-inferentielt forståelsesmål. Pitts og Thompsons (1984) resultater tyder endvidere på, at EFT-scorer i bedste fald er en *svag* prædiktør af inferensfærdigheder, idet korrelationen mellem EFT og to mål for inferensfærdigheder i deres undersøgelse er hhv. 0,15 og 0,31.

I andre undersøgelser har EFT-scorer vist sig at være associeret med mere generelle færdigheder. Der kan fx påvises en vis sammenhæng med årskarakterer i mange forskellige skolefag (Tinajero & Páramo, 1997), og EFT-scorer korrelerer med scorer på forskellige intelligenstag, især nonverbale test (McKenna, 1984). McKenna fandt således, at korrelationerne mellem EFT og nonverbal problemløsning (Ravens matricer) over en række af studier er moderate til stærke. En eventuel sammenhæng mellem EFT-scorer og inferensfærdigheder synes altså potentielt at kunne være udtryk for en fælles sammenhæng med problemløsningsfærdigheder og/eller andre generelle færdigheder.

Formål

Formålet med den aktuelle undersøgelse var at undersøge, hvilken relation der er mellem EFT-scorer og inferensfærdigheder, og herved indirekte afprøve, om en tendens til at fokusere på detaljer og lave forventninger til sammenhæng kan være sammenfaldende, og om EFT dermed vil kunne anvendes som negativ indikator på læserens forventninger til sammenhæng.

Forskningsspørgsmålene lød:

- 1. Er EFT-scorer negativt eller positivt relateret til inferensfærdigheder?*
- 2. Er en relation mellem scorer på EFT og inferensfærdigheder nærmere udtryk for en generel sammenhæng mellem EFT-scorer og læseforståelse end for en specifik sammenhæng mellem EFT-scorer og inferensfærdigheder?*
- 3. Kan en relation mellem scorer på EFT og inferensfærdigheder forklares ved en bagvedliggende evne til problemløsning?*
- 4. I hvilken grad kan EFT-scorer forudsige individuelle forskelle i inferensfærdigheder i sammenligning med en allerede etableret prædiktør af inferensfærdigheder som ordforråd?*

Metode

Design

Deltagernes færdigheder blev undersøgt med samtidige test som beskrevet i kapitel 3 (testning ved T2) og kapitel 4.

Deltagere

Deltagerne ($n = 54$) er beskrevet i kapitel 3.

Materialer

Inferensfærdigheder

Målet for inferensfærdigheder var testen af inferensfærdighed fra fagtekst, som blev beskrevet i kapitel 4.

Tendens til at fokusere på detaljer

Tendens til at fokusere på detaljer blev afdækket med en udgave af EFT, der kan afvikles som gruppetest The Group Embedded Figures Test (GEFT). Testen blev dog her afviklet som individuel test i lighed med de andre test i testbatteriet. Manualen til GEFT (Witkin, Oltman, Raskin & Karp, 1971/2002) beskriver primært afprøvninger med collegestuderende, men det fremgår også, at en version af testen er blevet anvendt med børn på 10 år. Før den aktuelle undersøgelse blev testen piloteret blandt 5. og 6. klasseselever, og det blev sikret, at formatet fungerede med denne aldersgruppe. GEFT består af et hæfte, der er opdelt i tre sektioner – en øvesektion (1. sektion) og to testsektioner med hver 9 opgaver (2. og 3. sektion). I hver opgave ses en kompleks geometrisk figur, og deltageren skal optegne (tydeliggøre) de streger i figuren, der tilsammen udgør en bestemt mindre figur. Hvilken lille figur, der skal findes i den enkelte opgave, fremgår af hæftets bagside. Deltageren kan til enhver tid vende hæftet for at gense, hvilken lille figur der skal findes i den større figur, men deltageren har altså aldrig samtidigt adgang til de to figurer. Der er en tidsgrænse på fem minutter til løsningen af hver af de to testsektioner. Manualen foreskriver, at antal rigtige i de to sektioner lægges sammen til én score, hvormed den maksimale score er 18. Manualen opgiver en split half korrelation (korrelationen mellem sektion 2 og 3), som er korrigeret for halvering af items, på 0,82. Testmaterialet findes i den selvstændige bilagssamling (bilag F).

Bogstavelig forståelse

Målet for bogstavelig forståelse bestod af besvarelser på ti spørgsmål til bogstavelig information i teksterne i testen af inferensfærdigheder ind imellem de inferenskrævende spørgsmål. Der blev givet to point for fyldestgørende svar og ét point for et delvist fyldestgørende svar. Hvor teksten fx fortæller: *I Asien har man i mange år tæmmet elefanter og trænet dem op til at arbejde*, blev svar på spørgsmålet: "Hvad har man gjort i Asien?" som "man har tæmmet og trænet elefanterne op til at arbejde" tildelt 2 point, og svar som "tæmmet elefanter" eller "man har trænet elefanter" blev tildelt 1 point. Den maksimalt opnåelige score var 20. Den interne homogenitet var lav, Cronbach's alfa = 0,43.

Problemløsning

Evne til nonverbal problemløsning blev undersøgt med deltesten Matricer fra testmaterialet *Wechsler Nonverbal Scale of Ability* (Wechsler & Naglieri, 2009). I hver opgave ser deltageren en ufuldstændig matrix og vælger den manglende del fra fire eller fem svarmuligheder. Testen afvikles med nonverbale instruktioner, og der er tre øveopgaver. I overensstemmelse med manualen startede de 12-årige deltagere med opgave 7. Hver korrekt opgave udløste ét point. Testen blev afbrudt efter 4 opgaver til nul point inden for fem på hinanden følgende opgaver. Den maksimalt opnåelige score var 41. Manualen opgiver koefficienter for testens interne homogenitet på 0,88 og 0,93 for 12-årige. Koefficienterne er baseret på de stikprøver, der ligger til grund for den amerikanske og canadiske normering af materialet.

Ordforråd

Målet for ordforråd var det receptive ordforrådsmål, som blev beskrevet i kapitel 5.

Fremgangsmåde

Fremgangsmåden var den fremgangsmåde, der blev beskrevet i kapitel 3 (for testning ved T2) og kapitel 4.

Resultater

Tabel 1. Deskriptiv statistik

	Max score	Middelværdi	Standardafvigelse	Spænd
Inferensfærdigheder	36	25,9	6,0	11-36
EFT	18	9,1	4,0	3-17
Bogstavelig forståelse	20	17,5	2,2	11-20
Problemløsning	41	22,7	4,9	13-38
Ordforråd	121	74,8	13,3	57-108

Alle scorefordelinger for målene i tabel 1 havde acceptable værdier for skævhed (inden for $\pm 1,1$). Der var tendens til loftseffekt på målet for bogstavelig forståelse, hvor middelværdien lå ca. en standardafvigelse under den maksimale score.

Nogle undersøgelser har fundet kønsforskelle på EFT, oftest sådan at drenge scorer lidt højere end piger (Witkin et al., 1971/2002). I den aktuelle undersøgelse scorede piger (middelværdi = 9,9, standardafvigelse = 4,0) gennemsnitligt lidt højere end drenge (middelværdi = 8,5, standardafvigelse = 4,0) på EFT. Denne forskel var ikke signifikant $t(52) = -1,3$. Resultater for piger og drenge blev i de følgende analyser behandlet under ét.

Det første spørgsmål var, om EFT-scorer er negativt eller positivt relateret til inferensfærdigheder. Tabel 2 viser korrelationer mellem alle mål.

Tabel 2. Korrelationer mellem inferensfærdigheder og øvrige mål

	1	2	3	4
1. Inferensfærdigheder				
2. EFT	0,38**			
3. Bogstavelig forståelse	0,28*	0,21		
4. Problemløsning	0,30*	0,42**	0,29*	
5. Ordforråd	0,50**	0,38**	0,37**	0,17

Note. ** $p < 0,01$, * $p < 0,05$.

Af tabel 2 ses det, at der var en signifikant positiv korrelation mellem EFT-scorer og inferensfærdigheder ($r = 0,38$). Højere scorer på EFT var således relateret til højere scorer på testen af inferensfærdigheder.

Det andet spørgsmål var, om en relation mellem scorer på EFT og inferensfærdigheder nærmere er udtryk for en generel sammenhæng mellem EFT-scorer og læseforståelse end en specifik sammenhæng mellem EFT-scorer og inferensfærdigheder. Af tabel 2 ses, at EFT-scorer korrelerede svagt og ikke signifikant med bogstavelig forståelse. Til sammenligning var korrelationen mellem EFT-scorer og inferensfærdigheder signifikant og svag til moderat. Der var dog ikke signifikant forskel på korrelationen mellem EFT og inferensfærdigheder og korrelationen mellem EFT og bogstavelig forståelse ($df = 51$, $t = 1,10$, $p > 0,05$).

Det tredje spørgsmål var, om relationen mellem præstation på EFT og inferensfærdigheder kan forklares ved en bagvedliggende evne til problemløsning. Af tabel 2 fremgår det, at både inferensfærdigheder og EFT-scorer var signifikant korreleret med nonverbal problemløsning. For at afdække om EFT-scorer bidrog til at forklare forskelle i inferensfærdigheder, når der blev kontrolleret for forskelle i nonverbal problemløsning, blev der gennemført en hierarkisk regressionsanalyse. Resultaterne fremgår af tabel 3.

Tabel 3. Resultater fra hierarkisk regressionsanalyse med inferensfærdighed som afhængig variabel og nonverbal problemløsning (trin 1) og EFT-scorer (trin 2) som uafhængige variable

Trin	Mål	R^2	ΔR^2	F ændring	β
1	Problemløsning	0,09	0,09	5,20*	0,17
2	EFT	0,17	0,08	4,77*	0,31*

Note. * $p < 0,05$.

Når der blev kontrolleret for evne til nonverbal problemløsning, forklarede EFT-scorer selvstændigt 8% af variationen i inferensfærdigheder. Tilsammen forklarede nonverbal problemløsning og EFT-scorer 17% af variationen i inferensfærdigheder.

Det fjerde spørgsmål angik prædiktionsværdien af EFT for inferensfærdigheder. Spørgsmålet var, i hvilken grad EFT-scorer kan forudsige individuelle forskelle i inferensfærdigheder i sammenligning med en allerede etableret prædiktør som ordforråd. Af tabel 2 fremgår det, at ordforråd var moderat korreleret med inferensfærdigheder ($r = 0,50$) og forklarede 25% af variationen i inferensfærdigheder ($r^2 = 0,50^2 = 0,25$). Til sammenligning forklarede EFT-scorer 14% af variationen i inferensfærdigheder ($r^2 = 0,38^2 = 0,14$). Korrelationen mellem EFT og inferensfærdigheder var dog ikke signifikant forskellig fra korrelationen mellem ordforråd og inferensfærdigheder ($t = 0,91$, $df = 51$, $p > 0,05$). Af tabel 2 fremgår det, at scorer på EFT og ordforråd var signifikant korreleret indbyrdes ($r = 0,38$). For at afdække i hvilken grad EFT-scorer bidrog til at forklare forskelle i inferensfærdigheder, når der blev kontrolleret for forskelle i ordforråd, blev der derfor gennemført en hierarkisk regressionsanalyse. Resultaterne fremgår af tabel 4.

Tabel 4. Resultater fra hierarkisk regressionsanalyse med inferensfærdigheder som afhængig variabel og ordforråd (trin 1) og EFT-scorer (trin 2) som uafhængige variable

Trin	Mål	R^2	ΔR^2	F ændring	β
1	Ordforråd	0,25	0,25	17,07**	0,41**
2	EFT	0,29	0,04	4,77	0,23

Note. * $p < 0,05$, ** $p < 0,01$.

Efter kontrol for ordforråd bidrog EFT-scorer til 4% af variationen i inferensfærdighed. Dette bidrag var ikke signifikant ($p = 0,08$).

Diskussion

Det var hensigten at undersøge relationen mellem EFT-scorer og inferensfærdigheder for derved at afdække, om der er indikation på, at en tendens til at fokusere på hhv. detaljer eller helhed kan udgøre en grundindstilling i læserens forventninger om sammenhæng, sådan at lave forventninger til sammenhæng i nogen grad kan henføres til en tendens til detaljefokus.

Det første forskningsspørgsmål lød: 1. *Er EFT-scorer negativt eller positivt relateret til inferensfærdigheder?* Resultaterne viste klart, at EFT-scorer var positivt korreleret med inferensfærdigheder. Det tyder ikke på, at en tendens til at fokusere på detaljer eller helheder kan udgøre en grundindstilling i læserens forventninger til sammenhæng, sådan at detaljefokus er sammenfaldende med lave forventninger om sammenhæng. Resultatet er umiddelbart i modstrid med den almindelige opfattelse eller hypotese, der findes i forskningslitteraturen om en sammenhæng mellem en tendens til at fokusere på detaljer/svag evne til at etablere sammenhæng (målt med fx EFT) og ringe inferensfærdigheder. Resultatet er imidlertid i overensstemmelse med den forventning, som enkelte FI/FD-forskere har formuleret om præstationer på EFT. Pitts og Thompson forventede således, at EFT-scorer og inferensfærdigheder ville være positivt korreleret. De anså EFT-scorer for at være et udtryk for evne til at finde kritiske træk *på trods* af en overordnet struktur og evne til at forholde sig analytisk til tvetydige stimuli og forventede, at disse træk ville være en force for inferensfærdigheder. Tolkningen af en specifik sammenhæng mellem EFT og inferensfærdigheder i et FI/FD-perspektiv kompromitteres dog som nævnt af undersøgelser, der viser at EFT-scorer kan relateres til en lang række generelle færdigheder.

Nogle FI/FD-forskere har fx vist sammenhænge mellem EFT-scorer og mål for generel læseforståelse. Det tredje forskningsspørgsmål lød netop: 2. *Er en relation mellem scorer på EFT og inferensfærdigheder nærmere udtryk for en generel sammenhæng mellem EFT-scorer og læseforståelse end for en specifik sammenhæng mellem EFT-scorer og inferensfærdigheder?* Resultaterne kan i nogen grad tyde på, at en sammenhæng med EFT ikke er specifik for inferensfærdigheder, men også gælder bogstavelig forståelse. Der kunne ikke påvises en signifikant stærkere sammenhæng mellem EFT-scorer og inferensfærdigheder end mellem EFT-scorer og et mål for bogstavelig forståelse, skønt den sidstnævnte korrelation ikke var signifikant. Den interne homogenitet for målet for bogstavelig forståelse var desuden lav. Med et mere internt homogent mål kunne korrelationen med EFT muligvis have vist sig at være stærkere og signifikant.

EFT-scorer har vist sig at være korreleret med problemløsningsfærdigheder i en sådan grad, at nogle forfattere opfatter målet som et udtryk for problemløsningsfærdighed. Det fjerde forskningsspørgsmål lød derfor: 3. *Kan en relation mellem scorer på EFT og inferensfærdigheder forklares ved en bagvedliggende evne til problemløsning?* Resultaterne viser, at selv om problemløsning er relateret til både EFT-scorer og inferensfærdigheder, kan EFT-scorers betydning for inferensfærdigheder ikke alene tilskrives problemløsningsfærdigheder, som de her er undersøgt med en konventionel (nonverbal) test. Det sidste forskningsspørgsmål viste sig overraskende at kunne give indikationer på, hvori den resterende sammenhæng mellem EFT og inferensfærdigheder kan bestå.

Det sidste forskningsspørgsmål omhandlede prædiktionsværdien af EFT for inferensfærdigheder:

4. *I hvilken grad kan EFT-scorer forudsige individuelle forskelle i inferensfærdigheder i sammenligning med en allerede etableret prædikator af inferensfærdigheder som ordforråd?* De sammenhænge mellem EFT-scorer og inferensfærdigheder, som Pitts og Thompson fandt i deres undersøgelse, indikerede, at EFT i bedste fald ville være en svag prædikator af inferensfærdigheder. I den aktuelle undersøgelse var EFT-scorers prædiktionsværdi for inferensfærdigheder dog størrelsesmæssigt ikke uvæsentlig i sammenligning med en allerede etableret prædikator i form af ordforråd. Ordforråd og EFT delte dog varians i inferensfærdigheder i en sådan udstrækning, at EFT-scorer ikke bidrog signifikant til at forudsige forskelle i inferensfærdigheder, når der kontrolleredes for forskelle i ordforråd. Det er et overraskende resultat, da der ikke er krav til viden om ords betydning eller sproglige færdigheder i det hele taget i EFT. Det er dog muligt, at nogle deltagere støttes i deres løsning af EFT-opgaverne ved at ræsonnere sig frem verbalt, fx ved at beskrive figurene for sig selv. En sådan strategi synes dog ikke at sætte store krav til viden om ords betydning. Der er et visuelt element i ordforrådsopgaven (hvor deltageren skal vælge mellem fire billeder), der i princippet kunne forklare en vis grad af sammenhæng med EFT, men der er netop ikke nogle visuelle krav i testen af inferensfærdigheder, så det visuelle element kan ikke redegøre for den fælles variation i inferensfærdigheder. En forklaring skal måske nærmere findes i form af en bagvedliggende variabel, der har en fælles betydning for evnen til at løse ordforråds- og EFT- og inferensopgaver, enten fordi selve færdigheden er afhængig af denne bagvedliggende variabel, eller fordi opgaveformatet inducerer et krav til generelle færdigheder. En sådan generel færdighed kunne fx være kontrol af opmærksomhed. Kontrol af opmærksomhed er afgørende for løsning af enhver opgave af en vis kompleksitet. Ordforrådstestens format kan desuden siges at fordrer en vis færdighed i at vurdere flere muligheder, inden man træffer et valg. Det samme kan muligvis karakterisere EFT-opgaverne og inferensfærdigheder. Pitts og Thompson tillagde netop en betydning af at kunne reagere på flertydig stimuli på en analytisk måde til såvel inferensfærdigheder som løsning af EFT-opgaver. Generelle færdigheder som disse kan inkluderes under samlebetegnelsen *selvregulerende* eller *eksekutive funktioner*, der bl.a. dækker evne til at planlægge, vurdere muligheder og træffe beslutninger om handling samt evne til selv-monitorering (Tranel, Anderson & Benton, 1994). Eksekutive funktioner kan i nogen grad forklare forskelle i præstationer på sprogforståelses- og læseforståelsestest. Dog har man ikke undersøgt, om specifikke komponenter i sprogforståelse kan mediere betydningen af eksekutive funktioner for forståelse (fx Kieffer, Vukovic & Berry, 2013).

Man kan også forestille sig, at særligt inferensfærdigheder og løsning af EFT-opgaver (og almindelige problemløsningsopgaver) trækker på en fælles bagvedliggende variabel, fx en evne til at forholde sig analytisk til tvetydige eller komplekse stimuli, og at sammenhængen med ordforråd primært skyldes, at forskelle i inferensfærdigheder over tid afspejles i ordforråds-scorer, da inferensfærdigheder, som det fremgik af kapitel 5, sætter læseren i stand til at lære nye ord fra konteksten.

Begrænsninger

En række begrænsninger for tolkningen af resultaterne udspringer af anvendelsen af EFT som et udtryk for detaljefokus og den version af testen, der er anvendt, samt den konkrete afviklingsprocedure.

Som Happé & Booth (2008) anfører, er de fleste test, der er blevet brugt til at undersøge evne til at etablere sammenhæng og senere detaljefokus, generelt ikke i stand til at give information om, om der er tale om det ene eller det andet, det vil sige, hvorvidt en deltager scorer højt på testen på grund af overlegent detaljefokus eller på grund af en læderet evne til at fokusere på helhed og etablere sammenhæng. White og Saldaña (2011) pointerer, at dette forbehold også gælder for EFT. Da der er en positiv sammenhæng mellem EFT-scorer og inferensfærdigheder, er det dog svært at argumentere for, at høje EFT-scorer skulle være udtryk for en læderet evne til at fokusere på helhed og etablere sammenhæng. Det skal her understreges, at den aktuelle undersøgelse har beskæftiget sig udelukkende med sammenhængen mellem EFT og inferensfærdigheder hos personer uden autisme. Dette er blevet gjort med udgangspunkt i Happé og Friths (2006) hypotese om en normaldistribution på et kontinuum mellem tendens til detaljefokus og tendens til helhedsfokus i populationen uden autisme og i et forsøg på at afdække, om en tendens til at fokusere på detaljer kan være sammenfaldende med lave forventninger om sammenhæng hos børn uden autisme. Det vides således ikke, om der også vil kunne findes en positiv sammenhæng mellem EFT-scorer og inferensfærdigheder i populationen med autisme. Andre mønstre og sammenhænge kan vise sig for denne gruppe. Happé og Booth (2008) henviser fx til ikke publicerede resultater, hvor man på tværs af flere test finder, at to udledte faktorscorer for evne til 'integration' og evne til 'segmentering' er positivt korreleret hos typisk udviklede, hvorimod gode segmenterings-scorer hos personer med autisme ofte følges med dårlige integrations-scorer og omvendt.

I den aktuelle undersøgelse var der ikke en grænse for, hvor lang tid den enkelte deltager kunne bruge på en bestemt opgave. I stedet var der en øvre tidsgrænse for løsning af hver halvdel af opgavehæftet. Det giver mulighed for en mere strategisk tilgang til løsning af opgaverne. Man kan altså sætte spørgsmålstejn ved, om EFT, i den udformning den her er blevet brugt, nu også reelt tester en 'spontan tendens'. Det må som nævnt følge af Happé og Friths (2006) hypotese om kontrasterende spontane tendenser, at en spontan tendens til fx detaljefokus ikke udelukker, at personen samtidig kan være habil eller ligefrem god til at fokusere på sammenhæng og helhed under nogle betingelser. En tidsgrænse på de enkelte opgaver kunne i en opfølgende undersøgelse være med til at belyse om en positiv sammenhæng med inferensfærdigheder primært gælder strategisk færdighed i at finde detaljer i figurene.

Ud over spørgsmålet om en spontan tendens overfor strategisk løsning af opgaverne kan man også spørge, om EFT rent faktisk afdækker, hvor godt deltageren (spontan eller strategisk) finder en detalje inden i en helhed, som det er blevet antaget i tidligere forskning, eller om ikke deltageren faktisk finder en lille figur på trods af mange *andre mindre figurer og/eller streger*. Den lille figur kan på sin vis beskrives som en afsluttet helhed, og den store figur er ikke en entydig helhed, men består af mange mindre figurer og streger. Denne tolkning er et stykke af vejen i tråd med FI/FD-teorien (deltageren skal *se bort fra* noget). Men der er den væsentlige forskel, at det ikke kun er en helhed/struktur, men faktisk også *detaljer*, deltageren skal se bort fra, og at det kan siges at være en

helhed, som deltageren rent faktisk identificerer, når han eller hun finder den lille figur. Med denne tolkning er det muligt at se en meningsfuld forbindelse til inferensfærdigheder – hvis man antager at læseren udvælger kritiske informationer i teksten frem for andre for at drage de centrale inferenser. Denne anderledes tolkning af scorer på EFT kan måske særligt knyttes til den anvendte version af EFT, hvor der ikke er nogen mening knyttet til den store figur, hvilket der fx er i The Children's Embedded Figures Test (CEFT), hvor den store figur udgøres af en kendt figur, fx en barnevogn eller et ur (eksempler kan ses hos: Happé, 1999; Booth & Happé, 2008). CEFT udgives ikke længere (Witkin et al. 1971/2002, forlagets indledende kommentar), men det ville være interessant at sammenligne sammenhænge mellem scorer på de to forskellige udgaver af EFT med inferensfærdigheder for at se, om det har en effekt, at den store figur er et kendt objekt og ikke en kompleks geometrisk figur.

Konklusion

Ideen om læserens forventninger til sammenhæng har været indflydelsesrig i læseforståelsesforskningen. Læserens forventninger om sammenhæng er blevet kædet direkte sammen med, hvor mange inferenser læseren drager under læsning, og er blevet udpeget som potentiel bagvedliggende årsag til individuelle forskelle i inferensfærdigheder. Man har dog ikke undersøgt læseres forventninger om sammenhæng uafhængigt af læseres forståelse. I den aktuelle undersøgelse blev det afprøvet, om tendens til at fokusere på detaljer kan være (negativt) associeret til inferensfærdigheder, hvorved det indirekte afprøvedes, om et sammenfald mellem tendens til at fokusere på detaljer og lave forventninger om sammenhæng kan være muligt. Hertil blev der anvendt en test, der i tidligere forskning er blevet taget til udtryk for en tendens til at fokusere på detaljer, EFT.

Undersøgelsen viste imidlertid, at en tendens til at fokusere på detaljer, målt med EFT, nærmere er relateret til gode end dårlige inferensfærdigheder hos børn uden kendte vanskeligheder. Det kunne desuden ikke påvises, at relationen var særlig for inferensfærdigheder i læsning. Det er således muligt, at fokus på detaljer er relateret til læseforståelse generelt frem for inferensfærdigheder specifikt. Individuelle forskelle i problemløsningsfærdigheder kan til dels, men ikke fyldestgørende, ligge bag denne sammenhæng. Sammenhængen kan muligvis skyldes, at test af inferensfærdigheder og EFT begge trækker på (andre) generelle færdigheder. Den praktiske relevans af EFT-scorer som prædiktor for inferensfærdigheder er meget begrænset, da EFT ikke bidrager til prædiktion af inferensfærdigheder, når man tager højde for prædiktionsværdien af ordforråd.

En række vigtige begrænsninger for konklusionen er blevet fremhævet i diskussionsafsnittet. Begrænsningerne omhandler primært en væsentlig usikkerhed om, hvad EFT-scorer er udtryk for. Det synes dog klart, at EFT-scorer *ikke* kan tages til udtryk for en negativ indikator på læserens forventninger om sammenhæng.

Kapitel 7: Sammenfatning med perspektiver

Hvem er dårlige forståere?

I introduktionen til afhandlingen blev der rejst fire overordnede spørgsmål angående dårlige forståere. Det var, hvordan dårlige forståere kan defineres i forhold til andre læsere med læsevanskeligheder; hvilke mulige årsager der er til dårlig sprogforståelse i læsning; hvordan man konkret kan afgrænse dårlige forståere; og hvilke udfordringer der er, når man følger elever med specifikke vanskeligheder med sprogforståelse i læsning.

Hvordan dårlige forståere kan defineres, og hvilke årsager af psykolingvistisk karakter, der kan være til dårlig sprogforståelse i læsning, var emnet i kapitel 1. Ifølge den enkle læsemodel kræver god tekstforståelse sikre afkodningsfærdigheder *og* god sprogforståelse. Læsevanskeligheder kan dermed *både* opstå som følge af svage afkodningsfærdigheder *og* som følge af utilstrækkelig sprogforståelse. Dårlige forståere kan ud fra den enkle læsemodel defineres som elever med dårlig tekstforståelse, der ikke skyldes dårlig afkodning. Undersøgelser har vist, at problemer med en række færdigheder med betydning for sprogforståelse kan karakterisere dårlige forståere. Det er ordforråd og ordkendskab, grammatiske færdigheder, viden om teksters struktur, monitorering af egen forståelse, inferensfærdigheder og arbejdshukommelse. For nogle af disse færdigheder synes belægget for at opfatte dem som årsager til dårlig læseforståelse meget stærkt. Det gælder især for ordforråd og inferensfærdigheder. Resultater fra en lang række sammenligningsstudier og langtidsundersøgelser samt effektundersøgelser konvergerer og peger på en årsagssammenhæng mellem disse færdigheder og læseforståelse.

Efter læseundervisningens start, når eleverne er kommet godt i gang med at afkode, er det muligt at identificere elever med specifikke forståelsesvanskeligheder i læsning. Hvordan man kommer fra den generelle definition af dårlige forståere til en konkret afgrænsning af dårlige forståere, var emnet i kapitel 2. I forskning identificerer man som oftest dårlige forståere med det formål at undersøge potentielle årsager til forståelsesvanskelighederne nærmere. Det gør man ved at sammenligne dårlige forståere med læsere med god læseforståelse (gode forståere) på de færdigheder, man er interesseret i. En gennemgang af tidligere forskning viste, at det i sådanne undersøgelser er almindeligt at undlade at udtage læsere, der med udgangspunkt i deres baggrundshistorier kan have særlig risiko for forståelsesvanskeligheder i læsning. Dog blev der ikke anvendt samme eksklusionskriterier på tværs af undersøgelser. Eksklusionskriterier kunne fx gælde for elever med kendte sproglige og/eller kognitive vanskeligheder og elever, der læste på et andet- eller fremmedsprog. I nogle af undersøgelserne inddrog man desuden en kontrolvariabel. Også her var der dog stor variation i, *hvilken* kontrolvariabel der var tale om. I flere tilfælde var der en sammenhæng mellem tilvalg eller fravalg af kontrol med en bestemt færdighed og forskningsformålet. Det kunne fx være, at deltagerne var matchet på ordforråd, når man ønskede at undersøge betydningen af inferensfærdigheder for læseforståelse (som hos Cain og Oakhill (1999)). Disse fremgangsmåder, der i forskning anvendes til at afgrænse relativt homogene grupper kan føre

til indsigt i betydningen af udvalgte komponenter for dårlig sprogforståelse i læsning, som nye undersøgelser og ikke mindst praksis kan profitere af. I praksis er selve formålet med at identificere dårlige forståere dog et andet end i forskning. I praksis er formålet med en afgrænsning at blive opmærksom på eventuelt oversete vanskeligheder, så man kan iværksætte målrettet undervisning for de elever. I en afgrænsning med relevans for praksis synes det derfor ikke hensigtsmæssigt at inddrage kontrolvariable som fx ordforråd eller IQ. I den generelle definition af dårlige forståere indgår kun de to nøglefærdigheder; tekstforståelse og afkodning. I overensstemmelse med formålet med en afgrænsning i praksis synes det dog at være hensigtsmæssigt at anvende baggrundsoplysninger om fx kendte sproglige eller kognitive vanskeligheder (med henblik på en nærmere undersøgelse af elevernes tekstforståelse). Dette kan være et første trin i afgrænsningen af elever med forståelsesvanskeligheder. Derefter kan man så gøre brug af screeningsresultater (læseforståelses- og afkodningstest) til at afgrænse elever med specifikke forståelsesvanskeligheder, der ikke er indikeret i baggrundsoplysningerne.

Hvilke testmaterialer, der anvendes i afgrænsningen, er endnu en faktor, der kan have betydning for, hvilke deltagere der udpeges som dårlige forståere. Der er evidens for, at den valgte læseforståelsestest er helt central (Keenan et al., 2014; Nation & Snowling, 1997). Det er her en væsentlig pointe, at læseforståelsestest varierer i følsomhed for variation i sprogforståelse (fx Cutting & Scarborough, 2006). Hvis resultaterne på testen næsten udelukkende kan forklares af afkodningsfærdigheder, kan man *ikke* forvente at kunne afgrænse dårlige forståere: Der vil ikke være deltagere, der scorer lavt på læseforståelsestesten, som ikke også scorer lavt på en test af afkodningsfærdigheder. Brug af en læseforståelsestest, der ikke er følsom for sprogforståelse kan medføre, at *det ser ud* som om, at der ikke er elever, der har problemer med sprogforståelse i læsning.

Vægtningen af de krav som en læseforståelsestest stiller til de forskellige komponenter i sprogforståelse kan sandsynligvis også have betydning for, hvilke dårlige forståere der afgrænses med testen: Anvendes en test med store krav til ordkendskab, men få krav til inferensfærdigheder, kan man forvente, at primært dårlige forståere med ordforrådsproblemer afgrænses, og omvendt kan en test med store krav til inferensfærdigheder, men simple krav til ordkendskab, forventes først og fremmest at medføre, at elever med ringe inferensfærdigheder afgrænses.

Der er evidens for, at anvendelse af forskellige afkodningstest kan have en betydning for, hvilke deltagere der udpeges som dårlige forståere (Keenan et al., 2014; Rønberg & Petersen, 2015a). For at sikre uafhængighed af læseforståelsesscoren og afkodningsscoren, der anvendes i afgrænsningen, kan man undersøge afkodningsfærdigheder med en separat test. Det synes desuden oplagt, at man for at sikre et mål for *automatiserede* afkodningsfærdigheder *kombinerer* mål for afkodning af nonord og afkodning af ord (som hos Catts et al. (2006)), og at man undersøger *effektivitet* (og ikke bare præcision) af afkodningsfærdigheder (som hos Tong et al. (2011)).

Den konkrete grænsedragning på de anvendte test er endnu en faktor, der varierer stærkt på tværs af undersøgelser. I nogle undersøgelser afgrænses dårlige forståere udelukkende ved hjælp af simple grænseværdier på testene af læseforståelse og afkodningsfærdigheder hver for sig (cut-off definition). I andre undersøgelser er det grundlæggende princip at sikre, at dårlige forståere skal have væsentlig dårligere læseforståelse, end man ville forvente ud fra deres afkodningsfærdigheder (diskrepansdefinition). Der skal altså være et misforhold af en vis størrelse mellem de to

færdigheder. Dette misforhold fremgår også af den generelle definition af dårlige forståere – vanskelighederne med at forstå tekster skal ikke kunne forklares af dårlig afkodning. Jeg argumenterede i kapitel 2 for, at en afgrænsning af dårlige forståere bør indeholde elementer fra *både* en cut-off- og en diskrepansdefinition. En øvre simpel grænseværdi for læseforståelse er nødvendig for at sikre, at de udpegede dårligere forståere reelt har dårlig læseforståelse sammenlignet med deres jævnaldrende. En diskrepans mellem afkodningsfærdigheder og læseforståelsesniveau er desuden nødvendig for at sikre, at den dårlige læseforståelse ikke kan forklares ud fra afkodningsniveauet. En nedre grænseværdi for afkodningsfærdigheder kan desuden sættes for at afgrænse dårlig forståere fra elever med sammensatte vanskeligheder.

Hvilken stabilitet af en afgrænsning af dårlige forståere, man kan forvente, var emnet i kapitel 3. Jeg undersøgte stabiliteten af profilen 'dårlig forstår' over kort tid og på tværs af to indholdsmæssigt parallelle læseforståelsestest, der blev afviklet under forskellige betingelser (knyttet til henholdsvis gruppetestning og individuel testning). Undersøgelsen viste en moderat stabilitet af profilen 'dårlig forstår'. Andre undersøgelser understøtter undersøgelsens resultater: Stabiliteten synes at være meget høj, når man følger dårlige forståere over tid, men stabiliteten på tværs af forskellige test synes at være begrænset.

Undersøgelsen i kapitel 3 supplerede desuden listen fra kapitel 2 med faktorer, der kan have betydning for, hvilke læsere der bliver udpeget som dårlige forståere, idet *testbetingelsen* så ud til at have en selvstændig betydning for dårlige og gode forståeres resultater på læseforståelsestest. De anvendte test var henholdsvis en gruppetest og en individuel test med hver sit sæt af betingelser. Det syntes muligt, at ustabilitet på tværs af testene kunne komme fra testenes forskellige muligheder for at udøve selvregulerende adfærd, eventuelt i kombination med forskelle som højtlesning over for stillelæsning, tidsgrænse over for ingen tidsgrænse og skriftlige besvarelser over for mundtlige besvarelser. Forskning i opgaveorienteret læsning tyder netop på, at forskelle i læsers selvregulerende adfærd kan påvirke læsernes opgaveløsning og dermed deres resultater på læseforståelsestest (Cerdán et al., 2009), og anden forskning tyder på, at test med højtlesning generelt er lettere end test med stillelæsning, måske særligt for elever med vanskeligheder (Elgart, 1978; Swalm, 1972). Undersøgelsen kan ikke pege på, hvilke af disse forhold der er mest afgørende, eller på, i hvilken grad der er interaktion mellem fx betydningen af krav til selvregulerende adfærd i en test og tilstedeværelse af en tidsgrænse eller ej. Betydningen af de forskellige elementer for afgrænsning af dårlige forståere er et oplagt emne for fremtidig forskning.

Inferensfærdigheders betydning for læseforståelse

Adskillige undersøgelser har vist, at inferensfærdigheder er en vigtig komponent i læseforståelse. I kapitel 4, 5 og 6 prøvede jeg at komme et spadestik dybere. I introduktionen til afhandlingen rejste jeg en række spørgsmål angående betydningen af inferensfærdigheder i læseforståelse. Spørgsmålene var, om det er selve det at kunne drage en følgeslutning, der er afgørende for læseforståelse; om den type af viden, det er nødvendigt at aktivere for at drage inferensen, har en betydning; og om det er et springende punkt at kunne udnytte sproglige signaler om inferens i teksten. Desuden stillede jeg spørgsmål til relationen mellem inferensfærdigheder og andre

færdigheder. Spørgsmålene var, hvordan inferensfærdigheder spiller sammen med ordforråd og arbejdshukommelse i læseforståelse, og om et fokus på detaljer er i modstrid med gode inferensfærdigheder.

I kapitel 4 undersøgte jeg betydningen af forskellige aspekter af inferensfærdigheder for læseforståelse i stedet for at fokusere på forskellige inferenstyper, som det ofte har været tilfældet i forskning (fx Kendeou et al., 2008; Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999). Jeg undersøgte først betydningen af modalitet (om inferensen drages fra sprogligt eller visuelt materiale). Undersøgelsen viste, at der er sammenhæng mellem nonverbal og verbal inferensfærdighed, sådan at man med god grund kan tale om en i hvert fald delvist generaliseret færdighed på tværs af modaliteter. Resultaterne støtter dermed konklusioner fra tidligere undersøgelser med bl.a. voksne og børn med sprogvanskeligheder (Gernsbacher et al., 1990; Bishop & Adams, 1992; Yussen et al. 1989), der dog har undersøgt nonverbal inferensfærdighed i sprogligt krævende opgaveformater. Men undersøgelsen viste også, at det først og fremmest er den *verbale* inferensfærdighed, der er vigtig for *læseforståelse*. Det er altså ikke selve det at kunne drage en følgeslutning (uanset modalitet), der er afgørende for inferensfærdighedens betydning for læseforståelse. Det er afgørende at kunne drage følgeslutninger fra *sprogligt* materiale. Dette er ikke tidligere vist, og det er derfor oplagt, at undersøgelser forsøger at replicere resultaterne med andre deltagere og nye materialer. En vigtig pointe i sådanne undersøgelser må være, at man – som i den aktuelle undersøgelse – undersøger nonverbal inferensfærdighed i et opgaveformat, der ikke stiller krav til sproglig processering. Man kan her indvende, at det selv i den nonverbale opgave i den aktuelle undersøgelse er oplagt, at deltageren kan ræsonnere sig verbalt frem til et svar. Og man kan spørge, i hvilken grad det er muligt at drage inferenser helt uden støtte i sproget. Det kan denne undersøgelse ikke svare på, men resultaterne tyder på, at der er indsigter at hente ved at minimere de sproglige krav i inferensopgaver, når man ønsker at undersøge betydningen af inferenser fra ikke-sprogligt materiale for læseforståelse.

Jeg undersøgte dernæst, om den type af viden, det er nødvendigt at aktivere for at drage en inferens (emnespecifik viden eller almenmenneskelige erfaringer), spiller en rolle for inferensfærdighedens betydning for læseforståelse. Undersøgelsen indikerede, at den videnstype, der skal aktiveres i forbindelse med inferensdragning, *kan* have betydning for forskelle i læseforståelse. Det synes at være sådan, at man har et forspring i læseforståelse, hvis man ud over at kunne trække på umiddelbar tilgængelig viden i form af personlige erfaringer i forbindelse med inferenser – hvilket ofte er nødvendigt i narrativer – også er god til at mobilisere andenhånds, tillært emnespecifik viden om verden – hvilket ofte er nødvendigt i fagtekster. Dette resultat er i overensstemmelse med tidligere undersøgelser, der har vist, at emnespecifik viden har større betydning for forståelse af fagtekster end forståelse af narrativer (Best et al., 2008), og at videns tilgængelighed har betydning for sandsynligheden for, at læseren drager en inferens, der kræver aktivering af denne viden (Barnes & Haefele-Kalvaitis, 1996). I den aktuelle undersøgelse var en distinktion mellem forståelse af tekster, der krævede inferenser med emnespecifik viden og forståelse af tekster, der krævede inferenser med erfaringsbaseret viden baseret på en forsøgsvis opdeling af teksterne i en læseforståelsestest, der ikke var udviklet til dette formål, og opdelingen kan naturligvis diskuteres. Undersøgelsen kan derfor ses som en indledende undersøgelse af videnstypers betydning, og det er oplagt at efterprøve undersøgelsens resultater i mere kontrollerede

forsøg med materialer udviklet specielt til undersøgelsesformålet. Et oplagt design vil være et design, der tager udgangspunkt i test af 'genre-løse inferenser', der udelukkende adskiller sig på den type af (kendt) viden, deltagerne må aktivere for at drage inferenserne. Man kan så sammenligne sammenhængene mellem disse to inferens-scorer og forståelse af henholdsvis fagtekster og narrativer. På baggrund af resultaterne i kapitel 3, hvor testbetingelsen syntes at have betydning for præstationerne, er det ligeledes oplagt at sikre, at alle de anvendte inferens- og forståelsestest i en undersøgelse af videnstypens betydning har samme format og afvikles under samme betingelser. I den aktuelle undersøgelse adskilte testen af inferenser fra narrativer sig fra de andre test ved at være en test med stillelæsning og valgmuligheder (de øvrige test inkluderede højtlesning og åbne spørgsmål) og scoren var en effektivitetsscore (for de øvrige test var der tale om præcisionsscorer). Målet for inferenser fra narrativer korrelerede netop gennemgående svagere med de forskellige forståelsesmål, end målet for inferenser fra fagtekster gjorde.

Det tredje aspekt, der blev undersøgt i kapitel 4, var sproglig signalering. Jeg undersøgte sammenhængen mellem færdighed i at drage inferenser, der var signaleret sprogligt i teksten og inferenser, der ikke var signaleret i teksten, og desuden deres selvstændige betydning for læseforståelse. Der var sammenhæng mellem færdighed i at drage inferenser med sproglige signaler (henvisende substantiver med bestemthedsendelse eller konnektiver) og inferenser uden sådanne signaler. Men evne til at drage de signalerede og ikke-signalerede inferenser havde dertil hver især en selvstændig betydning for læseforståelse. Konklusionen om en selvstændig betydning var især stærk for de signalerede inferenser. Selv om udnyttelse af sproglige signaler ikke på baggrund af denne undersøgelse kan siges at være *det* springende punkt ved inferensfærdigheder, når det kommer til betydning for læseforståelse, så tyder resultaterne på, at udnyttelse af sproglige signaler i hvert fald er en væsentlig faktor. Et konventionelt fokus på en distinktion mellem forskellige inferenstyper kunne eventuelt danne grundlag for en anden fortolkning af resultaterne. Der er ikke i det aktuelle datasæt mulighed for at undersøge betydningen af konventionelt beskrevne inferenstyper.

I kapitel 5 undersøgte jeg først, om inferensfærdigheder spillede en rolle for læseforståelse, når man tager højde for betydningen af ordforråd og verbal arbejdshukommelse for læseforståelse, og jeg undersøgte dernæst samspillet mellem de tre færdigheder i læseforståelse. Mange undersøgelser har vist, at inferensfærdigheder og ordforråd er relateret til forskelle i læseforståelse, men ofte er effekten af færdighederne blevet undersøgt enkeltvis (fx Nation et al., 2004; Cain og Oahill, 1999). Et samspil mellem ordforråd og inferensfærdigheder i læseforståelse synes dog oplagt. Semantiske relationer mellem ord i teksten udgør ofte delvist det vidensgrundlag, som en inferens kan drages på, og omvendt er det ofte nødvendigt at inferere de enkelte ords præcise betydning ud fra konteksten (det vil sige resten af teksten). Verbal arbejdshukommelse var en selvskreven kontrolvariabel i forbindelse med inferensfærdigheders betydning for læseforståelse, da arbejdshukommelse ofte bliver fremhævet som nødvendig for inferensfærdigheder.

Undersøgelsen viste, at inferensfærdigheder havde en selvstændig betydning for læseforståelse i konkurrence med både ordforråd og verbal arbejdshukommelse. Dette resultat styrker dermed det eksisterende belæg for inferensfærdigheders betydning for læseforståelse (fx Cain et al., 2004). Resultatet er endvidere interessant set i lyset af konklusionen om den sproglige modalitets betydning for inferensfærdigheders rolle i læseforståelse (kapitel 4). Konklusionen fra kapitel 4 kan

med resultatet om inferensfærdigheders selvstændige betydning for læseforståelse nuanceres. Den særlige betydning af *verbal* inferensfærdighed for læseforståelse synes *ikke* at kunne reduceres til et spørgsmål om sproglige færdigheder som ordforråd og verbal arbejdshukommelse.

Undersøgelsen viste desuden, at ordforrådets effekt på læseforståelse var fuldt medieret af inferensfærdigheder. Denne mediationseffekt var signifikant selv efter kontrol for evne til at holde og manipulere sprogligt materiale i arbejdshukommelsen. Ordforråd synes således at støtte inferensfærdigheder, og inferensfærdigheder er nødvendige for god læseforståelse. Undersøgelsen understøtter dermed resultater fra andre nyere undersøgelser, der tyder på afgørende samspil mellem ordforråd og inferensfærdigheder i læseforståelse (Cain & Oakhill, 2014; Cromley & Azevedo, 2007; Li & Kirby, 2014). Undersøgelsen bidrager desuden til belysningen af dette samspil mellem ordforråd og inferensfærdigheder i læseforståelse ved at tage højde for verbal arbejdshukommelse. Cromley og Azevedo (2007) påpegede netop, at de ikke i deres undersøgelse havde taget højde for arbejdshukommelse, som ellers var en oplagt interfererende variabel.

Det var overraskende, at den direkte effekt af ordforråd på læseforståelse ikke var signifikant i den aktuelle undersøgelse, som den var i undersøgelsen af Cromley & Azevedo. En del af forklaringen kan muligvis ligge i den anvendte ordforrådstest, som var en test af ordforrådets bredde. Ordforrådets *dybde* har netop vist sig at have en særlig betydning for læseforståelse sammenlignet med ordforrådets bredde (Ouellette, 2006). Det synes derfor muligt, at den direkte effekt af ordforråd målt med test af ordforrådets dybde, vil vise sig at være signifikant. Ligeledes kan andre mål for arbejdshukommelse vise sig at kunne forklare en større del af den indirekte effekt af ordforråd gennem inferensfærdigheder. Dette er særligt sandsynligt, hvis testene stiller højere krav til semantisk viden og procesering end den anvendte cifferspændviddetest, som det fx var tilfældet i de verbale arbejdshukommelsesopgaver hos Seigneuric et al. (2000) og Oakhill et al. (2011). Spørgsmålet er imidlertid, hvordan resultater med sådanne arbejdshukommelsesmål ville skulle tolkes. Hvis arbejdshukommelsestest indebærer processering af ordbetydninger og færdiggørelse af sætninger som hos Seigneuric et al. og Oakhill et al., er der risiko for, at forskelle i scorer på testene er konfunderet med forskelle i ordforråd og inferensfærdigheder.

I kapitel 6 undersøgte jeg, om en tendens til at fokusere på detaljer var relateret negativt eller positivt til inferensfærdigheder. Beskrivelsen af en tendens til at fokusere på detaljer, også kaldet 'svag central kohærens' (*weak central coherence*), stammer fra autismedforskningen (Frith, 2003) og har visse ligheder med beskrivelsen af dårlige forståeres lave forventninger om sammenhæng, som man finder i læseforskningen (van den Broek, Ridsen, & Husebye-Hartmann, 1995). Såvel tendens til at fokusere på detaljer som lave forventninger om sammenhæng er blevet sat i forbindelse med ringe inferensfærdigheder (fx Norbury & Nation, 2011; Oakhill & Cain, 2007). Den anvendte test af tendens til at fokusere på detaljer i den aktuelle undersøgelse, *the Embedded Figures Test* (EFT), er blevet anvendt som en test af detaljefokus/svag central kohærens i mange tidligere undersøgelser. EFT er dog også blevet anvendt i et andet forskningsområde (Field (in)dependence), inden for hvilket man kan finde forskere, der forventer en sammenhæng mellem EFT-scorer og *gode* læseforståelses- og inferensfærdigheder (fx Bruininks, 1969; Pitts & Thompson, 1984). Undersøgelsen viste, at der var en positiv sammenhæng mellem EFT-scorer og inferensfærdigheder. Hvis EFT tages til udtryk for en tendens til at fokusere på detaljer, er der altså ikke umiddelbart en sammenhæng mellem detaljefokus og ringe inferensfærdigheder. Tværtimod synes høje EFT-scorer

at følges med gode inferensfærdigheder. Det samme viste sig dog at være tilfældet med EFT-scorer og scorer for bogstavelig (ikke-inferentiel) forståelse. Den positive sammenhæng synes altså ikke at være specifik for inferensfærdigheder i læseforståelse. Sammenhængen kunne i nogen grad forklares af nonverbale problemløsningsfærdigheder, og EFT bidrog desuden ikke signifikant til prædiktionen af inferensfærdigheder efter kontrol for ordforråd. Selv om der må siges at være usikkerhed om, *hvad* EFT tester, tyder resultaterne fra den aktuelle undersøgelse på, at EFT-scorer i hvert fald ikke kan tages til udtryk for en bagvedliggende årsag til ringe inferensfærdigheder, som lave forventninger om sammenhæng. Hvorvidt tendens til at fokusere på detaljer undersøgt på en anden måde end med EFT er positivt eller negativt relateret til inferensfærdigheder, er et spørgsmål, der nu står åbent. Resultatet fra kapitel 4 om betydningen af den sproglige modalitet for inferensfærdigheders betydning for læseforståelse kan tyde på, at det i fremtidige undersøgelser af detaljefokus og inferensfærdigheder kan være hensigtsmæssigt at undersøge detaljefokus med sprogligt materiale.

Årsager til dårlig forståelse

Undersøgelserne i afhandlingen har overordnet set haft to fokuser. I første del var dårlige forståere specifikt i fokus. I anden del blev betydningen af inferensfærdigheder for læseforståelse undersøgt med data fra et bredt udvalg af læsere. Hvorvidt resultaterne om betydningen af aspekter af inferensfærdigheder for læseforståelse og samspillet med andre færdigheder også gælder for dårlige forståere specifikt, kan undersøges i opfølgende undersøgelser, men det er formodningen, at fx et samspil mellem færdigheder som ordforråd og inferensfærdigheder ikke mindst vil være relevant for læsere med dårlig forståelse. Som nævnt i den kortfattede oversigt i kapitel 1 over de mulige årsager til forståelsesvanskeligheder, som er bedst dokumenteret i den hidtidige forskning, er der netop *en række* mulige årsager til dårlig forståelse på tværs af dårlige forståere – og et samspil mellem færdigheder/problemer hos den enkelte synes derfor oplagt. Tabel 1 kan være med til at illustrere, at det ikke altid er et spørgsmål om *enten* ordforrådsproblemer *eller* dårlige inferensfærdigheder og arbejdshukommelse hos dårlige forståere.

		Ordforråd	Inferensfærdighed	Arbejdshukommelse
	1	-1,3	-2,5	-2,1
	2	-0,5	-0,7	-0,7
	3	-1,3	-1,8	0,0
	4	-0,7	1,2	0,7
Dårlige	5	0,6	-0,8	-1,4
forståere	6	-0,1	0,0	1,4
	7	-0,8	-0,2	-2,1
	8	-1,1	-1,7	0,0
	9	-1,3	0,5	-0,7
	10	-1,3	-2,2	-0,7

Tabel 1. Resultater (z-scorer) for 10 dårlige forståere på test af ordforråd, inferensfærdigheder og verbal arbejdshukommelse. Mørkeblå farve markerer scorer mellem 2,1 og 3 standardafvigelser under gennemsnittet for de 54 elever, der deltog ved T2. Mellembå markerer scorer mellem 1,1 og 2 standardafvigelser under gennemsnittet, og lyseblå markerer scorer mellem 0,1 og 1 standardafvigelse under gennemsnittet.

I tabellen indgår resultater fra 10 elever, der ved T1 blev afgrænset som dårlige forståere ved hjælp af gruppetest, og som også adskilte sig fra majoriteten af de gode forståere ved gentestning (T2) med en individuel læseforståelsestest. De 10 elevers resultater på test af ordforråd, inferensfærdigheder og arbejdshukommelse er udtrykt i z-scorer. En z-score på 0 svarer til det gennemsnitlige resultat for alle 54 elever, der deltog i undersøgelsen ved T2. Tabel 1 viser, at der sjældent er tale om en isoleret vanskelighed med én af de tre færdigheder, og at dårlige forståere kan have meget forskellige profiler. Cornoldi, De Beni og Pazzaglia (1996) fandt på lignende vis forskellige profiler for dårlige forståere på tværs af forskellige test af færdigheder med betydning for forståelse.

Tabel 1 viser desuden indirekte, at der også må være andre årsager til dårlig sprogforståelse i læsning på spil i elevgruppen end de undersøgte færdigheder, idet ikke alle de dårlige forståere har (store) problemer med én eller flere af disse. Det gælder helt tydeligt for deltager 6, der har gennemsnitligt ordforråd og gennemsnitlige inferensfærdigheder samt god arbejdshukommelse, men også for deltager 4 og deltager 2 er det rimeligt at sætte spørgsmålstejn ved om ordforråd, inferensfærdigheder og arbejdshukommelsesressourcer kan forklare forståelsesvanskelighederne. Oplagte kandidater som årsag til dårlig sprogforståelse i læsning hos disse elever er de kendte årsager som viden om teksters struktur og monitorering af egen forståelse.

Ved at sammenligne forståelsen hos elever med forskellige profiler og følge elevernes færdigheder over tid, kan man sandsynligvis opnå yderligere indsigt i samspillet mellem færdigheder i læseforståelse og kompleksiteten i forståelsesvanskeligheder. Hvis man derimod negligerer samspillet, kan man risikere at bidrage til vedligeholdelse af den fejlopfattelse, at *enten* begrænsede sproglige færdigheder (som ordforråd og grammatiske færdigheder) *eller* svage højere

kognitive færdigheder (som inferensfærdigheder og monitorering) er årsagen til dårlig læseforståelse.

Forskningsperspektiver

I den gennemgåede forskning og i forslaget til en afgrænsning af dårlige forståere i kapitel 3 blev dårlige forståere afgrænset blandt elever uden kendte vanskeligheder. Det er en separat forskningsopgave at undersøge ligheder og forskelle mellem forskellige grupper med forståelsesvanskeligheder. Det er fx oplagt at undersøge, om deltagere med autisme og kendte sproglige vanskeligheder har de *samme* forståelsesvanskeligheder som gruppen af dårlige forståere uden andre vanskeligheder – og om fx samspillet mellem ordforråd og inferensfærdigheder gælder i samme grad for disse grupper, hvor andre kognitive faktorer også må medregnes.

Den aktuelle undersøgelse tyder på, at testbetingelsen for nogle deltagere kan have en betydning for deres scorer på læseforståelsestest, sådan at de muligvis kun har dårlig (eller god) læseforståelse under *nogle* betingelser. Det indikerer, at det kan være vigtigt at sikre overensstemmelse på tværs af de anvendte færdighedstest i undersøgelser af læseforståelse og komponenter i sprogforståelse i læsning. Hvis der er uoverensstemmelser i testbetingelserne for de anvendte færdighedstest i en undersøgelse, *kan* uventede resultater muligvis skyldes sådanne uoverensstemmelser. Resultaterne fra den aktuelle undersøgelse tyder på, at det fx kan gælde forskelle, når det kommer til, hvor meget eller lidt af opgaveløsningens forløb der styres af deltageren selv, hvilken måde deltageren skal svare (skriftligt eller mundtligt), og om testene afvikles under et tidspres eller ej. Undersøgelser kan desuden afdække den selvstændige betydning af de nævnte elementer i forskellige testbetingelser knyttet til typiske gruppetest og individuelle test. Hvor stor en rolle spiller varierende krav til selvregulerende adfærd fx på tværs af test? Og hvilken effekt af tidspres er der på forståelsesniveauet?

I den videre afdækning af stabiliteten for grupper af dårlige forståere på tværs af test, synes det relevant, at undersøgelser til en start ser nærmere på stabiliteten, når man tester med præcis samme test og præcis samme betingelser. Hvilken stabilitet kan man i udgangspunktet forvente på trods af eventuel begrænset testpålidelighed samt regressionseffekt? Dernæst vil det være interessant at se på forskelle i *indholdet* i forskellige test og undersøge, i hvilken grad indholdsmæssige forskelle i testene kan påvirke stabiliteten af fx en afgrænsning af dårlige forståere. Kan ustabilitet skyldes, at læseforståelsestestene i forskellig grad er følsomme for sprogforståelse? Og videre: Kan ustabilitet skyldes, at testene stiller forskellige krav til de forskellige komponenter i sprogforståelse i læsning? Sådanne undersøgelser kan være med til at kaste lys over tilsyneladende modstridende resultater fra sammenligningsundersøgelser med dårlige og gode forståere. Undersøgelsen af Keenan et al. (2014) er den første af sin art og er dermed et vigtigt bidrag i afdækningen af stabiliteten af grupper af dårlige forståere på tværs af forskellige test. Men konklusionen om en begrænset stabilitet på tværs af test synes at kunne nuanceres i fremtidige undersøgelser. Er det reelt stabiliteten af forståelsesvanskeligheder, der er begrænset, eller er den begrænsede stabilitet i højere grad et tegn på indholdsmæssige forskelle på test? Kan den begrænsede stabilitet ligefrem være et symptom på begrænset indholdsvaliditet for nogle læseforståelsestests vedkommende?

Effektundersøgelser kan være med til at belyse betydningen af henholdsvis sprogligt signalerede og ikke-signalerede inferenser for læseforståelse yderligere. Sådanne undersøgelser kan fx vise, i hvilken grad der er effekt af undervisning i inferenser med sproglig signalering, og i hvilken grad en eventuel effekt overføres til færdighed i at drage ikke-signalerede inferenser – og omvendt. Effektundersøgelser kan også være et redskab til yderligere belysning af samspillet mellem inferensfærdigheder og ordforråd. Kan inferensfærdigheder fx forbedres ved at arbejde med centrale ordbetydninger for udvalgte inferenser i undervisningstekster? Og kan en sådan undervisning måske føre til generelt bedre inferensfærdigheder, når eleverne læser andre tekster? Kan inferensundervisning have effekt på elevers tilegnelse af ordbetydninger under selvstændig læsning? Effektundersøgelser kan også være med til at afdække yderligere samspil mellem inferensfærdigheder og *andre* færdigheder med betydning for sprogforståelse i læsning. Kan et indledende arbejde med tekstgenrer og strukturer fx være en støtte for inferensfærdigheder? Og kan en introduktion til vigtigheden af at monitorere egen forståelse øge effekten af inferensundervisning?

Undervisningsperspektiver

Den skitserede metode til en konkret grænsedragning mellem dårlige forståere og andre læsere, der er i overensstemmelse med den generelle definition af dårlige forståere synes at være mulig at indføre i praksis. Imidlertid er der en række udfordringer for en sådan gennemførelse. For det første fastlægges grænserne (både de simple grænseværdier for læseforståelse og afkodning og diskrepansgrænsen) på baggrund af resultaterne fra en større elevgruppe. Metoden kræver altså adgang til resultater på de valgte test af nøglefærdighederne for et bredt udsnit af læsere på mange forståelses- og afkodningsniveauer. Et alternativ er anvendelse af normerede test (som hos fx Cain & Oakhill, 1999), hvilket der dog i dansk praksis er begrænset mulighed for. De mange mulige årsager til dårlig sprogforståelse i læsning er også en udfordring, der stiller store krav til den anvendte læseforståelsestest. Testen må være følsom over for de forskellige komponenter i sprogforståelse. Udvikling af gyldige læseforståelsestest (der er følsomme for komponenter i sprogforståelse) og normering af disse vil således forbedre mulighederne for at identificere dårlige forståere i praksis. Udvikling af test af de forskellige komponenter i sprogforståelse, der kan anvendes i praksis, vil desuden kunne bidrage til en kvalificeret udredning af vanskelighederne og øge mulighederne for at kunne sætte ind med en målrettet undervisning.

At alle de undersøgte aspekter af inferensfærdighed i kapitel 4 havde betydning for læseforståelse kan have interessante implikationer for undervisning i at drage inferenser. Effekten af undervisningsideer med udgangspunkter i disse resultater kan afprøves i træningsundersøgelser. Eksistensen af en generaliseret (dvs. tvær-modal) inferensfærdighed kan umiddelbart lede til den antagelse, at billedhistorier kan anvendes i inferensundervisning, fx til at tydeliggøre behovet for at drage inferenser fra tekster – og måske endda anvendes som eneste stimulusmateriale i en inferensundervisning for yngre børn, der endnu ikke læser. Men når kun den sproglige inferensfærdighed har betydning for læseforståelse er det muligt, at et fokus på billedmateriale i undervisningen ikke vil lede til en overføringseffekt til læseforståelse, fordi undervisningen primært

styrker den inferensfærdighed, der ikke har stor betydning for læseforståelse. At kunne redegøre sprogligt for sin forståelse og ræsonnementer er netop et element, der fremhæves igen og igen i succesfuld forståelsesundervisning (National Reading Panel, 2000), og det er derfor sandsynligt, at dette element fortsat vil være en hjørnesten i fremtidig succesfuld inferensundervisning – trods det at en generaliseret inferensfærdighed formodentlig findes.

At lære om de forskellige slags sproglige signaler til inferens i tekster kan muligvis være en indgangsvinkel til at blive opmærksom på den implicit udtrykte information, man skal inferere sig til, når man læser tekster. Efter en introduktion til forskellige signaler til inferens kan man gå på jagt efter disse i både konstruerede og autentiske tekster og studere de inferenser, der signaleres. En anden indgangsvinkel til inferensundervisning kan være at gøre eleverne opmærksomme på de forskellige typer af viden, man kan drage inferenser med. Sidst, men ikke mindst, kan den aktuelle undersøgelses resultater indikere, at elever på mellemtrinnet særligt vil kunne have udbytte af undervisning med fokus på inferenser, der kræver aktivering af emnespecifik andenhånds tillært viden om verden.

Samspillet mellem ordforråd og inferensfærdigheder, som blev belyst i kapitel 5, giver også stof til eftertanke i forbindelse med undervisning. Det synes således muligt, at en inferensundervisning med fordel kan indebære elementer af undervisning i ordkendskab. Det kan fx være undervisning i udvalgte ordbetydninger, der er centrale for inferenser i tekster, som elever senere skal læse og løse opgaver til. Det er også oplagt, at ordkendskabsundervisning kan indeholde elementer af inferensundervisning. Vi ved, at en stor del af børns ordtilegnelse sker indirekte via de tekster, de læser, og at man ofte kan inferere sig til dele af nye ords betydning ud fra den kontekst, de nye ord indgår i. Inferensundervisning med fokus på nye ordbetydninger i tekster synes således at have et stort potentiale. Som nævnt skal der effektundersøgelser til, for at man kan give nærmere anvisninger om, hvilket indhold i inferensundervisning der fører til de bedste resultater for forskellige aldersgrupper og elever med forskellige behov.

Referencer

- Anderson, R. C. & Freebody, P. (1981). Vocabulary knowledge. I J. T. Guthrie (Red.), *Comprehension and teaching: research reviews* (s. 77–117). Newark, Del.: International Reading Assn.
- Andreassen, R. & Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading*, 33(3), 263–283.
- Arnbak, E. & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research*, 44(3), 229–251.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559.
- Baker, L. & Stein, N. (1978). *The development of prose comprehension skills* (Technical report No. 102). University of Illinois at Urbana-Champaign.
- Barnes, M. A., Dennis, M. & Haefele-Kalvaitis, J. (1996). The Effects of Knowledge Availability and Knowledge Accessibility on Coherence and Elaborative Inferencing in Children from Six to Fifteen Years of Age. *Journal of Experimental Child Psychology*, 61(3), 216–241.
- Barnes, M., Johnston, A. M. & Dennis, M. (2007). Comprehension in a Neurodevelopmental Disorder, Spina Bifida Myelomeningocele. I K. Cain & J. Oakhill (Red.), *Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective* (s. 193–217). New York: Guilford Press.
- Baumann, J., Edwards, E., Boland, E., Olejnik, S. & Kame'enui, E. (2003). Vocabulary Tricks: Effects of Instruction in Morphology and Context on Fifth-Grade Students' Ability to Derive and Infer Word Meanings. *American Educational Research Journal*, 40(2), 447–494.
- Baumann, J., Edwards, E., Font, G., Tereshinski, C., Kame'enui, E. & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly*, 37(2), 150–176.
- Beck, I. L. & McKeown, M. G. (2013). *Bringing Words to Life, Second Edition, Robust Vocabulary Instruction* (2. ed.). Guilford Publications.
- Beck, I. L., Perfetti, C. A. & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, 74(4), 506–521.
- Best, R. M., Floyd, R. G. & McNamara, D. S. (2008). Differential Competencies Contributing to Children's Comprehension of Narrative and Expository Texts. *Reading Psychology*, 29(2), 137–164.
- Bishop, D. V. & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential Meaning. *Journal of Speech & Hearing Research*, 35(1), 119.
- Bishop, D. V. & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological bulletin*, 130(6), 858.

- Booth, R. & Happé, F. (2010). «Hunting with a knife and fork»: Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, 107(4), 377–393.
- Bowers, P. N., Kirby, J. R. & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills A Systematic Review of the Literature. *Review of Educational Research*, 80(2), 144–179.
- Bowyer-Crane, C. & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology*, 75(2), 189–201.
- Bruininks, R. H. (1969). Auditory and visual perceptual skills related to the reading performance of disadvantaged boys. *Perceptual and Motor skills*, 29, 179–186.
- Buch-Iversen, I. (2010). *Betydningen av inferens for leseforståelse: effekter av inferenstrening*. Universitetet i Stavanger, Nasjonalt senter for leseopplæring og leseforskning, Det humanistiske fakultet.
- Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology*, 21(3), 335–351.
- Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory*, 14(5), 553–569.
- Cain, K. (2007a). Deriving word meanings from context: does explanation facilitate contextual analysis? *Journal of Research in Reading*, 30(4), 347–359.
- Cain, K. (2007b). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, 28(4), 679–694.
- Cain, K. & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology*, 14(2), 187–201.
- Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, 11(5-6), 489–503.
- Cain, K. & Oakhill, J. (2004). Reading Comprehension Difficulties. I T. Nunes & P. Bryant (Red.), *Handbook of Children's Literacy* (s. 313–338). Springer Netherlands.
- Cain, K. & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *The British Journal of Educational Psychology*, 76(4), 683–696.
- Cain, K. & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, 114(4), 647–662.
- Cain, K., Oakhill, J. & Bryant, Peter. (2004). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology*, 96(1), 31–42.
- Cain, K., Oakhill, J. & Lemmon, K. (2004). Individual Differences in the Inference of Word Meanings From Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology*, 96(4), 671–681.
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, 29(6), 850–859.
- Calvo, M. (2004). Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading. *Learning and Individual Differences*, 15(1), 53–65.

- Calvo, M. G. & Castillo, M. D. (2001). Bias in Predictive Inferences During Reading. *Discourse Processes*, 32(1), 43–71.
- Cantor, J., Engle, R. W. & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, 15(2), 229–246.
- Carretti, B., Borella, E., Cornoldi, C. & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, 19(2), 246–251.
- Catts, H. (2009). The narrow view of reading promotes a broad view of comprehension. *Language, Speech, and Hearing Services in Schools*, 40(2), 178–183.
- Catts, H., Compton, D., Tomblin, J. & Bridges, M. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, 104(1), 166–181.
- Catts, H. W., Adlof, S. M. & Weismer, S. E. (2006). Language Deficits in Poor Comprehenders: A Case for the Simple View of Reading. *Journal of Speech, Language & Hearing Research*, 49(2), 278–293.
- Catts, H. W., Fey, M. E., Tomblin, J. B. & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of speech, Language, and hearing Research*, 45(6), 1142–1157.
- Cerdan, R., Vidal-Abarca, E., Martinez, T., Gilabert, R. & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, 19(1), 13–27.
- Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.
- Chrysochoou, E., Bablekou, Z. & Tsigilis, N. (2011). Working memory contributions to reading comprehension components in middle childhood children. *The American journal of psychology*, 124(3), 275–89.
- Cornoldi, C., De Beni, R. & Pazzaglia, F. (2013). Profiles of Reading Comprehension Difficulties: An Analysis of of Single Cases. I C. Cornoldi & J. V. Oakhill (Red.), *Reading Comprehension Difficulties: Processes and Intervention* (s. 113–136). Routledge.
- Cromley, J. G. & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311–325.
- Crumpler, M. & McCarty, C. T. (2004). *Diagnostic reading analysis: reading booklet (forms A & B)*. London: Hodder Education.
- Cunningham, A. E. & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction*, 1(2), 137–149.
- Currie, N. K. & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, 137, 57–75.
- Cutting, L. E. & Scarborough, H. S. (2006). Prediction of Reading Comprehension: Relative Contributions of Word Recognition, Language Proficiency, and Other Cognitive Skills Can Depend on How Comprehension Is Measured. *Scientific Studies of Reading*, 10(3), 277–299.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, 19(4), 450–466.

- Daneman, M. & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422–433.
- Daugaard, H. T. & Elbro, C. (2013). *Inferensundervisning med grafiske modeller. En effektundersøgelse i specialundervisningen og på skolens 7. klassetrin*. Kbh.: Center for Læseforskning og Undervisningsministeriet (Ministeriet for Børn og Undervisning).
- Davey, B. (1990). Field dependence-independence and reading comprehension questions: Task and reader interactions. *Contemporary Educational Psychology*, 15(3), 241–250.
- De Beni, R., Palladino, P., Pazzaglia, F. & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology: Section A*, 51(2), 305–320.
- De Sousa, I. & Oakhill, J. (1996). Do levels of interest have an effect on children's comprehension monitoring performance? *British Journal of Educational Psychology*, 66(4), 471–482.
- Dixon, P., LeFevre, J.-A. & Twilley, L. C. (1988). Word Knowledge and Working Memory as Predictors of Reading Skill. *Journal of Educational Psychology December 1988*, 80(4), 465–472.
- Drane, M. J., Halpin, G. M., Halpin, W. G., vonEschenbach, J. F. & Worden, T. W. (1989). Relationships between reading proficiency and field dependence/field independence and sex. *Educational Research Quarterly*, 13(2), 2–10.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188.
- Ehrlich, M.-F. & Remond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. *British Journal of Developmental Psychology*, 15(3), 291–309.
- Ehrlich, M.-F., Remond, M. & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing*, 11(1), 29–63.
- Elbro, C. (2007). *Læsevanskeligheder* (1. udgave, 1. oplag). Kbh.: Gyldendal.
- Elbro, C. (2014). *Læsning og læseundervisning* (3. udgave). Kbh.: Hans Reitzels Forlag.
- Elbro, C. & Buch-Iversen, I. (2013). Activation of Background Knowledge for Inference Making: Effects on Reading Comprehension. *Scientific Studies of Reading*, 17(6), 435–452.
- 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. & Poulsen, M. (2015). *Hold i virkeligheden: Statistik og evidens i uddannelse*. Kbh.: Hans Reitzels Forlag.
- Elgart, D. B. (1978). Oral Reading, Silent Reading, and Listening Comprehension: a Comparative Study. *Journal of Literacy Research*, 10(2), 203–207.
- Elleman, A. M., Lindo, E. J., Morphy, P. & Compton, D. L. (2009). The Impact of Vocabulary Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis. *Journal of Research on Educational Effectiveness*, 2(1), 1–44.

- Elwér, Å., Gustafson, S., Byrne, B., Olson, R. K., Keenan, J. M. & Samuelsson, S. (2015). A retrospective longitudinal study of cognitive and language skills in poor reading comprehension. *Scandinavian Journal of Psychology*, 56(2), 157–166.
- Francis, D. J., Fletcher, J. M., Catts, H. W. & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. *Children's reading comprehension and assessment*, 369–394.
- Frith, U. (2003). *Autism : explaining the enigma* (2nd ed.). Malden MA: Blackwell Pub.
- Garcia, J. & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74–111.
- Garton, A. & Pratt, C. (1989). *Learning to be literate: the development of spoken and written language*. Oxford, UK; New York, NY, US: Blackwell.
- Gathercole, S. E. & Baddeley, A. D. (1993). *Working memory and language*. Hove, UK; Hillsdale, USA: L. Erlbaum Associates.
- Gellert, A. & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment*, 31(1), 16–28.
- Gernsbacher, M. A., Varner, K. R. & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 430–445.
- 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, New Jersey: Lawrence Earlbaum Associates, Publishers.
- Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371–395.
- Hansen, J. (1981). The Effects of Inference Training and Practice on Young Children's Reading Comprehension. *Reading Research Quarterly*, 16(3), 391–417.
- Happé, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, 3(6), 216–222.
- Happé, F., Briskman, J. & Frith, U. (2001). Exploring the cognitive phenotype of autism: Weak «central coherence» in parents and siblings of children with autism: I. Experimental tests. *Journal of Child Psychology and Psychiatry*, 42(3), 299–307.
- Happé, F. & Frith, U. (2006). The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25.
- Happé, F. G. & Booth, R. D. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. *Quarterly Journal of Experimental Psychology*, 61(1), 50–63.
- Harlaar, N., Hayiou-Thomas, M. E. & Plomin, R. (2005). Reading and General Cognitive Ability: A Multivariate Analysis of 7-Year-Old Twins. *Scientific Studies of Reading*, 9(3), 197–218.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. New York: The Guildford Press.
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2),

- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, Mass.: Harvard University Press.
- Jolliffe, T. & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition*, 71(2), 149–185.
- Keenan, J. M., Betjemann, R. S. & Olson, R. K. (2008). Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. *Scientific Studies of Reading*, 12(3), 281–300.
- Keenan, J. M., Hua, A. N., Meenan, C. E., Pennington, B. F., Willcutt, E. & Olson, R. K. (2014). Issues in Identifying Poor Comprehenders. *L'annee psychologique*, 114(4), 753–777.
- Keenan, J. M. & Meenan, C. E. (2014). Test differences in diagnosing reading comprehension deficits. *Journal of Learning Disabilities*, 47(2), 125–135.
- Kelly, L. P. & Barac-Cikoja, D. (2007). The Comprehension of Deaf Skilled Readers: The Roles of Word Recognition and Other Potentially Critical Aspects of Competence. I K. Cain & J. Oakhill (Red.), *Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective* (s. 244–280). New York: Guildford Press.
- Kelso, K., Fletcher, J. & Lee, P. (2007). Reading comprehension in children with specific language impairment: an examination of two subgroups. *International Journal of Language & Communication Disorders*, 42(1), 39–57.
- Kendeou, P., Bohn-Gettler, C., White, M. J. & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading*, 31(3), 259–272.
- Kieffer, M. J., Vukovic, R. K. & Berry, D. (2013). Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Reading Research Quarterly*, 48(4), 333–348.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. Cambridge; New York, NY, USA: Cambridge University Press.
- Kintsch, W. & Rawson, K. (2005). Comprehension. I M. Snowling & C. Hulme (Red.), *The Science of Reading: A Handbook* (s. 209–226). Malden MA: Blackwell Pub.
- Leach, J. M., Scarborough, H. S. & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, 95(2), 211.
- Li, M. & Kirby, J. R. (2014). Unexpected Poor Comprehenders Among Adolescent ESL Students. *Scientific Studies of Reading*, 18(2), 75–93.
- Lyon, G., Shaywitz, S. & Shaywitz, B. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1),
- Masson, M. & Miller, J. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology*, 75(2), 314–318.
- McGee, A. & Johnson, H. (2003). The Effect of Inference Training on Skilled and Less Skilled Comprehenders. *Educational Psychology*, 23(1), 49.
- McKenna, F. P. (1984). Measures of field dependence: cognitive style or cognitive ability? *Journal of Personality and Social Psychology*, 47(3), 593.
- McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3), 440–466.

- McMaster, K. L., Espin, C. A. & van den Broek, P. (2014). Making Connections: Linking Cognitive Psychology and Intervention Research to Improve Comprehension of Struggling Readers. *Learning Disabilities Research & Practice*, 29(1), 17–24.
- McNamara, D. S. & Magliano, J. (2009). Towards a comprehensive model of comprehension. I B. Ross (Red.), *The psychology of learning and motivation* (Bd. 51, s. 297–284). New York, NY, US: Elsevier Science.
- Medo, M. & Ryder, R. (1993). The effects of vocabulary instruction on readers' ability to make causal connections. *Reading Research and Instruction*, 33(2), 119–134.
- Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291.
- Meyer, B. J. F. & Freedle, R. O. (1984). Effects of Discourse Type on Recall. *American Educational Research Journal*, 21(1), 121–143.
- Miller, A. C., Keenan, J. M., Betjemann, R. S., Willcutt, E. G., Pennington, B. F. & Olson, R. K. (2013). Reading Comprehension in Children with ADHD: Cognitive Underpinnings of the Centrality Deficit. *Journal of Abnormal Child Psychology*, 41(3), 473–483.
- Murray, J. D., Klin, C. M. & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory and Language*, 32(4), 464–473.
- Muter, V., Hulme, C., Snowling, M. J. & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental psychology*, 40(5), 665.
- Nagy, W. E. (2007). Metalinguistic Awareness and the Vocabulary-Comprehension Connection. I R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Red.), *Vocabulary Acquisition: Implications for Reading Comprehension* (s. 52–77). Guilford Press.
- Nash, H. & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: a controlled evaluation of the definition and context methods. *International Journal of Language & Communication Disorders*, 41(3), 335–354.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Institute of Child Health and Human Development, National Institutes of Health.
- Nation, K. (2009). Reading comprehension and vocabulary. I R. K. Wagner, C. Schatschneider & C. Phythian-Sence (Red.), *Beyond decoding. The Behavioral and Biological Foundations of Reading Comprehension*. New York: The Guildford Pres.
- Nation, K., Adams, J., Bowyer-Crane, C. & Snowling, M. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, 73(2), 139–158.
- Nation, K., Clarke, P., Marshall, C. M. & Durand, M. (2004). Hidden Language Impairments in Children: Parallels Between Poor Reading Comprehension and Specific Language Impairment? *Journal of Speech Language and Hearing Research*, 47(1), 199.
- Nation, K., Clarke, P. & Snowling, M. (2002). General cognitive ability in children with reading comprehension difficulties. *British Journal of Educational Psychology*, 72, 549–560.

- Nation, K., Cocksey, J., Taylor, J. S. H. & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51(9), 1031–1039.
- Nation, K. & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67(3), 359–370.
- Nation, K. & Snowling, M. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*, 69(4), 996–1011.
- Nation, K. & Snowling, M. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, 70(1), B1–B13.
- Nation, K. & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics*, 21(02), 229–241.
- Neale, M. D. & Whetton, C. (1997). *Neale analysis of reading ability, revised*. Windsor: NFER-Nelson.
- Nielsen, I. & Petersen, D. K. (1993a). *Diavok, en diagnostisk læse- og stavetest*. AOF's Landsforbund.
- Nielsen, I. & Petersen, D. K. (1993b). Udvilling og afprøvning af et diagnostisk materiale for voksne dyslektikere. Specialeafhandling, Københavns Universitet
- Nielsen, J. B. (2008). Det mentale leksikon og testning af receptivt ordforråd, ændring af Peabody-testen. Specialeafhandling, Københavns Universitet.
- Norbury, C. & Nation, K. (2011). Understanding variability in reading comprehension in adolescents with autism spectrum disorders: Interactions with language status and decoding skill. *Scientific Studies of Reading*, 15(3), 191–210.
- Oakhill, J. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology*, 73(1), 13–20.
- Oakhill, J. (1984). Inferential and Memory Skills in Children's Comprehension of Stories. *British Journal of Educational Psychology*, 54(1), 31–39.
- Oakhill, J. & Cain, K. (2007). Issues of Causality in Children's Reading Comprehension. I D. S. McNamara (Red.), *Reading comprehension strategies theories, interventions, and technologies* (s. 47–71). Lawrence Erlbaum Associates, Taylor & Francis Group.
- Oakhill, J., Cain, K. & Elbro, C. (2015). *Læseforståelse: indsigt og undervisning*. Kbh.: Hans Reitzels Forlag.
- Oakhill, J., Cain, K., McCarthy, D. & Nightingale, Z. (2012). Making the link between vocabulary knowledge and comprehension skill. I A. Britt, S. Goldman, & J.-F. Rouet (Red.), *From words to reading for understanding*. New York: Routledge.
- Oakhill, J., Hartt, J. & Samols, D. (2005). Levels of Comprehension Monitoring and Working Memory in Good and Poor Comprehenders. *Reading and Writing*, 18(7-9), 657–686.
- Oakhill, J., & Cain, K. (2012). The Precursors of Reading Ability in Young Readers: Evidence From a Four-Year Longitudinal Study. *Scientific Studies of Reading*, 16(2), 91–121.
- Oakhill, J., Yuill, N. & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension. *International Electronic Journal of Elementary Education*, 4(1), 83–106.

- O'Brien, E. J., Shank, D. M., Myers, J. L. & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 410–420.
- Ouellette, G. & Beers, A. (2009). A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, 23(2), 189–208.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554–566.
- Pérez, A. I., Paolieri, D., Macizo, P. & Bajo, T. (2014). The role of working memory in inferential sentence comprehension. *Cognitive Processing*, 15(3), 405–413.
- Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357–383.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A. (1993). Why inferences might be restricted. *Discourse Processes*, 16(1-2), 181–192.
- Perfetti, C. A. (1994). Psycholinguistics and reading ability. I M. A. Gernsbacher (Red.), *Handbook of psycholinguistics* (s. 849–894). San Diego, CA, US: Academic Press.
- Perfetti, C. A., Landi, N. & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. I M. J. Snowling & C. Hulme (Red.), *The Science of Reading: A Handbook* (s. 227–247). Oxford, UK: Blackwell Publishing Ltd.
- Perfetti, C. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37.
- Perfetti, C., Yang, C.-L. & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, 22(3), 303–318.
- Pitts, M. M. & Thompson, B. (1984). Cognitive Styles as Mediating Variables in Inferential Comprehension. *Reading Research Quarterly*, 19(4), 426–435.
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S. & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841–865.
- Richter, T., Isberner, M.-B., Naumann, J. & Neeb, Y. (2013). Lexical Quality and Reading Comprehension in Primary School Children. *Scientific Studies of Reading*, 17(6), 415–434.
- Ricketts, J. (2011). Research review: Reading comprehension in developmental disorders of language and communication. *Journal of Child Psychology and Psychiatry*, 52(11), 1111–1123.
- Ricketts, J., Nation, K. & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11(3), 235–257.
- Ricketts, J., Sperring, R. & Nation, K. (2014). Educational attainment in poor comprehenders. *Frontiers in Psychiatry*, 5, 445.
- Rost, D. H. (2007). *Interpretation und Bewertung pädagogisch-psychologischer Studien: eine Einführung*. Weinheim: Beltz.
- Rowe, E. W., Miller, C., Ebenstein, L. A. & Thompson, D. F. (2012). Cognitive predictors of reading and math achievement among gifted referrals. *School Psychology Quarterly*, 27(3), 144–153.

- Rust, J., Golombok, S., Trickey, G. & Psychological Corporation. (1993). *WORD, Wechsler Objective Reading Dimensions manual*. London: Psychological Corporation.
- Rønberg, L. F. & Petersen, D. K. (2015a). How Specific are Specific Comprehension Difficulties? An Investigation of Poor Reading Comprehension in Nine-Year-Olds. *Scandinavian Journal of Educational Research*, online version.
- Rønberg, L. F. & Petersen, D. K. (2015b). It matters whether reading comprehension is conceptualised as rate or accuracy. *Journal of Research in Reading*, online version.
- Sanders, T. J. M. & Pander Maat, H. L. (1996). Cohesion and Coherence: Linguistic Approaches. I K. Brown (Red.), *Encyclopedia of Language and Linguistics* (Bd. 2, s. 591–595). Amsterdam: Elsevier.
- Sarfati, Y., Brunet, E. & Hardy-Baylé, M.-C. (2003). Comic-Strip Task: Attribution of Intentions to Others. Le Chesnay: Service de Psychiatrie Adulte, Ho[^]pital de Versailles.
- Sarfati, Y., Hardy-Baylé, M.-C., Besche, C. & Widlöcher, D. (1997). Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophrenia research*, 25(3), 199–209.
- Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V. & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and writing*, 13(1-2), 81–103.
- Semel, E., Wiig, E. H. & Secord, W. (2013). *CELF 4 Clinical Evaluation of Language Fundamentals. Fourth Edition. Vejledning. Dansk Version*. Pearson Assessment.
- Shears, C. & Chiarello, C. (2004). Knowledge-Based Inferences Are Not General. *Discourse Processes*, 38(1), 31–55.
- Shears, C., Hawkins, A., Varner, A., Lewis, L., Heatley, J. & Twachtmann, L. (2008). Knowledge-based inferences across the hemispheres: Domain makes a difference. *Neuropsychologia*, 46(10), 2563–2568.
- Shears, C., Miller, V., Ball, M., Hawkins, A., Griggs, J. & Varner, A. (2007). Cognitive Demand Differences in Causal Inferences: Characters' Plans are More Difficult to Comprehend Than Physical Causation. *Discourse Processes*, 43(3), 255–278.
- Simkin, Z. & Conti-Ramsden, G. (2006). Evidence of reading difficulty in subgroups of children with specific language impairment. *Child Language Teaching and Therapy*, 22(3), 315–331.
- Singer, M., Andruslak, P., Reisdorf, P. & Black, N. (1992). Individual differences in bridging
- Spear-Swerling, L. (2004). Fourth Graders' Performance on a State-mandated Assessment Involving Two Different Measures of Reading Comprehension. *Reading Psychology*, 25(2), 121–148.
- Spencer, M., Quinn, J. M. & Wagner, R. K. (2014). Specific Reading Comprehension Disability: Major Problem, Myth, or Misnomer? *Learning Disabilities Research & Practice*, 29(1), 3–9.
- Spooner, A. L. R., Baddeley, A. D. & Gathercole, S. E. (2004). Can reading accuracy and comprehension be separated in the Neale Analysis of Reading Ability? *The British Journal of Educational Psychology*, 74(2), 187–204.
- Stanovich, K. E. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. *Advances in Child Development and Behavior*, 24, 133–180.

- Sternberg, R. J. & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 38(8), 878–893.
- Stothard, S. E. & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing*, 4(3), 245–256.
- Swalm, J. E. (1972). A Comparison of Oral Reading, Silent Reading and Listening Comprehension. *Education*, 92(4), 111–115.
- Tannenbaum, K. R., Torgesen, J. K. & Wagner, R. K. (2006). Relationships between Word Knowledge and Reading Comprehension in Third-Grade Children. *Scientific Studies of Reading*, 10(4), 381–398.
- Tinajero, C. & Paramo, M. (1997). Field dependence-independence and academic achievement: A re-examination of their relationship. *Journal of Educational Psychology*, 67(2), 199–212.
- Tomesen, M. & Aarnoutse, C. (1998). Effects of an Instructional Programme for Deriving Word Meanings. *Educational Studies*, 24(1), 107–128.
- Tong, X., Deacon, S. H., Kirby, J. R., Cain, K. & Parrila, R. (2011). Morphological Awareness: A Key to Understanding Poor Reading Comprehension in English. *Journal of Educational Psychology*, 103(3), 523–534.
- Tranel, D., Anderson, S. W. & Benton, A. (1994). Development of the concept of “executive function” and its relationship to the frontal lobes. I F. Boller, H. Spinnler, & J. A. Hendler (Red.), *Handbook of neuropsychology* (Bd. 9, s. 125–148). Amsterdam: Elsevier.
- Vadasy, P. F., Sanders, E. A. & Herrera, B. L. (2014). Efficacy of Rich Vocabulary Instruction in Fourth and Fifth Grade Classrooms. *Journal of Research on Educational Effectiveness*, *accepted manuscript*.
- Van den Broek, P. (1994). Comprehension and memory of narrative texts: Inference and coherence. I M. A. Gernsbacher (Red.), *Handbook of psycholinguistics* (s. 423–445). Academic Press.
- Van den Broek, P. & Kremer, K., E. (2000). The Mind in Action. I B. M. Taylor, M. F. Graves, & P. van den Broek (Red.), *Reading for meaning: fostering comprehension in the middle grades* (s. 1–31).
- Van den Broek, P., Ridsen, K. & Husebye-Hartmann, E. (1995). The Role of Readers’ Standards for Coherence in the Generation of Inferences During Reading. I R. F. Lorch & E. J. O’Brien (Red.), *Sources of coherence in reading* (s. 353–373). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Van Dyke, J. A. & Shankweiler, D. P. (2013). From Verbal Efficiency Theory to Lexical Quality: The Role of Memory Processes in Reading Comprehension. I M. A. Britt, S. R. Goldman, & J.-F. Rouet (Red.), *Reading – from Words to Multiple Texts* (115–131). New York: Routledge.
- Van Silfhout, G., Evers-Vermeul, J. & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, 52(1), 47–76.
- Van Wingerden, E., Segers, E., van Balkom, H. & Verhoeven, L. (2014). Cognitive and linguistic predictors of reading comprehension in children with intellectual disabilities. *Research in Developmental Disabilities*, 35(11), 3139–3147.

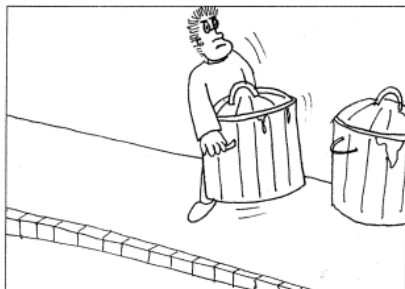
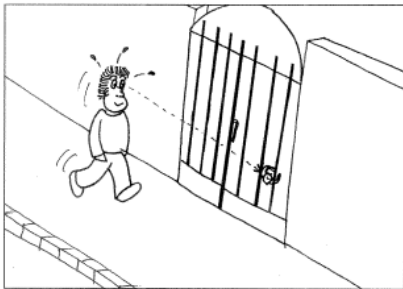
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1), 2–40.
- Verhoeven, L. (2013). Reading Acquisition in a Transparent Orthography. *Reading—from Words to Multiple Texts*, 34.
- Vidal-Abarca, E., Salmeron, L. & Mana, A. (2011). Individual differences in task-oriented reading. I J. Magliano, M.T. McCrudden & G. Shraw (Red.), *Text relevance and learning from text* (267–293). Charlotte NC: Information Age Pub.
- Wagner, R. K., Muse, A. E. & Tannenbaum, K. R. (2007). Promising avenues for better understanding implications of vocabulary development for reading comprehension. I R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Red.), *Vocabulary acquisition: Implications for reading comprehension* (s. 276–291). New York, NY, US: Guilford Press.
- Wahlberg, T. & Magliano, J. (2004). The Ability of High Function Individuals With Autism to Comprehend Written Discourse. *Discourse Processes*, 38(1), 119–144.
- Walczyk, J. (1993). Are general resource notions still viable in reading research? *Journal of Educational Psychology*, 85(1), 127–135.
- Walczyk, J. (1995). Testing a compensatory-encoding model. *Reading Research Quarterly*, 30(3), 396–408.
- White, S. & Saldaña, D. (2011). Performance of children with autism on the Embedded Figures Test: A closer look at a popular task. *Journal of Autism and Developmental Disorders*, 41(11), 1565–1572.
- Williams, J. P., Hall, K. M. & Lauer, K. D. (2004). Teaching Expository Text Structure to Young At-Risk Learners: Building the Basics of Comprehension Instruction. *Exceptionality*, 12(3), 129–144.
- Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, k. B. & deCani, J. S. (2005). Expository Text Comprehension in the Primary Grade Classroom, 97(4), 538–550.
- Williams, J. P., Nubla-Kung, A. M., Pollini, S., Stafford, K. B., Garcia, A. & Snyder, A. E. (2007). Teaching Cause—Effect Text Structure Through Social Studies Content to At-Risk Second Graders. *Journal of Learning Disabilities*, 40(2), 111–120.
- Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M. & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology*, 101(1), 1–20.
- Witkin, H. A., Moore, C. A., Goodenough, D. R. & Cox, P. W. (1977). Field-Dependent and Field-Independent Cognitive Styles and Their Educational Implications. *Review of Educational Research*, 47(1), 1–64.
- Witkin, H., Oltman, P., Raskin, E. & Karp, S. (1971, 2002). *Group Embedded Figures Test Sampler Set Manual and Sample Figures and Scoring*. Mindgarden, Inc.
- Yuill, N. & Joscelyne, T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology*, 80(2), 152–158.
- Yuill, N. & Oakhill, J. (1988). Effects of Inference Awareness Training on Poor Reading Comprehension. *Applied Cognitive Psychology*, 2(1), 33–45.

- Yuill, N. & Oakhill, J. V. (1991). *Children's problems in text comprehension : an experimental investigation*. Cambridge; New York: Cambridge University Press.
- Yussen, S. R., Rembold, K. L. & Mazor, A. (1989). Identifying main ideas in picture stories and written narratives. *Journal of Applied Developmental Psychology*, 10(3), 313–335.

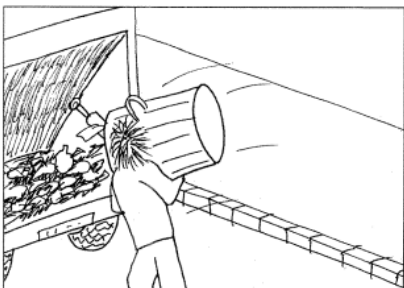
Bilag

Bilag 1: Eksempler på opgaver i <i>Test af nonverbale inferenser fra narrativer</i>	s. 151
Bilag 2: Eksempler på opgaver i <i>Test af verbale inferenser fra narrativer</i>	s. 152
Bilag 3: Eksempler på opgaver i <i>Test af inferenser fra fagtekster</i>	s. 153

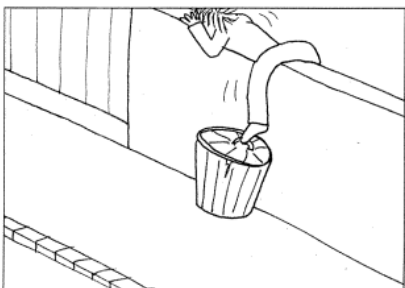
Bilag 1



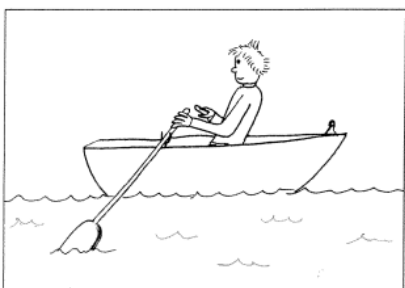
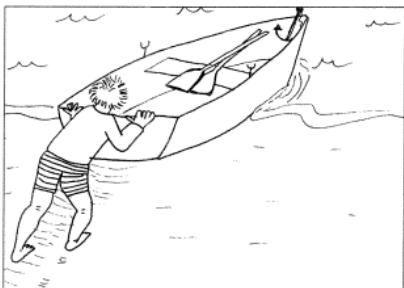
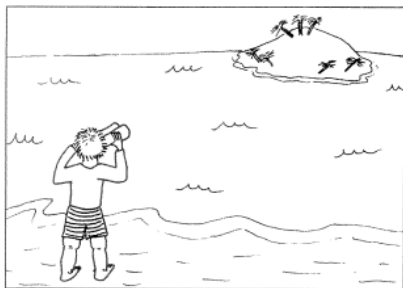
A



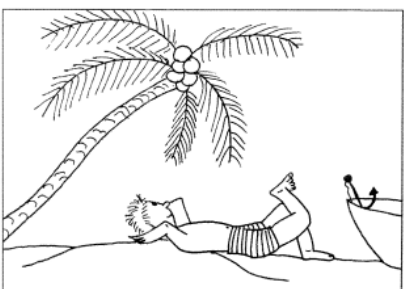
B



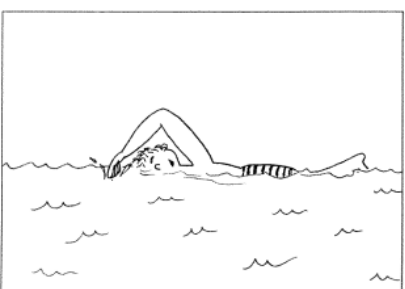
C



A



B



C

Bilag 2

Det drysser med brødkrummer ud over det hele, da han tager skærebrættet frem. Han bander for sig selv. Han går over til det høje skab.

- a) Han tager en ren trøje frem
- b) Han tager en kost frem
- c) Han tager havregryn frem

Hun kigger på den tørre blomst. Hun ser, at solen står højt på himlen. Hun mærker varmen i lejigheden.

- a) Hun drikker et glas vand
- b) Hun henter sine solbriller
- c) Hun henter en vandkande

Han spiser sin aftensmad og vasker op. Han sætter vand over. Han kigger ud af vinduet og nynner en lille melodi.

- a) Han går i seng
- b) Han finder et tebrev frem
- c) Han finder sin tandbørste frem

Hun har svært ved at holde sig vågen. Hun skynder sig hjem. Hun tager et lagen frem fra en skuffe.

- a) Hun reder sin seng
- b) Hun klæder sig ud som spøgelse
- c) Hun hænger lagenet op i vinduet

Bilag 3

Kroppens muskler sidder fast på kroppens knogler, som vi også kalder skelettet. Skelettet hos et foster består først af brusk, som er blødt og bøjeligt, men efterhånden som barnet vokser, dannes der mineralkrystaller af forskellig slags i skelettet.

Hvad kommer der til at ske, når der dannes mineralkrystaller i skelettet?

Signaleret inferens (konnektiv)

I perioden op til 1611 fik Christian den fjerde lavet nogle jordvolde rundt om København, og han byggede fæstninger ved landets grænser. Her holdt man sig parat til at kæmpe. I samme periode byggede Christian den fjerde også en ny krigshavn i København. I anlægget lastede man våben og fødevarer. Alt blev gjort i skjul.

Hvad foregik der i den nye krigshavn?

Signaleret inferens (henvisende substantiv med bestemthedsendelse)

I de amerikanske regnskove danner træernes toppe et tæt grønt tæppe. Oppe i trætoppene lever der en masse dyr, fx aber, fugle og tilmed dovendyr. I skovbunden er der ikke mange planter at finde for skovens dyr.

Hvorfor er der ikke mange planter at finde i skovbunden?

Ikke-signaleret inferens

I de seneste år er ny teknologi som GPS blevet taget i brug i landbruget. GPS medfører, at markerne kan dyrkes mere præcist og effektivt. Mejetærskere og traktorer, som i dag er helt nødvendige for landmanden, blev taget i brug i løbet af 1800- og 1900-tallet. Mejetærskere og traktorer medførte, at mange folk blev arbejdsløse og flyttede fra landet ind til byerne.

Hvad medfører GPS i landbruget?

Bogstavelig

Hvad medførte mejetærskere og traktorer?

Bogstavelig