

# Betydningen av inferens for leseforståelse

## Effekter av inferenstrening

av  
Ida Buch-Iversen

Avhandling for graden  
Philosophiae Doctor  
(PhD)



Nasjonalt senter for leseopplæring og leseforskning  
Humanistisk fakultet  
Universitetet i Stavanger  
2010

University of Stavanger  
P.O. Box 8002 Ullandhaug  
N-4068 Stavanger  
NORWAY  
[www.uis.no](http://www.uis.no)

©2010 Ida Buch-Iversen

ISBN: 978-82-7644-403-2  
ISSN: 1890-1387

---

## Forord

Denne avhandlingen er resultat av drøyt fire års arbeid som stipendiat ved Nasjonalt senter for leseopplæring og leseforskning, Lesesenteret, ved Universitetet i Stavanger. Lesesenteret har oppgaver både ved universitetet og utdanningsdirektoratet som spesialpedagogisk kompetansesenter og nasjonalt senter for leseopplæring. Stipendet har vært knyttet til virksomheten som nasjonalt senter for leseopplæring (§1-4).

Det er mange som har hjulpet meg i arbeidet med denne avhandlingen. Først vil jeg takke min hovedveileder, Carsten Elbro, som har fulgt meg gjennom hele prosessen. Han var med på å stake ut kursen i oppstartsfasen, har kommet med både ideer og realitetsorienteringer underveis og har hjulpet meg med å samle trådene i innspurten. Jeg er også takknemlig for hjelpen jeg har fått av biveileder Liv Engen. I tillegg har jeg også fått innspill fra mine kolleger Kjersti Lundetræ, Anne Elisabeth Dahle, Charlotte Begnum, Oddny Solheim og Nina N. Gabrielsen i stipendiatgruppen, og gruppens mentor, Knud Knudsen. Halvveis i stipendiatperioden fikk jeg også gode ideer til effektstudien fra Kate Cain da jeg hadde et opphold ved University of Lancaster. Dessuten vil jeg rette en takk til Pekka Niemi som med stort engasjement for mitt arbeid ga meg gode råd som opponert på 90 %-seminar.

En stor takk til tålmodige elever, engasjerte lærere og positive rektorer og inspektører ved skolene som deltok i forundersøkelsen og effektstudien. Universitetsbiblioteket i Stavanger fortjener også takk for veiledning i EndNote og for at de har skaffet meg nødvendig litteratur.

Lesesenteret har vært en god arbeidsplass med engasjerte kolleger som jeg setter stor pris på. Senterlederne Ingvold Austad og Ragnar Gees Solheim skal ha takk for å ha lagt til rette for gode arbeidsforhold gjennom stipendperioden. En særlig takk til min gode venn og kollega Charlotte Begnum, som jeg har hatt utallige diskusjoner med både over SPSS output'er og over strikkesøyet. Også en stor takk for støtte og oppmuntring fra Victor van Daal og Anne Mangen.

Til slutt en varm takk til Hans, min ektemann og bestevenn, som ble med meg over fjellet fra Oslo til Stavanger da jeg fikk stipendiatstillingen. Jeg setter stor pris på all den tålmodighet og forståelse jeg har fått på hjemmefronten.

Stavanger, november 2009  
Ida Buch-Iversen



---

## Sammendrag

Denne avhandlingen fokuserer på det å gjøre inferenser under lesing. Det overordnede målet med avhandlingen var å undersøke effekter av inferenstrening blant skolebarn. Avhandlingen hadde både en teoretisk og en pedagogisk motivasjon. Den pedagogiske motivasjonen var å finne en måte å arbeide med inferens på som kan anvendes i en norsk skolehverdag. Teoretisk var motivasjonen å påvise en årsakssammenheng mellom inferens og leseforståelse. Effekter av inferenstrening ble undersøkt på to måter. For det første ved å sammenfatte tidligere effektstudier av inferenstrening i en metaanalyse, og for det andre ved å gjennomføre en effektstudie av inferenstrening blant norske skolebarn.

Inferenser er nødvendig for å få en sammenhengende forståelse av tekst. Fordi inferensbegrepet defineres og kategoriseres på mange måter i faglitteraturen, ble det i avhandlingen laget en egen taksonomi av inferenstyper. Særlig fokuseres det på to typer inferenser som begge er nødvendige for å oppnå forståelse, brobyggende og utfyllende inferenser. Brobyggende inferenser er nødvendige for å få en sammenheng mellom ord og setninger på lokalt nivå i teksten. Utfyllende inferenser krever at leseren stiller opp et scenario eller en situasjonsmodell av innholdet i teksten.

Effekter av inferenstrening blant skolebarn i tidligere studier ble undersøkt gjennom en metaanalyse. Metaanalysen inkluderte ti studier som alle var gjennomført blant engelskspråklige barn. I alt 18 eksperimentgrupper inngikk i de ti studiene. Effektstørrelser, Cohen's  $d$ , og vektete gjennomsnittlige effektstørrelser ble beregnet på bakgrunn av opplysninger i studiene. Metaanalysen viste sterke effekter av trening både for inferens og leseforståelse målt umiddelbart etter avsluttet inferenstrening. Kun tre studier undersøkte effekter av inferenstreningen etter avsluttet trening. Resultatene viste at effekten avtok etter avsluttet trening. Studiene i utvalget ble kategorisert i forhold til hvilke inferenstyper som ble trent, resultatene viste at effektene var sterkest for studier som trente brobyggende inferens. Organisatoriske trekk ved inferenstreningen, for eksempel hvem som gjennomførte treningen, og hvor store grupper treningen ble gjennomført i, hadde lite betydning for effekt. Metaanalysen viste videre at inferenstreningen i de 18 eksperimentgruppene besto av mange av de samme komponentene, det var derfor ikke hensiktsmessig å beregne effekter for de enkelte komponenter. Tvert om kan denne overlappingen i komponenter betraktes som at det var enighet om hvordan inferens kan trenes.

---

På bakgrunn av metaanalysen ble det utviklet materiell til inferenstrening blant norske elever på sjette trinn. Effekter av inferenstreningen ble undersøkt i en effektstudie med pre- postdesign med kontrollgruppe og utsatt posttest. I tillegg til effektvariablene inferens og leseforståelse ble mål på ordavkodning, ordforråd, nonverbale ferdigheter, motivasjon og matematikk inkludert som kontrollvariabler. Utvalget i effektstudien besto av 236 elever, hvorav 151 elever deltok i eksperimentgruppe og 85 elever i kontrollgruppe. Inferenstreningen som ble prøvd ut gjennom effektstudien, var tilpasset en norsk skolekontekst basert på tre av de fire komponentene i inferenstrening som ble identifisert gjennom metaanalysen. I inferenstreningen ble det brukt tekster og spørsmål som krevde at elevene gjorde inferenser, og grafiske modeller som hadde til hensikt å illustrere sammenhengen mellom opplysninger i teksten. Inferenstreningen ble administrert av lærere og besto av åtte økter som hver varte i ca. 30 minutter. Resultatene viste at eksperimentgruppen hadde signifikant større framgang i inferens og leseforståelse fra pre- til posttest sammenlignet med kontrollgruppen. Effektene av inferenstreningen var sterke både med hensyn til inferens og leseforståelse. Effektene for inferens vedvarte på den utsatte posttesten, men det var ikke ytterligere framgang hos eksperimentgruppen. For leseforståelse var det ikke signifikant forskjell mellom gruppene over tid. For item i målet på leseforståelse som krevde tolkning og refleksjon, vedvarte imidlertid effekten, men det var ingen ytterligere framgang. Det var ingen forskjell i hvordan elever med svake, middels og gode ferdigheter i inferens og leseforståelse målt ved pretest profitterte på treningen, slik effekten var målt. Det var heller ingen forskjell i hvordan jenter og gutter profitterte på inferenstreningen, verken med hensyn til inferens eller leseforståelse. Effekten av inferenstreningen var i liten grad betinget av ordforråd, ordavkodning og nonverbale ferdigheter, slik disse variablene var målt.

I tråd med avhandlingens pedagogiske motivasjon viste effektstudien at inferenser kan fremmes gjennom inferenstrening via grafiske modeller. Særlig verdifullt var det at treningen hadde overføringseffekt på leseforståelse. I forhold til den teoretiske motivasjonen viste både metaanalysen og effektstudien at trening i inferens førte til framgang i leseforståelse, noe som innebærer at inferens er årsak til, ikke et produkt av, leseforståelse.

---

## Innhold

Forord .....	I
Sammendrag .....	III
<b>1. Innledning.....</b>	<b>1</b>
1.1 Valg av tema .....	1
1.2 Avhandlingens oppbygning .....	4
<b>2. Leseforståelse .....</b>	<b>7</b>
2.1 Hva innebærer det å kunne lese med forståelse? .....	7
2.2 Ferdigheter av betydning for leseforståelse .....	9
2.2.1 Ferdigheter som årsak til eller konsekvens av leseforståelse.....	9
2.2.2 Ferdigheter på ordnivå .....	11
2.2.3 Bevissthet om tekststruktur.....	14
2.2.4 Monitorering .....	16
2.2.5 Inferens .....	17
2.2.6 Høyere standarder for sammenheng .....	19
2.3 Betingelser for leseforståelse .....	19
2.3.1 Verbale og nonverbale ferdigheter.....	19
2.3.2 Arbeidsminne.....	21
2.3.3 Leserens kunnskap om tekstens tema .....	22
2.3.4 Motivasjon .....	24
2.3.5 Sosiokulturell kontekst.....	25
2.2.6 Teksteksponering .....	26
<b>3. Typer av inferens .....</b>	<b>29</b>
3.1 Teoretiske perspektiver på inferens i lesing.....	29
3.2 Kategorisering av inferens .....	31
<b>4. Argument og motivasjon for effektstudier av inferenstrening .....</b>	<b>37</b>
4.1 Måling av effekt av inferenstrening.....	38
4.2 Eksperimentelle design .....	39
4.2.1 Validitetsmessige problemstillinger.....	40
4.2.1.1 Statistisk validitet.....	40
4.2.1.2 Indre validitet.....	42
4.2.1.3 Begrepsvaliditet .....	43
4.2.1.4 Ytre validitet .....	44
4.2.1.5 Hvordan imøtekomme validitetsmessige problemstillinger?..	44
4.3 Etske problemstillinger ved pedagogiske effektstudier .....	45
4.3.1 Etikk og bruk av kontrollgruppe .....	46
4.3.2 Informert samtykke.....	46
4.3.3 Hvordan imøtekomme etiske problemstillinger? .....	47
4.4 Oppsummering: Verdien av effektstudier av inferenstrening .....	48
<b>5. Metaanalyse av inferenstrening i tidligere effektstudier .....</b>	<b>51</b>
5.1 Mål og hypoteser .....	51

---

5.2 Utvalg .....	52
5.2.1 Beskrivelse av studiene i utvalget.....	53
5.2.1.1 Deltakere .....	53
5.2.1.2 Kategorier av inferens.....	55
5.2.1.3 Grupper .....	57
5.2.1.4 Komponenter i inferenstreningen.....	64
5.2.1.5 Organisering av inferenstreningen .....	66
5.3 Metode .....	67
5.3.1 Måling av effekt av inferenstreningen .....	67
5.3.2 Beregning av effektstørrelser .....	70
5.4 Resultater .....	72
5.4.1 Primær-, primær overført og generalisert effekt .....	72
5.4.2 Effekter knyttet til kjennetegn ved studiene .....	74
5.4.2.1 Effekter for inferenstyper.....	74
5.4.2.2 Effekter for komponenter i inferenstreningen.....	75
5.4.2.3 Effekter knyttet til hvem som gjennomførte infernstreningen.....	75
5.4.2.4 Effekter knyttet til gruppestørrelse og treningens varighet.....	75
5.4.2.5 Effekter knyttet til deltakernes leseferdigheter .....	76
5.4.3 Langtidseffekter .....	76
5.5 Diskusjon .....	77
5.5.1 Effekter av inferenstrening på fire nivåer .....	77
5.5.2 Effekter for studier med ulike kjennetegn.....	78
5.5.3 Begrensinger ved metaanalysen.....	80
5.5.4 Konklusjon.....	81
<b>6. Betydningen av situasjonsmodeller for inferens – en forundersøkelse</b>	<b>85</b>
6.1 Bakgrunn.....	85
6.2 Mål og problemstillinger .....	85
6.3 Metode .....	87
6.3.1 Design .....	87
6.3.2 Deltakere .....	87
6.3.3 Måleinstrumenter .....	88
6.3.3.1 Screening prøve .....	88
6.3.3.2 Ordforråd .....	89
6.3.3.3 Inferenstekster.....	90
6.3.4 Gjennomføring.....	92
6.4 Resultater .....	93
6.4.1 Beskrivende analyser .....	93
6.4.1.1 Deskriptive analyser.....	93
6.4.1.2 Korrelasjoner mellom variablene.....	96
6.4.2 Effekter av eksperimentbetingelser.....	97



---

6.4.2.1	Generell effekt av eksperimentbetingelsen på inferensspørsmålene?.....	98
6.4.2.2	Spesifikk effekt av instruksjonsspørsmålene for skåre på inferensspørsmålene?.....	99
6.4.2.3	Effekt av eksperimentbetingelsen på faktaspørsmålene?.....	100
6.5	Diskusjon .....	100
6.5.1	Problemstillinger og resultater .....	100
6.5.2	Erfaringer fra forundersøkelsen .....	103
<b>7.</b>	<b>Effekter av inferenstrening blant norske sjetteklassinger .....</b>	<b>105</b>
7.1	Mål og problemstillinger .....	105
7.2	Undervisning i inferens.....	107
7.2.1	Arbeidsform og rammer for inferenstreningen .....	107
7.2.2	Materiell.....	111
7.2.2.1	”Hva forteller setningene” .....	111
7.2.2.2	Inferenstekster.....	112
7.2.3	Instruksjon til lærerne i eksperimentgruppen.....	116
7.3	Metode .....	117
7.3.1	Design .....	117
7.3.2	Deltakere.....	117
7.3.3	Instrumenter .....	120
7.3.3.1	Mål på effekt .....	120
	Inferensprøve .....	120
	Prøve i leseforståelse .....	122
7.3.3.2	Kontrollvariabler.....	123
	Spørreskjema om motivasjon.....	123
	Ordforrådsprøve.....	124
	Ordkjedeprøve .....	125
	Raven .....	125
	Matematikkprøve – M5 .....	125
7.3.3.3	Mål på implementering .....	126
	Observasjonsskjema.....	126
	Loggskjema.....	126
7.3.4	Gjennomføring.....	126
7.3.4.1	Pretest.....	127
7.3.4.2	Inferenstrening .....	128
7.3.4.3	Posttest .....	128
7.3.4.4	Utsatt posttest.....	128
7.4	Resultater .....	128
7.4.1	Vurdering av gjennomføring av inferenstreningen .....	128
7.4.1.1	Klasseromsobservasjoner.....	129
7.4.1.2	Loggskjema.....	130

---

7.4.2 Innledende analyser .....	131
7.4.2.1 Deskriptiv statistikk og sammenligning av gruppene .....	131
7.4.2.2 Korrelasjoner mellom effektvariablene og kontrollvariablene ved pretest, samt språk, kjønn og alder.....	139
7.4.2.3 Kontroll av Hawthorneeffekt .....	140
7.4.2.4 Sammenligning av gutter og jenter .....	141
7.4.2.5 Sammenligning av enspråklige og tospråklige elever.....	143
7.4.3 Effekter av inferenstreningen.....	144
7.4.3.1 Måling av effekt.....	144
7.4.3.2 Forskjeller mellom eksperimentgruppe og kontrollgruppe... Sammenheng mellom endring på faktaspørsmål og tolkningsspørsmål .....	146 148
Sammenheng mellom endring på fiksjonstekster og nonfiksjonstekster .....	152
7.4.3.3 Endring i eksperimentklasser og kontrollklasser .....	154
7.4.3.4 Sammenhenger mellom klasseopplysninger og eksperimentklassenes endring.....	156
7.4.4 I hvilken grad forklarer kontrollvariablene endring? .....	158
7.4.4.1 Effekt av inferenstrening kontrollert for ordforråd, ordavkodning og nonverbale ferdigheter .....	158
7.4.4.2 Sammenhenger mellom endring og ordforråd, ordavkodning og nonverbale ferdigheter .....	161
7.4.5 Endring hos ulike elevgrupper .....	161
7.4.5.1 Forskjeller mellom elever på ulike ferdighetsnivå.....	161
7.4.5.2 Forskjeller mellom gutter og jenter.....	162
7.4.5.3 Forskjeller mellom enspråklige og tospråklige elever .....	166
7.4.6 Langtidseffekter av inferenstreningen?.....	168
7.5 Diskusjon .....	171
7.5.1 Resultater og problemstillinger .....	171
7.5.1.1 Effekter av inferenstreningen på inferens og leseforståelse..	171
7.5.1.2 Langtidseffekter .....	174
7.5.1.3 Betingelser for effekt .....	175
Kontrollvariabler.....	175
Ulike grupper elever .....	175
Forhold i eksperimentklassene.....	177
7.5.2 Muligheter for generalisering .....	178
7.5.2.1 Forutsetninger knyttet til gruppene .....	178
7.5.2.2 Forutsetninger knyttet til måleinstrumenter .....	180
7.5.2.3 Begrensninger i og muligheter for generalisering.....	183
7.5.3 Konklusjon.....	184
<b>8. Oppsummering og konklusjoner .....</b>	<b>187</b>

---

Referanser .....	191
Liste over tabeller .....	203
Liste over figurer .....	205
Oversikt over appendiks .....	207

## **1. Innledning**

### **1.1 Valg av tema**

Det å kunne lese er blitt en svært viktig kompetanse for å delta i utdanning, arbeid og generelt i samfunnet. Den internasjonale undersøkelsen Adult Literacy and Life Skills (ALL) som ble gjennomført i 2003, illustrerer dette. Undersøkelsen kartla voksnes kompetanse i lesing, tallforståelse og problemløsning. Resultatene fra de norske dataene viste at av de deltakerne som skåret lavest på de ovennevnte kompetanseområdene, var nesten 10 prosent uten jobb (Bratsberg, Raaum, Hægeland, Adult Literacy and Life Skills, & Nasjonalt senter for leseopplæring og leseforskning, 2006). Til sammenligning var arbeidsledigheten i befolkningen for øvrig på 4,5 prosent samme år (Statistisk sentralbyrå, 2004).

Det siste tiåret har Norge deltatt i flere internasjonale leseundersøkelser blant skoleelever, som for eksempel Programme for International Student Assessment (PISA) i 2000 (Programme for International Student Assessment, 2001) og 2003 (Programme for International Student Assessment, 2004) og Progress in International Reading Literacy Study (PIRLS) i 2001 (Mullis, Martin, Gonzalez, & Kennedy, 2003) og 2006 (Mullis, Martin, Kennedy, & Foy, 2007). PISA undersøkte blant annet leseferdigheter hos femtenåringene, mens PIRLS undersøkte leseferdigheter blant elever på fjerde og femte trinn.

De norske femtenåringene som deltok i PISA undersøkelsene i 2001 og 2003 (Kjærnsli, Lie, Olsen, Roe, & Turmo, 2004; Lie, Kjærnsli, Roe, & Turmo, 2001), viste gjennomsnittlige prestasjoner i lesing. Forfatterne av de norske rapportene uttrykker misnøye med de norske elevenes middelmådige resultater, tatt i betraktning at Norge er et av verdens rikeste land, og at knapt noe annet land bruker mer penger på skole og utdanning. Særlig urovekkende er det at hele 17 % av de norske femtenåringene ifølge forfatterne av den internasjonale PISA rapporten står i fare for å ikke dra nytte av utdanning på grunn av deres lave leseferdigheter (Programme for International Student Assessment, 2001, s. 45-48).

PIRLS-undersøkelsene viste tilsvarende resultater for norske fjerde og femte trinns elever (Senter for leseforskning & Progress in International Reading Literacy Study, 2003; Solheim et al., 2007). I PIRLS-undersøkelsene skåret de norske elevene like under det internasjonale gjennomsnittet, og klart under øvrige nordiske land på leseferdigheter både i 2001 og 2006. Forfatterne av de norske PIRLS-rapportene påpeker at noe av forklaringen på at norske elever

ikke gjør det bedre, kan ligge i at de norske elevene hadde hatt færre år med formell leseopplæring enn elevene i mange av de øvrige landene som var med i PIRLS. Likevel presterer de norske elevene lavere enn elever i de andre nordiske landene som er like gamle og har like mange års formell leseopplæring. PIRLS-undersøkelsene viste også at de norske lærerne, sammenlignet med lærere i øvrige land, i mindre grad inkluderte aktiviteter i undervisningen for å utvikle elevenes leseferdigheter, som for eksempel å generalisere og å trekke slutninger, å skildre stil og struktur i teksten og å trekke ut hovedpoenger i teksten. Dette er eksempler på aktiviteter som særlig utfordrer elevene til en dypere forståelse av teksten, ut over informasjonsuthenting. Med andre ord vektlegger de norske lærerne i mindre grad aktiviteter som fremmer dypere leseforståelse enn sine kolleger i andre land.

Resultater i de ovennevnte undersøkelsene førte til debatt om norske elevers leseferdigheter både i politikken og i pedagogiske fagkretser. PIRLS-undersøkelsen var ifølge Kavli (2008) viktig i arbeidet med å stake ut kursen for den nye læreplanen på begynnelsen av 2000-tallet, blant annet med hensyn til retningslinjer for leseundervisning og fokus på lesing gjennom hele skoleløpet. Da den nye læreplanen, Kunnskapsløftet, kom i 2006 (Kunnskapsdepartementet & Utdanningsdirektoratet, 2006), ble lesing definert som en av fem grunnleggende ferdigheter som inngikk i alle fag. Kunnskapsløftet inkluderte også flere kompetansemål som spesielt vektla leseforståelse. Etter fjerde trinn er det eksempelvis forventet at elevene skal kunne "lese barnelitteratur og fagtekster for barn med flyt, sammenheng og forståelse for innholdet og beskrive egne litteraturvalg" (s. 45). Videre er et av kompetansemålene etter sjuende trinn er at elevene skal kunne "strukturere tekst etter tidsrekkefølge og tema og skape sammenheng mellom setninger og avsnitt"(s. 46). Selv om leseforståelse har en sentral plass i målene for opplæring i Kunnskapsløftet, gir læreplanen imidlertid få svar på hvordan elevene kan tilegne seg en god leseforståelse.

Hvis vi vender oss til forskningen på leseundervisning, finner vi en ubalanse mellom forskning på avkoding og leseforståelse. Mens det er en overvekt av studier om avkoding, er studier av leseforståelse stadig en minoritet. Norske elevers resultater på internasjonale leseundersøkelser, ambisjoner i Kunnskapsløftet og ubalansen i forskningen på leseundervisning taler for at flere studier av undervisning i leseforståelse er på sin plass. Denne avhandlingen setter fokus nettopp på leseforståelse. Avhandlingen undersøker effekter av undervisning i leseforståelse, mer spesifikt, effekter av å trene inferens både på inferens spesielt og på leseforståelse generelt.

Elbro (2007) definerer lesing som det å gjenskape et forestillingsinnhold på bakgrunn av identifikasjon av ord i teksten og kunnskaper om den verden som teksten beskriver (s. 40). Denne definisjonen sier at leseforståelse innebærer gjenskaping av mening, meningen ligger allerede i teksten, og leseren må derfor rekonstruere meningen. Med andre ord må leserens oppfatning av teksten være i tråd med forfatterens idé og hensikt. Å rekonstruere mening innebærer at leseren danner seg indre bilder av situasjonen eller innholdet som beskrives i teksten. Dette krever at leseren har kunnskap og erfaringer om tekstens tema. Siden ingen lesere har eksakt den samme kunnskap og erfaring, vil forståelsen av en tekst variere noe fra leser til leser. Leserens kunnskaper og erfaringer har også konsekvenser for hvor dyp eller omfattende forståelse han eller hun får av teksten. Snow og Sweet (2003) understreker to nivåer, eller komponenter, i leseforståelse. For det første det å trekke ut informasjon, og for det andre å konstruere mening på bakgrunn av tekst. Dersom leseren mangler kunnskaper om tekstens tema, blir det vanskeligere å rekonstruere meningen og å skape indre bilder av innholdet. Mens informasjonsuthentning innebærer forholdsvis overfladisk forståelse, krever rekonstruksjon av mening og dermed dannelse av indre bilder en dypere forståelse.

Teksten gir oss hint om hva slags meningsinnhold vi skal rekonstruere, men disse hintene er ufullstendige. Dette er fordi nesten ingen tekster er fullstendig eksplisitte, tvert om er de fleste tekster ganske ufullstendige. Når sammenhengen mellom ord, setninger og avsnitt er implisitt, kreves det at leseren trekker slutninger om innholdet. Denne typen slutninger kalles inferens. Vi kan betrakte inferens som de "hullene" som leseren må fylle for at rekonstruksjonen av mening skal få sammenheng. De første linjene i fortellingen "Hold Banen" som er gjengitt, under illustrerer dette.

*"- Tar du sekken min?"*

*Per kastet et raskt blikk bort på Gøran. Han nikket kort, slang blyanten inn i pennalet og dro igjen glidelåsen. Begge to feide mattebøkene lynraskt ned i sekken og stilte seg ved pulten.*

*De måtte rekke det i dag også. De måtte rekke å holde banen før 6. klasse kom. Gøran stod klar til sprang."*

(Kristoffersen, 2006, s. 93)

En leser som har erfaringer fra skolen gjenkjenner lett situasjonen i denne teksten. Handlingen er lagt til i et klasserom. Timen er snart slutt, og guttene Per og Gøran står klare til å løpe ut for å spille fotball. Ingen ting av dette står direkte i teksten, leseren må selv gjøre disse inferensene om situasjonen i teksten.

Undervisning i inferens innebærer en pedagogisk utfordring – det å undervise i det implisitte. Avhandlingen angriper denne utfordringen to måter. For det første ved å sammenfatte tidligere studier av inferenstrening blant skolebarn i en metaanalyse, og for det andre ved å gjennomføre en effektstudie av inferenstrening blant norske elever. Metaanalysen undersøker hvilke effekter tidligere studier har gitt, og hva det har blitt fokusert på i treningen av inferens. Tidligere studier av inferenstrening som har blitt identifisert i forbindelse med denne avhandlingen, har kun vært gjennomført blant engelskspråklige barn. Resultatene kan derfor ikke uten videre overføres til norsk skole. Av den grunn har inferenstrening blant norske skolebarn en sentral plass i denne avhandlingen. På bakgrunn av erfaringer fra utenlandske studier av inferenstrening ble det utviklet en inferenstrening tilpasset norske elever og lærere. Effekten av treningen ble undersøkt i et pre- postdesign med kontrollgruppe. I tillegg ble en utsatt posttest inkludert for å undersøke effekter av inferenstreningen etter at den var avsluttet.

I tillegg til den pedagogiske motivasjonen om å undersøke hvordan man kan undervise i inferens, har avhandlingen også en teoretisk motivasjon. Den teoretiske motivasjonen er å undersøke årsaksforholdet mellom inferens og leseforståelse. Dersom trening i inferens fører til framgang i leseforståelse, kan inferens betraktes som en årsak til leseforståelse.

## **1.2 Avhandlingens oppbygning**

Flere ferdigheter og kompetanser er nødvendige for å kunne lese med forståelse, for eksempel ordforråd, bevissthet om tekststruktur, monitorering og inferenser. Betydningen av disse ferdighetene for leseforståelse er utledet fra studier av barn med svak leseforståelse tross adekvate avkodingsferdigheter. I kapittel 2 redegjøres det for disse ferdighetene, samt betingelser for leseforståelse.

Kapittel 3 gir en yterligere redegjørelse for inferensbegrepet. Fordi inferensbegrepet defineres og kategoriseres på flere måter i faglitteraturen, ble det i arbeidet med denne avhandlingen laget en egen taksonomi av inferenstyper. Denne taksonomien blir presentert i kapittel 3.

Kapittel 4 gir begrunnelser for effektstudier i leseforskning, med spesielt fokus på effekter av inferenstrening. Videre redegjøres det for kvasi-eksperimentelle design. Både validitetsmessige og etiske problemstillinger knyttet til dette designet diskuteres i kapittelet.

I kapittel 5 oppsummeres tidligere studier av inferenstrening i en metaanalyse. Metaanalysen undersøker hvilke effekter inferenstrening har både på elevenes evne til å gjøre inferenser spesielt, og på leseforståelse generelt. I tillegg søker metaanalysen svar på hvorvidt eventuelle effekter vedvarer, og hva som karakteriserer inferenstrening med sterk effekt.

Som et ledd i utviklingen av undervisningsmateriell i inferenstreningen som det gjøres rede for i kapittel 7, ble det gjennomført en forundersøkelse som presenteres i kapittel 6. Forundersøkelsen ble gjennomført blant elever på niende trinn, og har fokus på betydningen av situasjonsmodeller for elevenes evne til å gjøre inferenser. I kapittelet blir både materiell og resultater gjort rede for og diskutert.

I kapittel 7 presenteres en effektstudie av inferenstrening blant elever på sjette trinn. Effektstudien søker svar på hvilke effekter inferenstrening har på inferens og leseforståelse både umiddelbart etter at treningen var avsluttet og fem uker etter avsluttet trening. I tillegg søker studien svar på hvilke elever som har mest utbytte av å trene på inferens. I kapittelet redegjøres det for innholdet i og gjennomføringen av treningen. Videre diskuteres muligheter for generalisering av resultatene.

Til slutt, i kapittel 8, blir øvrige kapitler oppsummert.





## 2. Leseforståelse

### 2.1 Hva innebærer det å kunne lese med forståelse?

Lesing innebærer to hovedkomponenter: Avkoding og språkforståelse. Avkoding innebærer visuell gjenkjenning av bokstaver og ord, og karakteriseres som det tekniske aspektet ved lesingen. Avkodingen er avgjørende for lesingen, men avkoding alene gir ikke en adekvat lesing. Mens avkodingen er spesifikk for lesing, er språkforståelse involvert både i lytting og lesing. Gough og Tunmer (1986) og Hoover og Gough (1990) sin teori om *the simple view of reading* framstiller lesing som produktet av de to elementene ganget med hverandre: Lesing = Avkoding x Forståelse. Begge komponentene må beherskes for å kunne lese. Dersom en av komponentene ikke beherskes, blir produktet, lesingen, lik null.

Avkodingsprosessen, som innebærer gjenkjenning av skrifttegn, er grunnleggende for lesingen. Derfor er det viktig å etablere gode, automatiserte strategier for avkoding. Korrelasjonen mellom avkoding og leseforståelse varierer fra .03 til .07 i ulike studier (Paris & Hamilton, 2009), og det har lenge vært kjent at barn med svak leseforståelse bruker mer tid på å lese ord og nonord enn barn med en adekvat leseforståelse (Perfetti & Hogaboam, 1975). Dersom avkodingen ikke er automatisert, skaper den en "flaskehals" i arbeidsminnet, og leseren blir nødt til å bruke en større del av sin kognitive kapasitet til avkoding på bekostning av forståelsen (Perfetti, 1985). Avkodingsferdigheter har ikke bare betydning for unge og uerfarne lesere. I en studie av norske ungdomsskoleelever har Samuelstuen og Bråten (2005) vist at avkodingsferdigheter spiller en rolle også for erfarne leseres leseforståelse. En longitudinell studie av Catts, Hogan og Adlof (2005) har imidlertid vist at den unike variansen av lytteforståelse og avkoding i leseforståelse endret seg etter som elevene ble eldre. Mens den unike variansen av avkoding og lytteforståelse var henholdsvis 27 og 9 prosent i andre klasse, endret den seg til 13 og 21 prosent i fjerde klasse, og 2 og 36 prosent i åttende klasse. Rapp, van den Broek, McMaster, Kendeou og Espin (2007) har gjennom en oppsummering av tidligere forskning vist at basisferdigheter i lesing, som for eksempel fonologisk bevissthet, avkoding og leseflyt, og ferdigheter som er involvert i leseforståelse, utvikler seg simultant. De understreker at det derfor er viktig å fokusere på både basisferdigheter og leseforståelse i begynneropplæringen.

Som det ble gjort rede for i kapittel 1, hevder Snow og Sweet (2003) at leseforståelse er resultatet av to prosesser, det å trekke ut informasjon og det å

konstruere mening, eller, for å bruke Elbro (2007) sine ord, gjenskape mening. Det å hente ut informasjon fra tekst er nært knyttet til det teksten eksplisitt uttrykker. Det å gjenskape eller rekonstruere mening krever at leseren må integrere kunnskaper og erfaringer fra utenfor teksten med det som er direkte uttalt i teksten. Denne tosidige prosessen, tekstnær forståelse og rekonstruksjon av mening, finner vi også i Kintsch sin konstruksjon-integrasjonsmodell (Kintsch, 1998, 2005). Kintsch sin modell deler forståelsen inn i tre nivåer. Ifølge Kintsch er en begrepsmessig og perseptuell avkodning det første steget i forståelsesprosessen. Ord og setninger kan ha flere betydninger, det er derfor viktig at leseren danner spesifikke sammenhenger mellom objekter og deres kontekst. Denne begrepsmessige og perseptuelle avkodningen danner grunnlaget for det andre nivået i leseforståelse, som Kintsch kaller tekstbase. Tekstbasen består av mikro- og makrostrukturer. De semantiske forbindelsene mellom ord og setninger utgjør mikrostrukturen. Når vi leser, finner vi hint i teksten om hvordan ord og setninger henger sammen. Vi forventer at det er sammenheng mellom den setningen vi leser og den vi nettopp leste. Ord som *derfor* og *fordi* hjelper leseren å knytte setningene sammen. Makrostrukturer er forbindelser på tekstnivå. Når vi forstår en tekst, knytter vi ifølge Kintsch hovedideene i teksten sammen. Videre strukturerer leseren relasjoner i teksten hierarkisk, fra overordnede tema til detaljinformasjon. Denne hierarkiske inndelingen av informasjonen i teksten hjelper leseren med å skille ut det som er viktig fra det som er mindre viktig i teksten. Makrostrukturen er også betydningsfull i forhold til å kunne gjenkalle innholdet i teksten, fordi overordnede ideer setter oss på sporet av informasjon lengre ned i hierarkiet. Teksten gir oss ofte hint om makrostrukturen i form av for eksempel overskrift eller tittel, punktvis kapitteinndeling eller illustrasjoner.

Forståelse basert på tekstbase er forholdsvis bokstavelig og overfladisk, og i de fleste lesesituasjoner ikke tilstrekkelig for en adekvat leseforståelse. Å lese med forståelse krever at leseren også knytter innholdet i teksten til egen erfaring og skaper indre bilder, eller representasjoner, av situasjoner og hendelser i teksten. En slik rekonstruksjon av mening karakteriseres gjerne som en situasjonsmodell (Kintsch, 1998) eller en mental modell (Johnson-Laird, 1983). Situasjonsmodeller er i stor grad knyttet til vår egen bakgrunn og kunnskap, og gir en dypere forståelse utover det teksten eksplisitt uttrykker.

Kintsch (1998) understreker at tekstbasen kan være mer eller mindre koherent og komplett, likeså kan situasjonsmodellen være mer eller mindre adekvat og presis. Vanligvis er forståelsen basert på en blanding av informasjon som er

trukket direkte ut fra teksten (tekstbase), og kunnskapsbasert og erfaringsbasert informasjon (situasjonsmodell). Forståelsen av en tekst kan imidlertid preges sterkere av tekstbasen enn situasjonsmodellen, eller omvendt. Dersom leserens kunnskap om tekstens tema er svært begrenset, vil forståelsen hovedsakelig bero på tekstbasen. Omvendt vil situasjonsmodellen være hovedkilde til forståelsen av teksten dersom den eksplisitte teksten gir lite informasjon og tekstbasen dermed er lite koherent.

## **2.2 Ferdigheter av betydning for leseforståelse**

Siden lesing avhenger av både avkoding og språkforståelse, kan vansker med lesing også knyttes til begge elementene. Elbro (2007) skiller mellom vansker knyttet til avkoding og vansker knyttet til språkforståelse, samt sammensatte vansker som innebærer problemer med begge elementer. De spesifikke vanskene med forståelse, hyperleksi, kan ifølge Elbro ofte ha sammenheng med andre gjennomgripende vansker som for eksempel autisme. Likevel er det mange barn som har svak leseforståelse sett i forhold til deres avkodingsferdigheter. Studier har anslått at som mye som 10 % av britiske skolebarn har problemer med å forstå det de leser på tross av at de har adekvate avkodingsferdigheter (Yuill & Oakhill, 1991).

Når kan vi si at vi har en god leseforståelse? Leseforståelse kan betraktes som en gjensidig prosess mellom leser og tekst. Leserens bakgrunn, leseerfaring og forutsetninger bidrar til en unik forståelse av teksten. Det kan derfor være vanskelig å sette fingeren på når leseforståelsen er god. Derimot kan det være lettere å identifisere når forståelsen av en tekst uteblir. Tegn på svak leseforståelse kan gi oss informasjon om hva som kreves for å oppnå en adekvat leseforståelse. Det er gjennomført en rekke studier som sammenligner barn som behersker både avkoding og forståelse med barn som har svak leseforståelse tross god avkoding. De spesifikke områdene som barn med svak leseforståelse sliter med, kan gi oss informasjon om prosessene som er involvert i leseforståelse. Studier som kartlegger profiler av barn med forståelsesvansker tross adekvat avkoding kan gi oss kunnskap om spesifikke årsaker til problemer med leseforståelse, og hvilke forutsetninger som må være til stede for å oppnå en god leseforståelse.

### **2.2.1 Ferdigheter som årsak til eller konsekvens av leseforståelse**

Studier av barn som har vansker med leseforståelse til tross for adekvat avkoding, kan bidra til å identifisere ferdigheter av betydning for leseforståelse. Disse studiene sier ikke nødvendigvis noe om retningen på forholdet mellom en ferdighet og leseforståelse. Høy korrelasjon mellom en ferdighet av betydning for leseforståelse og den overordnede leseforståelsen

bekrefter en sammenheng, men kan ikke fortelle oss om retningen på sammenhengen. En sammenheng kan skyldes flere forhold: 1) At ferdigheten er en følge av leseforståelse, 2) at leseforståelse er en følge av ferdigheten, eller 3) at sammenhengen mellom en ferdighet og leseforståelsen er en følge av en bakenforliggende betingelse. Årsaksretninger mellom ferdigheter av betydning for leseforståelse og den overordnede leseforståelse kan bekreftes gjennom studier med følgende design: Comprehension age match-studier, longitudinelle studier og effektstudier.

Studier som anvender comprehension age match (CAM) design, sammenligner to grupper barn: Barn med svak leseforståelse og adekvate avkodningsferdigheter og yngre barn som matcher førstnevnte gruppe med hensyn til leseforståelse. Ofte inkluderes også ofte en tredje gruppe som matcher den førstnevnte gruppen i alder, men som har adekvate ferdigheter både i avkodning og leseforståelse. Dersom gruppen med yngre barn presterer bedre enn gruppen med svak leseforståelse på et mål på en ferdighet av betydning for leseforståelse, utelukker dette at ferdigheten er en følge av leseforståelse. Dette fordi det ikke er sannsynlig at de yngre barnas ferdigheter skyldes mer erfaring med tekst (Oakhill & Cain, 2007). Dette designet kan imidlertid ikke utelukke at sammenhengen er en følge av en bakenforliggende betingelse.

I longitudinelle studier følger man de samme barna over lengre tid, og har minst to måletidspunkter. Dermed unngår man problemer med forutsetningen om asymmetri i kausale slutninger, fordi i dette designet kommer årsaken nødvendigvis før virkningen (Skog, 2004). Designet er derfor godt egnet til å kartlegge hvordan ferdigheter kan predikere leseforståelse. Ifølge Oakhill og Cain (2007) er det imidlertid fare for at de potensielle årsakene (ferdighetene) påvirker virkningen (leseforståelse) via en medierende ferdighet når longitudinelle design benyttes i forskning på leseforståelse.

En tredje måte å undersøke årsaksretningen i forholdet mellom en ferdighet og leseforståelse, er gjennom effektstudier. I effektstudier manipuleres årsaksvariabelen ved at man trener en gruppe elever i en ferdighet. I denne typen studier har man som regel minst to betingelser, en eksperimentgruppe som mottar trening i en ferdighet av betydning for leseforståelse og en kontrollgruppe som enten ikke får endringer i ordinær undervisning, eller får en trening som ikke angår leseforståelse. Effekten måles enten ved å sammenligne de to gruppens resultater etter treningen, eller ved å sammenligne de to gruppens framgang fra pretest til posttest. Dersom eksperimentgruppens posttestresultater eller framgang er bedre eller større enn

kontrollgruppens, er det rimelig å anta at leseforståelse er en følge av den trente ferdigheten. Dette forutsetter naturligvis at treningen er en rimelig operasjonalisering av det teoretiske begrepet. Styrker og svakheter ved effektstudier, samt ulike varianter av effektstudier, vil bli videre diskutert i kapittel 4.

### **2.2.2 Ferdigheter på ordnivå**

Ordforråd er den viktigste enkeltforutsetningen for leseforståelse. Hvis mer enn 20 % av ordene i en tekst er ukjente, uteblir forståelsen (Høien & Lundberg, 2000). Beck og McKeown (1991) har stilt spørsmål ved hva et ordforråd egentlig er, og hva det innebærer å vite betydningen av et ord. Det er ikke slik at vi enten kan eller ikke kan et ord. Beck og McKeown hevder at det er mange nivåer i det å kunne et ord, helt fra det å aldri ha hørt et ord før, det å ha hørt ordet men ikke kjenne betydningen av det, til det å kunne gjenkjenne ordet i kontekst og det å kjenne ordet godt. Ifølge Beck og McKeown er også ulik kunnskap involvert i det å kunne ord. For det første kan man snakke om kunnskap om generalisering, det å kunne definere et ord ved hjelp av andre ord. For det andre kunnskap om hvilket ord som passer best i en gitt situasjon. For det tredje innebærer kunnskap om ord å kjenne til flere betydninger av samme ord. Videre innebærer kunnskap om ord også presisjon i bruken og i gjenkjennelsen av ord og det å kunne bruke ord fleksibelt i tenking og diskusjon.

Ordforråd korrelerer høyt (0.66 til 0.75) med leseforståelse (Joshi, 2005), likevel er kausalitet mellom ordforråd og leseforståelse vanskelig å påvise. Kamil (2004) hevder at dette kan ha sammenheng med at ordforråd gjerne blir knyttet til enkeltord, mens forståelse handler om helhet, det å se sammenhenger. Ordforråd er nødvendig for leseforståelse, men det er ikke tilstrekkelig. Ordforråd og forståelse er tett knyttet sammen, det er vanskelig å skille de to prosessene fra hverandre. Ordforråd bidrar til leseforståelse, men lesing bidrar også til å utvide ordforrådet. Med andre ord er det en gjensidig avhengighet mellom de to variablene.

Hvor mye man eksponeres for tekst har betydning for ordforrådet. Cunningham og Stanovich (1997) fant at forskjeller i ordforråd kunne forklares gjennom forskjell i teksteposponering. Ifølge Joshi (2005) er det imidlertid forskjell på i hvilken grad sterke og svake lesere profitterer på lesing med tanke på å utvide ordforrådet. Svake lesere velger ofte å lese enklere bøker med færre nye ord enn sterke lesere, og lærer derfor færre nye ord av å lese. Sterke og svake lesere er også forskjellige med hensyn til hvordan de angriper ukjente ord i en tekst. Cain, Oakhill og Elbro (2003) fant

at barn med svak leseforståelse hadde vanskeligere for å trekke slutninger om betydningen av ukjente ord på bakgrunn av kontekst, enn barn som hadde leseforståelse som forventet for alder.

Det er ikke bare størrelsen på ordforrådet som har betydning for leseforståelsen. Det er også forskjell mellom barn med adekvat og svak leseforståelse når det gjelder dybde i ordforrådet og begrepsforståelsen. Nation og Snowling (1999) har undersøkt i hvilken grad ”priming”, det vil si evnen til å aktivere semantiske assosiasjoner mellom ordpar, er forskjellig hos barn med aldersadekvat leseforståelse og hos barn med svak leseforståelse på tross av adekvat avkodning. Nation og Snowling fant ingen forskjell mellom sterke og svake lesere med hensyn til funksjonsrelaterte ordpar som for eksempel *vaskekost – gulv*. I forhold til kategorirelaterte ordpar som for eksempel *katt – hund*, var det heller ingen forskjell mellom sterke og svake lesere så lenge ordene var velkjente. Barn med svak leseforståelse klarte imidlertid ikke, i motsetning til barn med adekvat leseforståelse, å aktivere assosiasjoner for ordpar som ikke var dagligdagse. Med andre ord hadde barn med svak leseforståelse svakere forbindelser mellom ord som ikke inngikk i deres hverdag.

Tannenbaum, Torgesen, og Wagner (2006) har undersøkt sammenhenger mellom leseforståelse og tre dimensjoner av ordkunnskap: dybde og bredde i ordforrådet og leseflyt. Dybde handler om graden av kunnskap om det enkelte ord, for eksempel det å kjenne ulike betydninger av samme ord. Bredde i ordforrådet handler om omfanget av ord en leser har kunnskaper om. Tannenbaum et al. fant at bredde i ordforrådet var sterkere relatert til leseforståelse enn dybde og leseflyt. I hvilken grad leseren er bevisst doble betydninger av ord og uttrykk har også innvirkning på leseforståelsen. Effektstudier av Yuill (2007) har vist at strukturerte diskusjoner om tvetydighet i språket ga signifikant forbedring i leseforståelse, og det var særlig de svake leserne som profitterte på diskusjoner om doble betydninger av ord og uttrykk.

Beck og McKeown (1991) skiller mellom to hovedtilnærminger til å lære nye ord, gjennom tilfeldige læringssituasjoner og ved bevisst tilrettelegging av læringssituasjoner. Swanborn og Glopper (1999) har gjennomført en metaanalyse av studier som undersøkte førstnevnte tilnærming til å lære nye ord. Felles for alle studiene i utvalget var at læring av nye ord skjedde gjennom tilfeldig lesing uten at det ble fokusert på ordforråd overfor elevene. Metaanalysen viste at elevene i studiene lærte seg betydningen av om lag 15 % av alle nye ord de støtte på gjennom lesing. Videre viste studien at jo eldre

elevene var, desto flere nye ord lærte de gjennom lesing. Tilsvarende viste resultatene fra metaanalysen at jo høyere skåre elevene hadde på tester i leseferdighet, desto flere ord lærte de gjennom tilfeldige læresituasjoner.

Når det gjelder den andre tilnærmingen til å lære nye ord, bevisst tilrettelegging av læringssituasjoner, har flere effektstudier vist at barns ordforråd kan utvides gjennom systematisk trening. For eksempel har Apthorp (2006) undersøkt effekter av å trene på nye ord. Studien ble gjennomført blant tredjeklassinger i to skoledistrikter, med en eksperimentgruppe og en kontrollgruppe i hvert distrikt. Eksperimentgruppene fikk undervisning i ordforråd hvor det ble fokusert på 7 nye ord hver uke i en periode på 24 uker. Elevene i eksperimentgruppene trente på ordene ved blant annet å lese tekster som inneholdt ordene, matche ord med bilder og skrive ned sine egne erfaringer med ordene. Effektene av ordforrådstreningen ble målt ved pre- og posttest. Resultatene viste en signifikant interaksjon mellom tid og gruppe for ekspressivt vokabular i det ene skoledistriktet, men ikke det andre. Vokabulartreningen hadde ikke signifikant effekt på impressivt ordforråd, det var heller ingen signifikant effekt på leseforståelse i noen av de to skoledistriktene.

En metaanalyse av Mezynski (1983) viste samme tendens til at ordforrådstrening ikke hadde generalisert effekt på leseforståelse. Mezynski mente at en forklaring på hvorfor ordforrådstrening ikke hadde effekt på leseforståelse, kunne ha med automatisering av ordkunnskap å gjøre. Stahl og Nagy (2006) utdyper dette argumentet, og hevder at personer som har mye ordkunnskap, generelt har dypere kunnskap om ordenes betydning og gjør raskere koblinger mellom det skrevne ord og ordets betydning enn personer med lite ordforråd. Manglende effekter av ordforrådstrening på leseforståelse kan også ha sammenheng med hvilke mål på leseforståelse som ble brukt. En metaanalyse av ordforrådstrening av Stahl og Fairbanks (1986) viste at den gjennomsnittlige effektstørrelsen av ordforrådstrening på leseforståelse målt med tekster som inneholdt ord som inngikk i treningen var .97, hvilket kan karakteriseres som en sterk effekt. Det var også en gjennomsnittlig effekt av treningen på leseforståelse målt med prøver som ikke inneholdt ord som det var blitt trent. Denne effekten var imidlertid betraktelig mindre, bare .30. Metaanalysen viste også at ordforrådstrening som ga sterk effekt hadde til felles at elevene ble eksponert for ordene i flere kontekster og at de fikk informasjon om både definisjoner av ordet og kunnskap om hvilken kontekst ordet hørte hjemme i.



I skolen er utviklingen av ordforråd en stor utfordring fordi forskjeller i ordforråd hos elevene oppstår tidlig i livet, før de begynner på skolen. Dickinson og Tabors (2001) har gjennom en longitudinell studie påvist sammenheng mellom aktiviteter i hjemmet og ferdigheter relatert til lesing og skriving i barnehage og skole. Studien viste at det var forskjell i hvordan foreldrene støttet barnas språkutvikling. Tabor og Dickinson fant blant annet at i familier med lav inntekt hadde foreldrene færre samtaler med barna, noe som hadde sammenheng med for barnas ordforråd. Studien viste videre en korrelasjon mellom skåre på ordforråd i barnehagen og leseforståelse i fjerde og sjuende klasse, korrelasjonen ble sterkere etter som elevene ble eldre. Barn som hadde gode lese- og skriveferdigheter i tidlig alder, hadde en tendens til å forbli gode lesere, mens de svake forble svake. Matteuseffekt er en vanlig betegnelse på en slik type utvikling, og henspiller på en passus i Matteusevangeliet som sier: "For den som har, han skal få, og det i overflod. Men den som ikke har, han skal bli fratatt selv det han har" (Matt. 25, 29). Matteuseffekten forsøker å beskrive hvordan spriket mellom elevenes ferdigheter øker med årene, og hvordan flinke elever blir enda flinkere mens de svake elevene forblir svake selv i samme klasserom. Vi kan snakke om et gjensidig forhold der leseferdigheter og individuelle egenskaper og /eller miljø bidrar til å bedre eller hemme leseutviklingen. Ordforråd bidrar til matteuseffekt i lesing fordi det å kunne mange ord bidrar til at det blir lettere å lære nye ord. Ordforråd bidrar til leseforståelse, samtidig som lesing bidrar til å utvide ordforrådet. Dermed forsterkes forskjeller i ordforråd gjennom at de som kan mange ord, har bedre forståelse og lærer flere ord fordi de leser mye. Penno, Wilkinson og More (2002) gjennomførte en studie som hadde til hensikt å forbedre barns ordforråd gjennom eksponering for litteratur. I tillegg til å undersøke den generelle effekten av ordforrådstrening, var Penno et al. også interessert i å finne ut om treningen bidro til å utjevne voksende forskjeller i barnas ordforråd. Med andre ord om treningen kunne redusere matteuseffekt. I studien fikk barn i alderen fem til åtte år lære nye ord under forskjellige betingelser. Penno et al. fant at høytlesning bidro til å utvide barnas ordforråd, men effekten ble større når høytlesningen ble kombinert med at lærer ga forklaringer til de nye ordene i historiene som ble lest høyt. Resultatene viste at matteuseffekten ikke ble redusert som følge av treningen, tendensen var snarere tvert om. De sterkeste barna profitterte mest på ordforrådstreningen.

### **2.2.3 Bevissthet om tekststruktur**

Måten teksten er organisert på har sammenheng med innholdet i teksten, for eksempel hvilken type informasjon vi kan vente å finne i teksten og hvilken informasjon i teksten som er over- og underordnet. Strukturen varierer i ulike

typer tekster. For eksempel har tekster som har til hensikt å underholde, en annen struktur enn tekster som er informerende. I narrative tekster er tekststrukturen basert på fortellingens grammatikk, eller "story grammar" på engelsk. Ifølge Pearson og Fielding (1991) består fortellingens grammatikk av abstrakte lingvistiske representasjoner av blant annet rammen for fortellingen, hendelser, personlige motiver, problem og mål. Argumentet for betydningen av fortellingens grammatikk for forståelsen er, ifølge Pearson og Fielding, at disse kjennetegnene ved fortellingen gjør det lettere for leseren å kode informasjonen i langtidsmindet. Å benytte seg av strukturen i fagtekster for å oppnå forståelse, innebærer blant annet å følge forfatterens argumentasjon og hvordan ideene i teksten blir presentert.

En studie av Cain (1996) har vist at barn som har vansker med leseforståelse tross adekvate avkodingsferdigheter, er mindre bevisste på underliggende strukturer i teksten enn barn uten vansker. I Cains studie ble CAM-design benyttet, hvor barn med og uten vansker med leseforståelse ble sammenlignet med en gruppe yngre elever som matchet den førstnevnte gruppen med hensyn til leseforståelse. Cain fant at gruppen elever med svak leseforståelse produserte fortellinger som var mindre strukturerte enn de yngre elevene med samme nivå i leseforståelse. Med andre ord var elevgruppen med spesielle vansker med leseforståelse særlig lite bevisste på tekststruktur. Cain sin studie tok utgangspunkt i narrative tekster. I en oppsummering av forskning på tekststruktur og forståelse av fagtekster fant Pearson og Fielding (1991) at elever som hadde mer kunnskap om tekstens struktur, i større grad husket innholdet. Videre konkluderte de også med at sterke lesere i større grad benyttet seg av forfatterens struktur når de skulle gjenkalle en tekst sammenlignet med svake lesere.

To deskriptive studier av Boulineau, Fore, Hagan-Bruke og Bruke (2004) og Gardill og Jitendra (1999) viste at elever med lærevansker hadde en tydelig forbedring i å gjenkjenne elementer i fortellingens grammatikk, som for eksempel rammen for og hendelser i fortellinger, etter å ha mottatt instruksjon i dette. Videre har en studie av Idol (1987), som inkluderte både elever med og uten lærevansker, vist at trening i tekststruktur med narrative tekster bidro til forbedring i elevenes leseforståelse. Pearson og Fielding (1991) sin oppsummering av forskning på tekststruktur viste også at trening i tekststruktur på narrative tekster bidro til økt leseforståelse for deltakerne, der hvor leseforståelse ble målt med tekster fra ordinære lesebøker.

Williams med kolleger (Williams, 2005; Williams et al., 2007; Williams, Stafford, Lauer, Hall, & Pollini, 2009) har gjennom flere effektstudier vist at

ekspisitt og strukturert undervisning i tekststruktur i fagtekster har effekt på andreklassingers leseforståelse. Tilsvarende funn har også blitt gjort av Dickson, Simmons og Kameenui (1998) i en sammenfatning av studier av forholdet mellom tekststruktur (både i forhold til narrative tekster og fagtekster) og leseforståelse. Dickson et al. konkluderte med at elever som hadde bevissthet om, eller som hadde fått undervisning i tekststruktur, hadde en bedre overordnet leseforståelse enn elever som i mindre grad var bevisste eller som ikke hadde fått undervisning i tekststruktur. Det var imidlertid ingen forskjell med hensyn til forståelse for detaljer i teksten.

#### **2.2.4 Monitorering**

Monitorering handler om å overvåke egen kognisjon under lesing, for eksempel å oppdage motsigelser i teksten eller uoverensstemmelser mellom informasjon i teksten og kunnskap om tekstens tema. Monitorering innebærer også å ta i bruk strategier når forståelsen uteblir. Ifølge Pearson og Fielding (1991) er det å oppsummere innholdet i teksten og det å stille spørsmål til seg selv eksempler på strategier for å monitorere egen forståelse under lesing. Det er ikke oppsummeringen eller spørsmålene i seg selv som bidrar til forståelsen, men spørsmålene fører til at kognitive prosesser settes i gang hos leseren. Spørsmålene kan for eksempel bidra til at leseren oppdager motsetninger mellom sin tidligere kunnskap og informasjon i teksten, eller erkjennelsen av å ikke ha forstått innholdet i teksten. Pearson og Fielding viser til en rekke studier som har vist at barn med ulike ferdigheter kan lære seg å monitorere sin egen forståelse gjennom å trene på de ovennevnte strategiene.

Oakhill, Hartt og Samols (2005) har undersøkt evne til overvåking av egen leseforståelse hos barn i alderen 9 til 10 år. I studien sammenlignes barn med svak og god leseforståelse. De to gruppene var matchet med hensyn til ordforråd og avkoding. Barna leste korte tekster som inneholdt enkelte uregelrette setninger der rekkefølgen på ordene var byttet om. Barna leste tekstene under to betingelser. Den ene betingelsen innebar at barna leste tekstene høyt og svarte på spørsmål uten å bli forberedt på at tekstene inneholdt uregelmessigheter. Tegn til at barna oppdaget uregelmessighetene, for eksempel at de stoppet eller nølte, ble registrert. Barna ble også spurt om de hadde oppdaget noe unormalt ved tekstene. Resultatene viste 67 % av de sterke leserne oppga at de hadde oppdaget uregelmessighetene i tekstene, mot 17 % av de svake leserne. Den andre betingelsen innebar at barna leste tilsvarende tekster etter å ha blitt gjort oppmerksomme på at tekstene inneholdt uregelmessigheter. Her viste resultatene at barn med god leseforståelse var mer presise i overvåkingen av sin egen leseforståelse, det vil si at de i

signifikant større grad kunne peke ut hvilke setninger som hadde uregelmessigheter. Videre fikk deltakerne i studien lese tekster som inneholdt to setninger med motstridende informasjon. I noen av tekstene var de to motstridende informasjonene i to påfølgende setninger, mens i de andre tekstene var det tekst mellom de to setningene som inneholdt motstridende informasjon. De svake leserne presterte generelt dårligere enn de sterke leserne på disse oppgavene. Det var også stor forskjell mellom de to gruppene med hensyn til tekster der det var avstand i teksten mellom de to motstridende informasjonene.

### **2.2.5 Inferens**

De færreste tekster er fullt ut eksplisitte. Inferens skaper sammenheng i leserens rekonstruksjon av teksten der hvor sammenhengen ikke er eksplisitt uttalt. Uten inferenser forblir teksten fragmenter av ord og setninger uten en overordnet mening. Derfor er inferens en ferdighet av stor betydning for leseforståelse. Ytterligere redegjørelse for inferensbegrepet og kategorier av inferenstyper vil bli gitt i kapittel 3.

Cain og Oakhill (1999) har vist at barn som strever med leseforståelse til tross for at de har adekvate avkodingsferdigheter, gjør færre inferenser når de leser enn barn uten forståelsvansker. Det er en sterk sammenheng mellom leseforståelse og evne til å gjøre inferenser (Cain, Oakhill, Barnes, & Bryant, 2001). Sammenhengen kan tolkes på to måter. For det første at svak leseforståelse er forårsaket av manglede evne til å gjøre inferenser. For det andre at evne til å gjøre inferenser påvirkes av leseerfaring, slik at de som har god leseforståelse får gode evner til å gjøre inferenser på grunn av at de leser mye.

Cain og Oakhill (1999) har kartlagt årsaksretningen i forholdet mellom inferens og leseforståelse ved bruk av CAM-design. To typer inferens ble inkludert i studien. For det første inferens som krevde at barna skapte sammenheng mellom flere setninger, og for det andre inferens som krevde at barna integrerte kunnskap fra utenfor teksten med det som var eksplisitt uttalt i teksten. Resultatene viste at kontrollgruppen med yngre barn skåret høyere på inferensoppgaver enn gruppen med svak leseforståelse. Siden det er lite trolig at de yngre barnas bedre evne til å gjøre inferenser skyldes at de hadde mer erfaring med tekst, kan det utelukkes at inferens er et produkt av god leseforståelse. Inferens er derfor snarere en årsak til leseforståelse. Studien viste også at barna med svak leseforståelse skåret signifikant lavere enn gruppen med yngre barn på inferensspørsmål som krevde at barna skapte sammenheng mellom setninger. På inferensspørsmål som krevde integrering

av kunnskap fra utenfor teksten, skåret barna med svak leseforståelse signifikant lavere enn jevnaldrende barn med god leseforståelse, men det var ikke signifikant forskjell mellom barn med svak leseforståelse og den yngre matchgruppen.

I en oppfølging av den ovennevnte studien fant Cain et al. (2001) at barn med svak leseforståelse hadde vanskeligere for å gjenkalle relevante premisser i teksten enn barn med god leseforståelse. Barn med god leseforståelse hadde lettere for både å skaffe seg ny kunnskap og lettere for å anvende nylig tilegnet kunnskap. Selv om barna med god leseforståelse i stor grad både benyttet seg av informasjon i teksten og kunnskap utenfor teksten, hadde de av og til problemer med å integrere informasjon fra teksten og kunnskap fra utenfor teksten. Dette tolker Cain et al. som at problemer med å gjøre inferens arter seg på forskjellig måte hos barn med god og svak leseforståelse; problemene oppstår på et tidligere stadium hos barn med svak leseforståelse.

En longitudinell studie har vist at inferens predikerer leseforståelse når det er kontrollert for avkodning, ordforråd og verbale ferdigheter (Cain, Oakhill, & Bryant, 2004). En rekke effektstudier kan også bekrefte årsaksretningen mellom inferens og leseforståelse, se kapittel 5 for en oversikt over disse studiene.

Kendeou og van den Broek med kolleger (Kendeou, Bohn-Gettler, White, & van den Broek, 2008; Kendeou, van den Broek, White, & Lynch, 2007) har gjort studier som gir innsikt i hvordan inferensferdigheter utvikles og hvilken betydning inferens har for leseforståelse. I en longitudinell kohortstudie ble barns ferdigheter i leseforståelse og inferens undersøkt. Det ble foretatt målinger av to grupper barn på to måletidspunkter; da den ene gruppen barn var 4 og 6 år, og samtidig av en annen gruppe barn da de var 8 og 10 år. Barnas ferdigheter i inferens og leseforståelse ble målt muntlig og via historier vist på TV, og i tillegg skriftlig for de eldste barna. Studien viste en sterk sammenheng mellom barnas evne til å gjøre inferenser på tvers av ulike medier, både for de yngste barna og for de eldste. Det ble gjort tilsvarende funn for forståelse. Disse funnene bekrefter at verken inferens eller leseforståelse er knyttet til lesing spesifikt. Det var en sterk korrelasjon mellom inferens og leseforståelse på tvers av medier på begge tidspunkter. Inferens predikerte leseforståelse i større grad enn språklige ferdigheter (fonologisk bevissthet, bokstav- og ordavkodning) og ordforråd i begge kohortene over begge tidspunkter. Verken for de yngste eller de eldste barna i studien var det sammenheng mellom inferensferdigheter og språklige ferdigheter som ligger til grunn for avkodning som fonologisk bevissthet,

bokstav- og ordavkodning, eller mellom forståelsesferdigheter og språklige ferdigheter.

### **2.2.6 Høyere standarder for sammenheng**

Å lese med forståelse innebærer å danne seg en sammenhengende forståelse av teksten. De områdene som elever med spesifikke vansker med leseforståelse sliter med på tekstnivå, handler alle om å skape en sammenhengende forståelse. *Bevissthet om tekststruktur* handler om oppmerksomhet på hvordan teksten er bygd opp og hvordan ideer i teksten henger sammen. *Å gjøre inferenser* der teksten er implisitt, innebærer å skape sammenheng i rekonstruksjonen av tekstens innhold. *Monitorering* er nødvendig for å vite hvorvidt forståelsen er en rimelig tolkning av teksten. Gjennom monitorering verifiserer leseren sin forståelse og reagerer dersom forståelsen uteblir. Disse tre komponentene handler alle om å skape sammenheng under lesing, og blir av noen forskere betraktet som eksempler på *en* generell ferdighet.

van den Broek med kolleger (Rapp et al., 2007; van den Broek et al., 2005; van den Broek & Kremer, 2000) understreker betydningen av sammenheng, eller koherens, som en generell ferdighet som er avgjørende for å oppnå en adekvat leseforståelse. De ser teksten som et nettverk bestående av tekstdeler. For å skape en sammenhengende oppfatning av teksten må leseren kunne oppfatte relasjonene mellom de ulike tekstdeler. van den Broek med kolleger hevder at forskjeller i leseforståelse kan forklares som en følge av ulike standarder for sammenheng. Det er altså forskjell på lesere med hensyn deres krav til egen oppfattelse av sammenhengen mellom ideene i teksten.

De svake leserne er ikke på samme måte som de gode leserne orientert mot å skape sammenheng mens de leser. Ikke bare har de sterke leserne mer "driv mot koherens" enn de svake leserne, de har også i stor grad automatisert prosessen hvor de knytter tekstdeler sammen (van den Broek & Kremer, 2000).

## **2.3 Betingelser for leseforståelse**

### **2.3.1 Verbale og nonverbale ferdigheter**

Siden tekst er en skriftlig form av det muntlige språket, er språkforståelsen grunnlaget for leseforståelse. Spira, Bracken og Fischel (2005) fant i en longitudinell studie at grunnleggende muntlige språkferdigheter målt i barnehagen bidro til leseferdigheter i fjerde klasse. Studien viste også at tidlige språklige ferdigheter var avgjørende for hvorvidt barna overvant

vansker med lesing når de begynte på skolen. I det norske utvalget i PIRLS-undersøkelsen fra 2006 fant man også sammenheng mellom språklige aktiviteter i hjemmet og leseferdigheter på fjerde trinn (Solheim et al., 2007).

Som nevnt tidligere kan lesing betraktes som produktet av avkodning ganger språkforståelse. Språkforståelsen er ikke spesifikt knyttet til lesing på samme måte som avkodningen. Vanskene til barn som har en diskrepans mellom avkodning og forståelse, er ikke kun knyttet til lesing. Disse barna har også ofte dårligere lytteforståelse enn barn med adekvat leseforståelse (Floyd, Bergeron, & Alfonso, 2006; Stothard & Hulme, 1992). De presterer også signifikant svakere på mål på generelle verbale ferdigheter enn jevnaldrende uten vansker med leseforståelse (Nation, Clarke, & Snowling, 2002) og yngre barn som matcher med hensyn til leseforståelse (Stothard & Hulme, 1996).

Når det gjelder nonverbale ferdigheter hos barn med god og svak leseforståelse, er resultater fra forskningen motstridende. Mens en studie av Stothard og Hulme (1996) ikke viste signifikante forskjeller mellom gruppene i forhold til nonverbale ferdigheter, har studier av Nation et al. (2002) og Cain og Oakhill (2006) avdekket at det *er* signifikante forskjeller mellom gruppene med hensyn til nonverbale ferdigheter. En rimelig tolkning av resultatene fra disse studiene vil være at dersom det er en sammenheng mellom nonverbale ferdigheter og leseforståelse, er denne sammenhengen ikke særlig sterk.

Dersom vi ser på generelle kognitive ferdigheter, både verbale og nonverbale, ligger barn med leseforståelsesvansker innenfor normalområdet, dog i nedre del av normalområdet (Nation et al., 2002). Det er imidlertid en signifikant forskjell mellom generelle ferdigheter hos barn med og uten vansker med leseforståelse men med adekvat avkodning (Floyd et al., 2006; Nation et al., 2002). Barn med svak leseforståelse tross adekvat avkodning kan deles i to grupper med hensyn til kognitive ferdigheter, 1) barn med gjennomsnittlige kognitive ferdigheter, og 2) barn med kognitive ferdigheter under gjennomsnittet (Cain & Oakhill, 2006; Nation et al., 2002). Nation et al. sin studie viste at den førstnevnte gruppen har avkodningsferdigheter som forventet i forhold til kognitive ferdigheter, men har leseforståelse som er mye svakere enn deres kognitive ferdigheter skulle tilsi. Videre viste resultater fra den samme studien at barn med forholdsvis lav skåre på kognitive ferdigheter hadde bedre avkodningsferdigheter enn kognitive ferdigheter skulle tilsi, mens skåre på leseforståelse var i samsvar med deres kognitive ferdigheter. For denne gruppen var det avkodningsferdighetene som er overraskende gode, ikke leseforståelsen, som var overraskende svak. Det var også forskjell mellom disse to gruppene med hensyn til deres progresjon i leseforståelse. En

longitudinell studie av Cain og Oakhill (2006) viste at barn med kognitive ferdigheter under gjennomsnittet hadde mindre framgang i leseforståelse fra de var 8 til 11 år enn barn med leseforståelsesvansker og gjennomsnittlige kognitive ferdigheter.

Floyd, Bergeron og Alfonso (2006) har også undersøkt kognitive ferdigheter hos barn og ungdommer som har spesifikke vansker med leseforståelse, men til forskjell fra de ovennevnte studiene trekker de inn matematikkferdigheter i tillegg til avkodning og forståelse i lesing. Floyd et al. sammenlignet tre grupper barn: 1) barn med svak leseforståelse men med adekvate avkodnings- og matematikkferdigheter, 2) barn som slet på alle de tre områdene, og 3) barn med adekvate ferdigheter på alle tre områdene. Barna med adekvate ferdigheter på alle tre områdene fikk høyest skåre i forhold til kognitive ferdigheter, barna med vansker på alle tre områder fikk lavest skåre. Både de barna som hadde problemer kun med leseforståelse, og de som hadde vansker på alle tre områder, hadde signifikant svakere skåre på nonverbal resonnering enn barn uten vansker.

Connors (2009) har undersøkt sammenhenger mellom avkodning, språkforståelse, oppmerksomhetskontroll (attentional control) og leseforståelse i et utvalg med 8 år gamle skoleelever. Studien viste at oppmerksomhetskontroll bidro med signifikant varians (mellom 5 og 10 %) i leseforståelse etter at det var kontrollert for avkodning og språkforståelse. Dette resultatet mener Connors er et argument for at oppmerksomhetskontroll bør anses som en tredje komponent i *the simple view of reading* (jevnfør avsnitt 2.1). Connors understreker imidlertid at det er sannsynlig at oppmerksomhetskontroll har mindre betydning for leseforståelse blant eldre elever som en følge av at automatisert avkodning frigjør kognitiv kapasitet hos leseren.

### 2.3.2 Arbeidsminne

I arbeidsminnet foregår en midlertidig lagring og bearbeiding av informasjon. Arbeidsminnet spiller derfor en viktig rolle i forhold til de prosesser i leseforståelse som krever lagring av for eksempel ord eller fakta. Hvorvidt vansker med leseforståelse kan skyldes problemer i verbalt arbeidsminne, har blitt undersøkt i flere studier. Stothard og Hulme (1992) har undersøkt denne problemstillingen ved CAM-design hvor barn med og uten vansker med leseforståelse i alderen 7 til 8 år ble sammenlignet med en gruppe yngre elever som matchet den førstnevnte gruppen med hensyn til leseforståelse. Studien viste ingen signifikante forskjeller mellom gruppene med hensyn til verbalt arbeidsminne, hvilket ble tolket som at problemer med leseforståelse ikke kan tilskrives underliggende problemer med verbalt arbeidsminne. Dette



resultatet står imidlertid i kontrast til hva Nation, Adams, Bowyer-Crane og Snowling (1999) fant noen år senere i en studie hvor de sammenlignet sterke og svake lesere i 10-årsalderen. De sterke og svake leserne i Nation et al. sin studie ble matchet med hensyn til kronologisk alder, nonverbale ferdigheter og avkoding målt ved nonordlesing. Resultatene viste at det var en signifikant forskjell mellom barn med og uten vansker med leseforståelse med hensyn til verbalt arbeidsminne. Forklaringen på disse motstridende resultatene kan være barnas alder. Studier av Seigneuric og Ehrlich med kolleger (Seigneuric & Ehrlich, 2005; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000) har nemlig vist at forholdet mellom arbeidsminne og leseforståelse blir sterkere etter som elevene blir eldre. I en longitudinell studie undersøkte Seigneuric og Ehrlich (2005) sammenhenger mellom blant annet avkoding, ordforråd, arbeidsminne og leseforståelse. Resultatene viste at det i første klasse først og fremst var avkoding som bidro til varians i leseforståelse, mens både ordforråd og avkoding var de sterkeste predikatorene på leseforståelse i andre klasse. I tredje klasse var i tillegg arbeidsminne en sterk prediktor på leseforståelse. I en studie av fjerdeklassinger undersøkte Seigneuric, Ehrlich, Oakhill og Yuill (2000) sammenhenger mellom arbeidsminne og leseforståelse. De fant sammenheng mellom mål på verbalt arbeidsminne og leseforståelse, kontrollert for vokabular og avkoding. Det var imidlertid ingen korrelasjon mellom spatialt minne og leseforståelse, hvilket tyder på at det ikke er arbeidsminne generelt, men verbalt arbeidsminne spesielt som ligger til grunn for leseforståelse.

### **2.3.3 Leserens kunnskap om tekstens tema**

Tekst i seg selv kan være ufullstendig uten våre egne erfaringer og kunnskaper. Den kunnskapen barn har om verden rundt seg, er viktig for å kunne lese med forståelse. Kunnskap om verden kan være kunnskap om situasjoner og hendelser som er en naturlig del av barns liv. Kunnskap om verden generelt er nødvendig, men ikke alltid tilstrekkelig for å forstå tekst. Leseren trenger også relevant kunnskap i forhold til det temaet eller den virkelighet som teksten beskriver. Gough, Hoover og Peterson (1996) illustrerer hvordan leserens kunnskap om tekstens tema er spesifikt knyttet til forståelseskomponenten i lesing, ved å vise til en upublisert studie hvor deltakernes kunnskap om baseball og computere ble målt. Videre ble også deltakeres leseferdigheter, lytteforståelse og avkoding målt med tekster om de to temaene. Resultatene viste korrelasjoner på mellom .56 og .82 mellom skåre på kunnskap om tema og lytteforståelse i tekster med samme tema. Korrelasjonen mellom lytteforståelse i tekster om baseball og lytteforståelse i tekster om computere var på mellom .40 og .62, mens korrelasjonen mellom avkoding i de to temaene var svært høy, hele .88. På bakgrunn av disse

resultatene konkluderte Gough et al. med at avkoding kan betraktes som en generell komponent i lesing, mens forståelsen er mer spesifikk og varierer fra tema til tema, og at det er sannsynlig at noe av variasjonen i forståelsen skyldes variasjon i lesers kunnskap om tekstens tema.

Det er flere teorier om hvordan kunnskap er strukturert og aktiveres hos leseren. Anderson og Pearson (1984) sin skjemateori betrakter leserens kunnskaper som bestående av skjemaer. Et skjema er en abstrakt struktur av generalisert kunnskap og erfaringer. Den kunnskapen vi har, både generell kunnskap om verden og relevant kunnskap i forhold til en gitt tekst, er representert i skjema. Ny informasjon fra tekst knyttes til det leseren allerede kan om temaet. Skjemaene våre hjelper oss med å identifisere hva som er utelatt i teksten, fordi skjemaene gjør at vi kan møte teksten med forventinger om innholdet. I motsetning til skjemateorien vil leseren ifølge konstruksjon–integrasjonsmodellen (Kintsch, 1998) også kunne aktivere informasjon og kunnskap som ikke er relevant for den gitte konteksten. Selv om leseren i første omgang aktiverer irrasjonelle betydninger og assosiasjoner til et ord eller en setning, vil disse betydningene og assosiasjonene raskt bli forkastet. Det å danne spesifikke sammenhenger mellom ord og kontekst utgjør det første nivået eller trinnet i leseforståelsesprosessen ifølge konstruksjon–integrasjonsmodellen.

Det er gjort flere studier som bekrefter betydningen av leserens kunnskap om tekstens tema for leseforståelse. I en studie av Samuelstuen og Bråten (2005) ble betydningen av ungdomsskoleelevers strategier, avkodingsferdigheter og forkunnskaper om tekstens tema undersøkt. Studien viste at kunnskaper om tekstens tema var den av de tre variablene som i størst grad bidro til leserens forståelse av teksten. Kendeou og van den Broek (2007) har undersøkt effekten av lesernes kunnskaper for forståelsen av naturvitenskaplige tekster hos universitetsstudenter. Deltakerne i studien ble delt inn i to grupper ut fra deres forkunnskaper i fysikk. Gruppen som hadde lite forkunnskaper, gjorde signifikant færre inferenser enn gruppen som hadde mye forkunnskaper. I tillegg gjorde gruppen som hadde lite forkunnskaper, signifikant færre riktige inferenser.

Cain et al. (2001) har studert hvorvidt forskjellene mellom barn med god leseforståelse og barn med svak leseforståelse tross adekvat avkoding har sammenheng med barnas kunnskap om tekstens tema. I undersøkelsen ble barna presentert for informasjon om en imaginær planet, ”Gan”. På denne måten kunne forskerne kontrollere for barnas forkunnskaper. Barna fikk opplest en historie hvor handlingen foregikk på planeten Gan. Forskerne

forsikret seg om at alle barna hadde samme kunnskapsbase ved å teste barnas kunnskaper om Gan. Deretter fikk barna spørsmål, hvis svar krevde at de integrerte informasjonen de hadde fått om planeten før historien ble lest, med informasjon i teksten. Selv om alle barna hadde samme utgangspunkt med hensyn til kunnskap til tema, gjorde barna med vansker med leseforståelse likevel signifikant færre inferenser enn barn som hadde god leseforståelse. Dette betyr at det ikke nødvendigvis bare er mengden forkunnskaper som har innflytelse på leserens forståelse av teksten, men også leserens evne til å benytte den kunnskapen han eller hun har. Barn med svak leseforståelse leste på tekstens forutsetning og aktiverte i mindre grad sin kunnskap sammenlignet med barn uten problemer med leseforståelse.

En studie av Brandao og Oakhill (2005) søkte svar på i hvilken grad barn i alderen 7 til 8 brukte sine egne kunnskaper og erfaringer om tekstens tema, og i hvilken grad de brukte informasjonen som var oppgitt i teksten til å svare på forståelsesspørsmål som krevde at barna gjorde to typer inferenser. For det første inferenser som krevde at leseren integrerte informasjon fra to setninger ved å bruke syntaktisk kunnskap og kunnskap av type kategori og overbegrep, for eksempel at koffert er en type bagasje. For det andre inferenser som krevde at leseren brukte sin kunnskap om verden for å oppfatte implisitt informasjon i teksten. Barna leste korte fortellinger og svarte på spørsmål. Etter hvert spørsmål ble barna bedt om å fortelle hvordan de visste svaret på spørsmålet, dette for å avgjøre om de brukte egne kunnskaper og erfaringer eller informasjonen fra teksten. Resultatene viste at teksten helt klart var barnas hovedkilde til å svare på spørsmålene. Noen av barna endret imidlertid svaret sitt da de ble bedt om å fortelle hvordan de visste svaret på spørsmålet. En slik endring mente Brandao og Oakhill talte for dette var en god måte å forbedre barnas leseforståelse på.

#### **2.3.4 Motivasjon**

Motivasjon gir leseren et driv og et mål som er nødvendig for å oppnå forståelse. Motivasjon for lesing predikerer hvor mye barn leser (Wigfield & Guthrie, 1997), som igjen predikerer leseforståelse (Cipielewski & Stanovich, 1992).

Motivasjon er ikke noe man har eller ikke har, vi kan snakke om grader av motivasjon og ulike typer av motivasjon. Motivasjon er viktig på mange områder både i og utenfor skolen. I forbindelse med lesing trekker Guthrie og Humenick (2004) fram tre typer motivasjon av særlig betydning, indre og ytre motivasjon og mestringsforventninger (self-efficacy). Ytre motivasjon handler om å søke ytre belønning for sine prestasjoner, for eksempel i form av gaver,

penger eller påskjønnelse og anerkjennelse fra lærer eller medelever. En studie av Konheim-Kalkstein og van den Broek (2008) blant collegestudenter viste at ytre motivasjon, i form av penger, hadde en positiv innvirkning på leseprestasjoner. Guthrie, Taboada og Coddington (2007) har funnet at indre motivasjon, det å involvere seg i lesing for lesingens skyld, bidrar til leseforståelse. Mestringsforventninger i lesing handler om leserens tro på egne ferdigheter i lesing. Barn som skårer høyt på mestringsforventning presterer bedre på leseprøver enn barn som skårer lavt mestringsforventning (Mucherah & Yoder, 2008).

Guthrie og Wigfield med kolleger (Guthrie, Wigfield, Barbosa et al., 2004; Guthrie, Wigfield, & Perencevich, 2004) har utviklet og undersøkt effekten av et undervisningsopplegg, Concept-Oriented Reading Instruction (CORI), som har til hensikt å øke elevens engasjement og motivasjon i lesing. I CORI knyttes instruksjon i leseforståelse opp mot undervisning i ulike tema. Videre legges det vekt på å bruke interessante tekster, samarbeid i leseaktiviteter, konkrete og følbare erfaringer relatert til leseaktiviteter, og systematisk og eksplisitt instruksjon i leseforståelsesstrategier. Effekten av CORI er blitt undersøkt i en effektstudie der barn som ble undervist etter CORIs prinsipper ble sammenlignet med en eksperimentgruppe som fikk undervisning i de samme leseforståelsesstrategiene som inngår i CORI og en kontrollgruppe som fikk ordinær undervisning (Guthrie, Wigfield, Barbosa et al., 2004). Lærere i CORI-gruppen rapporterte høyere grad av motivasjon for lesing enn de to øvrige gruppene. Barn i CORI-gruppen fikk høyere skåre på flere mål for leseforståelse enn barn som kun hadde fått undervisning i leseforståelsesstrategiene og signifikant høyere enn barn som fikk tradisjonell undervisning. Selv om studien viste at CORI hadde en positiv effekt på leseforståelse, kan ikke effekten alene fortelle oss hvorvidt det er strategiene og de øvrige aktivitetene i CORI eller om det er barnas økte motivasjon og engasjement som har innflytelse på barnas leseforståelse. Wigfield, Guthrie, Perencevich, Taboada, Klauda, McRae og Barbosa (2008) har imidlertid undersøkt effekten av CORI ved å sammenligne framgangen til de to eksperimentgruppene kontrollert for motivasjon og engasjement. Analysene viste da ingen signifikante forskjeller mellom eksperimentgruppene og kontrollgruppen med hensyn til leseforståelse. Med andre ord kunne effekten på leseforståelse tilskrives barnas motivasjon og engasjement i lesing.

### **2.3.5 Sosiokulturell kontekst**

Lesing skjer ikke i et vakuum. Tekster skrives og leses i en sosiokulturell kontekst som har innvirkning på hvordan de utformes og tolkes. Den sosiokulturelle konteksten er ikke en enkelt variabel. Gaskins (2003) betrakter

sosiokulturell kontekst i forbindelse med leseforståelse som bestående av tre grupper variabler: Kognitive, sosiale og kulturelle.

Kognitive variabler innebærer for eksempel lærers kunnskap, instruksjon i leseundervisning og klasse- og skolekultur. Lærers kunnskap om lesing har konsekvenser for hvilke aktiviteter som foregår i klasserommet i leseundervisningen. PIRLS-undersøkelsen (Solheim et al., 2007) viste at lærere som engasjerer elevene i flere aktiviteter etter at de har lest en tekst, har elever med bedre leseferdigheter enn de lærerne som i mindre grad aktiviserer elevene etter lesing. Samme undersøkelse viste også at det er sammenheng mellom skolemiljø og leseferdigheter, både når det er rektor som rapporterer om skolemiljø og når det er elevene som gjør det. Sosiale variabler handler om økonomi og utdanning. PIRLS-undersøkelsen har også vist sammenhenger mellom familiens økonomi og prestasjoner i lesing blant norske elever. Materielle goder, som for eksempel det å ha datamaskin, eget rom, egne bøker og så videre, har sammenheng med leseferdigheter. Samme undersøkelse har også vist at det er sammenheng mellom foreldrenes utdanning og barnas prestasjoner, de elevene som hadde mødre med høy utdanning, skåret best på leseferdigheter. Den tredje gruppen av variabler, kulturelle variabler, innebærer ifølge Gaskins (2003) for eksempel språk og verdier. I tillegg er det naturlig å inkludere kulturell kapital i denne gruppen variabler. Antall bøker i hjemmet blir ofte brukt som indikator på kulturell kapital. Sammenhengen mellom antall bøker i hjemmet og leseferdigheter ble også bekreftet i PIRLS-undersøkelsen.

### **2.2.6 Teksteksponering**

Som det gikk fram i avsnitt 2.2.2, er lesing en kilde til å lære nye ord, som igjen har innflytelse på forståelsen. Det å lese gir også erfaring med å gjøre inferenser og erfaringer i forhold til hva slags informasjon vi kan forvente å finne i ulike typer tekster. Barn som leser mye, blir ofte gode lesere, noe som igjen fører til at de så leser enda mer. Eksponeringen for tekst er med andre ord kilde til matteuseffekt. De gode leserne blir enda bedre, mens gapet til de svake leserne stadig vokser. Det er gjort en rekke studier hvor en forsøker å finne ut hvilken rolle erfaring med tekst spiller for spriket mellom elevens ferdigheter i lesing. Blant annet har Stanovich (2000) funnet at elevenes kontakt med tekst utenfor skolen kan være en faktor som er med på å skape dette gapet mellom elevene. Stanovich har blant annet gjort undersøkelser som viser at forskjellen i elevenes skoleferdigheter vokser i løpet av sommerferien. Han forklarer dette med at elevene har ulik kontakt med tekst på fritiden.

At øvelse gjør mester, burde også gjelde for lesing. I hvilken grad vi eksponeres for tekst og hvor mye vi leser, burde ha betydning for vår leseforståelse. Likevel er ikke forskningen på sammenhengen mellom teksteksponering og leseforståelse entydig. Cipielewski og Stanovich (1992) har forsøkt å forutsi utvikling av leseferdigheter blant barneskoleelever ut fra eksponering for tekst. I undersøkelsen ble en tittelgjenkjenningssprøve og en forfattergjenkjenningssprøve brukt som mål på teksteksponering, mens leseforståelse ble målt med to standardiserte prøver. Gjenkjennelsesprøvene går ut på at informanten blir presentert for en rekke titler på bøker eller forfattere, hvorav noen titler eller forfatternavn er feil, konstruert av forskeren. Resultatene viste korrelasjoner på mellom .31 og .58 mellom eksponering for tekst og leseforståelse avhengig av hvilke mål som ble brukt. Cipielewski og Stanovich konkluderer med at teksteksponering bør anses som en konsekvens av leseutvikling samtidig som eksponering for tekst kan bidra til utvikling av leseferdigheter. Både i Conners' (2009) studie som det ble referert til i avsnitt 2.3.1, og en studie av Spear-Swerling (2006) ble tittelgjenkjenningssprøver brukt i tillegg til standardiserte prøver i leseforståelse. I Conners' studie korrelerte teksteksponering og leseforståelse .33, mens det ikke var korrelasjon mellom teksteksponering og leseforståelse målt i Spear-Swerlings studie av tredje- og fjerdeklassinger. Spear-Swerling mener en mulig forklaring på funnet kan være at tittelgjenkjenningssprøven hadde få item. Spear-Swerlings studie målte teksteksponering med også et selvrapporteringskjema. Det var heller ingen signifikant korrelasjon mellom skårer på selvrapporteringskjemaet og leseforståelse, ei var det et signifikant forhold mellom de to målene på teksteksponering. Teksteksponering er ikke uproblematisk å måle. Ved selvrapportering er det alltid fare for at informantene svarer det han eller hun antar at er sosialt akseptert. Dette problemet unngår man ved tittel- og forfattergjenkjenningssprøver. På en annen side krever gjenkjennelsesprøvene at forskeren som konstruerer prøvene har svært god oversikt over hvilke bøker og forfattere som er tilgjengelige for målgruppen. Oppsummert viser de ovennevnte studiene at teksteksponering noen ganger, og da i ulik grad, er relatert til leseforståelse.



### **3. Typer av inferens**

Inferens handler om kommunikasjon som ikke er direkte uttalt. I hverdagen gjør vi stadig inferenser på bakgrunn av både språk og handlinger. I lesing er inferens en følgeslutning med den hensikt å skape sammenheng i rekonstruksjonen av teksten. Leseren gjør på den ene siden inferenser for å redusere informasjon i teksten, for eksempel ved å utelate irrelevante detaljer eller ved å trekke ut hovedpoenger. På en annen side utdypes leseren informasjonen i teksten ved å legge til opplysninger til teksten ut ifra sin kunnskap om tekstens tema. Ulike teoretiske tilnærminger strides om hvilke inferenser leseren gjør under lesing. Graesser, Singer og Trabasso (1994) hevder at de inferenser som leseren gjør, nærmest er ubegrenset både hva omfang og kompleksitet angår. McKoon og Radcliff (1992) derimot mener at leserens inferenser under lesing er svært begrenset. I avsnitt 3.1 under vil det bli gjort rede for begge disse perspektivene på inferenser. Disse teoretiske perspektivene på inferenser sammen med Kintsch sin konstruksjon-integrasjonsmodell danner grunnlaget for en kategorisering av ulike typer inferenser i avsnitt 3.2. En slik kategorisering, og dermed begrepsavklaring, er nødvendig både for å kunne sammenligne studier av inferenstrening gjennomført av ulike forskere (se kapittel 5), og for å kunne redegjøre for hvilken inferenstype som er utgangspunktet for inferenstreningen i kapittel 7.

#### **3.1 Teoretiske perspektiver på inferens i lesing**

Graesser et al. (1994) støtter seg til konstruktivistisk teori for å presisere hvorfor leseren gjør inferenser når han eller hun leser. De hevder at forståelse av narrative tekster, hvilket de definerer som tekster hvis handling involverer for eksempel handlinger for å oppnå mål og emosjonelle reaksjoner til hendelser, er forankret i et prinsipp om leserens jakt etter mening for å danne en meningsfull situasjonsmodell. De forsøker å presisere hva prinsippet om leserens stadige driv mot mening innebærer, og hvilke konsekvenser prinsippet har for de inferenser som leseren gjør. Konstruktivistisk teori tar utgangspunkt i at leseforståelse innebærer konstruksjon av representasjoner og inferenser på flere nivåer: Både overfladisk forståelse, som innebærer å skape forbindelser mellom eksplisitte ord og setninger i teksten, og den dypere leseforståelsen som krever at leseren trekker slutninger om det globale eller helhetlige budskapet i teksten. Det konstruktivistiske perspektivet på hvordan inferenser genereres under lesing, inkluderer ifølge Graesser et al. både vel anerkjente komponenter, i tillegg til komponenter som er spesifikt knyttet til konstruktivistisk teori. De gir flere eksempler på det de mener er anerkjente komponenter i generering av inferenser, for eksempel ulike kilder til informasjon, teksten i seg, leserens kunnskaper om tekstens tema og tekstens



budskap. Minne, både kortidsminne, arbeidsminne og langtidsminne, inngår også i genereringen av inferenser. De komponentene som er spesielle for generering av inferenser i et konstruktivistisk perspektiv er for det første målet ved lesingen, at det er sammenheng mellom den mening som leseren skaper og målet for lesingen. Leserens kan ha ulike mål for lesingen avhengig av konteksten som teksten leses i. Målene kan være fra svært generelle for å skape et bilde av situasjonen i teksten, til mer presise og særegne for lesesituasjonen. Tekstens sjanger har også betydning for leserens mål, for eksempel om det er en underholdende tekst eller en informasjonstekst. Den andre komponenten som konstruktivistisk teori særlig støtter seg til, er leserens forsøk på å danne seg et bilde av teksten som skaper sammenheng både lokalt, for eksempel på setningsnivå, og globalt, på tvers av større deler av teksten. Den tredje komponenten i generering av inferenser er leserens forsøk på å forklare hendelsesforløp, for eksempel ved å skape sammenheng mellom personers handlinger eller motiver og hendelser.

Graesser et al. (1994) fokuserer på det de kaller kunnskapsbaserte inferenser. Kunnskapsbaserte inferenser dannes under lesing ved at langtidsmindet aktiviseres, og informasjonen kodes i representasjonen av teksten. Denne representasjonen i teksten inkluderer både tekstbasen og situasjonsmodellen av teksten. Videre skiller de mellom kunnskapsbaserte inferenser som blir produsert "online" og "offline". Online inferenser inkluderer inferens som er nødvendig for å skape sammenhenger på lokalt eller globalt nivå i teksten, mens offline inferenser i hovedsak ikke er nødvendig for å skape sammenheng i teksten, men inkluderer blant annet logiske resonnementer og matematiske formler. Inferens som skaper lokal koherens mellom ord i teksten eller mellom setninger, kan for eksempel være slutninger om hvem eller hva et pronomen viser tilbake til. Global inferens er slutninger på tvers av større deler av tekst, for eksempel budskap eller tematikken i teksten, eller inferenser om hovedpersonens motiver eller følelser. Videre hevder Graesser et al. at inferenser som kan karakteriseres som pragmatisk kommunikasjon mellom forfatter og leser, som for eksempel inferenser om forfatterens intensjoner eller følelser som vekkes hos leseren, også er online inferenser. Elaborerende inferens klassifiserer de som offline inferens, og eksemplifiserer denne inferenstypen ved at leseren gjerne gjør inferensen "egg og bacon" på bakgrunn av ordet "frokost" i teksten. Med andre ord inferenser hvor leseren legger til informasjon.

McKoon og Ratcliffs (1992) minimalisthypotese ("minimalist hypothesis") er alternativ teori til konstruktivismen om hvorfor og hvilke inferenser leseren gjør under lesing. I motsetning til Graesser et al. som hevder at leseren gjør en

rekke inferenser under lesing, hevder McKoon og Ratcliff at omfanget av inferenser som gjøres automatisk under lesing er svært begrenset. De hevder at det kun er inferenser som baserer seg enten på informasjon som er lett tilgjengelig eller som skaper lokal koherens, som leseren gjør automatisk. Med andre ord utelukker de at leseren gjør slutninger på tvers av større deler av tekst, det Graesser et al. kaller global inferens, under lesing. McKoon og Ratcliff understreker også at automatiske inferenser gjøres ubevisst og uavhengig av leserens mål. Videre hevder de at leseren ikke gjør automatiske inferenser for å skape mentale representasjoner av teksten. Ifølge McKoon og Ratcliff kan imidlertid lesing ofte ha som mål å bruke strategiske inferenser. De automatiske inferensene utgjør da en base, eller et utgangspunkt for de strategiske inferensene. Strategiene som anvendes er i motsetning til de automatiske inferensene bevisste og tas først i bruk når de lokale inferensene ikke er tilstrekkelige for å oppnå forståelse.

McKoon og Ratcliff (1992) definerer resultatet av inferenser i lesing som *"any piece of information that is not explicitly stated in a text"* (s. 440). Denne definisjonen er svært vid og omfatter tilsynelatende mange former for inferens, fra de helt enkle slutningene til avanserte resonnementer, og står således i kontrast til presiseringen av inferens som slutninger som kun er basert på enten på informasjon som er lett tilgjengelig eller som skaper lokal koherens. Keenan (1993) har kritisert McKoon og Ratcliff i det hun hevder at de biter seg selv i halen med definisjon om at inferens er basert på lett tilgjengelig informasjon. Hun påpeker at dersom man skal presisere hva som er lett tilgjengelig informasjon, så vil svaret være at lett tilgjengelig informasjon er informasjon som kan aktiveres når leseren gjør inferenser. Med andre ord er forholdet mellom inferenser og informasjonen som inferensene baserer seg på, sirkulært. Ifølge Keenan kan heller ingen empiriske studier bevise at andre inferenser enn de som inngår i minimalisthypotesen faktisk finnes. Dette fordi en inferens ifølge McKoon og Ratcliff enten er nødvendig for å skape lokal koherens, er basert på lett tilgjengelig informasjon eller ble trukket fordi leseren anvendte en strategi. Videre påpeker Keenan at tilhengerne av minimalisthypotesen alltid vil kunne hevde at inferenser er resultat av strategibruk.

### **3.2 Kategorisering av inferens**

I litteraturen om leseforståelse defineres og kategoriseres inferens på mange måter, noe som metaanalysen i kapittel 5 vil gi flere eksempler på. En studie av Bowyer-Crane og Snowling (2005) har vist at standardiserte prøver i leseforståelse tapper ulike typer inferenser. Bowyer-Crane og Snowling viste at konsekvensen var at enkelte barn ble utpekt som svake lesere på bakgrunn

av resultater i en av prøvene, men ikke i den andre. Dette understreker viktigheten av en begrepsavklaring i studier av inferens.

konstruktivistisk teori og minimalisthypotesen som ble presentert i avsnittet over, skiller seg særlig på tre punkter. For det første i forhold til i hvilken grad inferenser faktisk er nødvendige for forståelsen. For det andre i forhold til hvilken kunnskap som er involvert i inferensene, og for det tredje i hvilken grad leseren er bevisst de inferensene som gjøres. Dette avsnittet har til hensikt å definere ulike typer av inferens med utgangspunkt i kontrastene i de to teoriene. En taksonomi av inferenstyper blir oppsummert i figur 1 under. Kategoriseringen av inferenstyper som blir presentert her, er ingen definitiv inndeling. Inndelingen er snarere et verktøy for å kunne sammenligne studier av inferens i forbindelse med metaanalysen av tidligere studier av inferenstrening i kapittel 5, og for å presisere hvilken type inferens som trenes i effektstudien i kapittel 7.

Det første punktet som skiller de to teoriene om generering av inferens, er hvorvidt inferensene er nødvendige for forståelsen eller ikke. Mens McKoon og Ratcliff (1992) argumenterer for kun et begrenset omfang av inferenser som alle er nødvendige for å oppnå forståelse, hevder Graesser et al. (1994) at leseren også gjør inferenser under lesing som strengt tatt ikke er nødvendige for å oppnå en adekvat forståelse av teksten. Det første steget i taksonomien av inferenser som vil bli benyttet er derfor skillet mellom inferenser som er nødvendige for å forstå, og tilleggsinferenser. Tilleggsinferenser er slutninger som strengt tatt ikke er nødvendige for forståelsen, men som beriker den. Denne type inferens er i tråd med det Graesser et al. kaller offline inferenser, som innebærer en utdypning eller elaborering av innholdet i tekstene ut over det teksten eksplisitt uttrykker. Tilleggsinferenser kan for eksempel bero på personlig kunnskap erfaring knyttet til tekstens tema. Eksempelvis vil et barn som har mye kunnskap om dinosaurer, gjøre andre assosiasjoner til en tekst om dinosaurer enn et barn som ikke har tilsvarende kunnskap.

De nødvendige inferensene er absolutt påkrevd for å oppnå en adekvat forståelse. Nødvendige inferenser kan kategoriseres ytterligere i forhold til hvilken grad av og type kunnskap de krever.

Det andre punktet som skiller konstruktivistisk teori og minimalisthypotesen er hvilken rolle kunnskap spiller i inferenser. Sammen illustrer de to teoriene at det er ulik grad og typer av kunnskap som er involvert i det å gjøre inferenser. Som det ble gjort rede for i avsnitt 2.2.5, er det å integrere informasjonen i teksten med kunnskap fra utenfor teksten noe som skiller

elever med god og svak leseforståelse. Her kan Kintsch (1998) sin teori om konstruksjon–integrasjonsmodellen belyse hvilken kunnskap som er involvert i å gjøre inferenser. Det kreves annen kunnskap for å gjøre inferenser i tilknytning til tekstbasen enn til situasjonsmodellen. Kunnskap knyttet til tekstbasen vil være assosiasjoner som for eksempel å trekke slutningen om at kjøleskap har en dør i setningen *"Han åpnet døra og satte melken inn i kjøleskapet"*. Inferenser knyttet til situasjonsmodellen derimot, krever at leseren stiller opp et scenario av tekstens innhold, som i følgende setninger: *"Simen veltet melkekartongen. Far hentet en klut."* For å forstå hvorfor far hentet en klut, må leseren gjøre inferensen om at den veltede melkekartongen skapte søl.

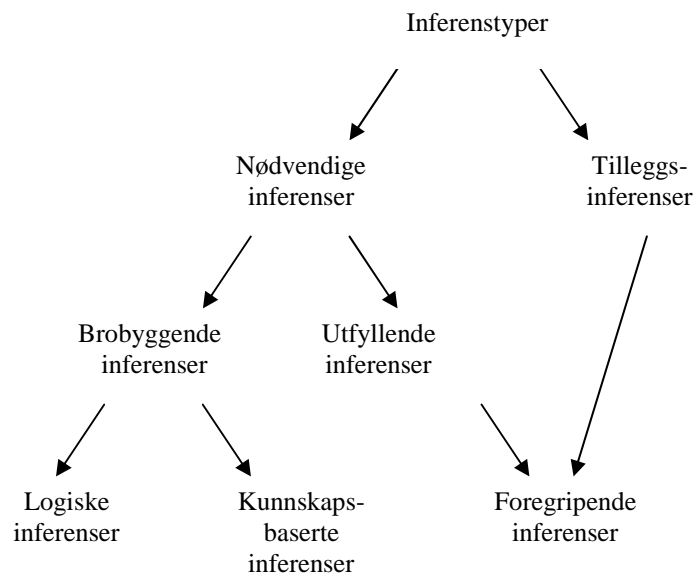
Blant de nødvendige inferensene kan man skille mellom inferenser knyttet til tekstbase og inferenser knyttet til dannelsen av situasjonsmodeller. Førstnevnte inferenstype vil heretter bli referert til som brobyggende inferens, mens sistnevnte inferenstype vil bli referert til som utfyllende inferens.

De brobyggende inferensene er viktige for å binde sammen ulike informasjonen som er eksplisitt uttrykt i teksten, og for å skape koherens mellom setninger og innad i setninger. Brobyggende inferens kan være både logiske, det vil si at de er utledet med utgangspunkt i premisser som er eksplisitt uttrykt i teksten, og kunnskapsbaserte, det vil si at leseren integrerer egen kunnskap for å skape koherens på lokalt nivå i teksten. Følgende eksempel illustrerer dette: *"Alexander Kielland ble født i Stavanger. Det er Norges fjerde største by."* En logisk, brobyggende inferens basert på eksempelet vil være at Alexander Kielland ble født i Norge. Inferensen om at "det" viser tilbake på Stavanger, kan være en logisk inferens, en resonnering på bakgrunn av teksten. Samtidig kan denne inferensen betraktes som kunnskapsbasert fordi man faktisk trenger kunnskap om tilbakevisende pronomener for å kunne gjøre denne inferensen.

De utfyllende inferensene krever at leseren stiller opp et scenario eller en situasjonsmodell på bakgrunn av teksten, noe som følgende eksempel illustrerer: *"Ole Jensen fikk ikke låne penger i banken til å kjøpe seg båt. Han begynte å se seg om etter en ekstrajobb."* For å skape sammenheng mellom de to setningene må leseren stille opp et scenario som innebærer at en ekstrajobb ville gi Ole Jensen muligheten til tjene pengene han trengte til å kjøpe båt siden han ikke fikk lån i banken.

Foregripelser er inferenser om hva som skjer videre i teksten ut over det som eksplisitt står i teksten. Foregripende inferenser kan være både nødvendige,

utfyllende inferenser og tilleggsinferenser. ”Demonstrantene brøt seg inn i ambassaden. Sikkerhetsvakten ringte politiet.” Slutninger om at sikkerhetsvakten ringte politiet for å opprettholde ro og orden, vil være en utfyllende foregripelse. Foregripende inferenser om at ambassaden i fremtiden kommer til å gjennomgå sine sikkerhetsrutiner, er derimot en tilleggsinferens som ikke er strengt tatt nødvendig for å forstå teksten, men som likevel er relevant.



Figur 1. Inferenstyper

Som det er beskrevet over, bidrar en del inferenser med å skape indre bilder av tekstens innhold. På en annen side kan de indre bildene gjøre det enklere å gjøre inferenser. Ifølge Kintsch (1998) sin konstruksjon–integrasjonsmodell er det et resiprokt forhold mellom tekstbase og situasjonsmodell, jevnfør avsnitt 2.1. De brobyggende inferensene på tekstbasenivå er en forutsetning for å kunne gjøre inferenser knyttet til situasjonsmodellen, samtidig som det kan bli enklere for leseren å gjøre inferenser på lokalt nivå når han eller hun har etablert en situasjonsmodell.

Når det gjelder det tredje punktet som skiller de to teoriene om generering av inferens, argumenterer McKoon og Ratcliff (1992) for automatiske inferenser, mens Graesser et al. (1994) snakker om online og offline inferenser. Selv om disse begrepene ikke er jamstilte, kan man snakke om ulik grad av bevissthet

og kontroll involvert i det å gjøre inferenser. Her kan igjen Kintsch sin teori belyse forholdet mellom bevissthet eller automatikk og kontroll. Kintsch (1993, 1998) hevder at inferenser varierer fra å være automatiske til kontrollerte prosesser. Inferens på grunnlag av automatiske prosesser innebærer at leseren umiddelbart henter opp informasjon fra langtidsminet ut fra stikkord i teksten. Ved inferens på bakgrunn av kontrollerte prosesser hevder Kintsch at stikkordene i teksten ikke er fullstendige nok til at de kan knyttes direkte opp mot langtidsminet. Stikkordene i teksten må suppleres med et kontrollert søk etter sammenhenger mellom to eller flere utsagn i teksten eller ulike tekstdeler. Dersom man betrakter prosessene som inngår i Kintsch sin teori om genereringen av inferens som individuelle, vil det være rimelig å anta at hvilke slutninger leseren trekker, og om slutningene krever automatiske eller kontrollerte prosesser, varierer fra leser til leser. Hvorvidt letingen eller søket etter sammenhenger mellom utsagn eller bruddstykker av tekst skjer rutinemessig, vil da også være individuelt. I hvilken grad letingen etter sammenhenger i teksten skjer automatisk eller kontrollert, vil også være avhengig av tidligere erfaringer og hvilken kunnskap en har.



## **4. Argument og motivasjon for effektstudier av inferenstrening**

Pedagogisk forskning er nødvendig for å finne faktorer som har innflytelse på barnets læring og for å vinne kunnskap om hvordan man kan endre disse faktorene til barnets fordel. Pedagogisk praksis bør være vitenskaplig forankret, derfor bør dens effekt undersøkes empirisk. Denne avhandlingen har til hensikt å undersøke effekter av å trene inferens for å gjøre deltakerne i bedre stand til å gjøre inferenser og for å fremme deres leseforståelse. Med andre ord fokuseres det på effektstudier av inferenstrening. I dette kapitlet gjøres det rede for hvilke og hvordan effekter av inferenstrening kan undersøkes gjennom effektstudier. Videre vil validitetsmessige og etiske problemstillinger knyttet til effektstudier bli diskutert.

I effektstudier av inferens vil inferenstreningen være den uavhengige variabelen, og dermed den antatte årsaksfaktoren. Den uavhengige variabelen kan betraktes fra to sider, teoretisk og pedagogisk. Motivasjonen for å undersøke forholdet mellom inferens og leseforståelse er i denne avhandlingen knyttet til både det teoretiske og det pedagogiske. Fra et teoretisk perspektiv undersøkes det om ferdigheter i å gjøre inferenser er årsaksforbundet med leseforståelse. Fra et pedagogisk perspektiv undersøkes det om inferenstrening har en praktisk betydning for elevenes leseforståelse. Fra et pedagogisk perspektiv bør forskningsspørsmålene være fruktbare i forhold til om kunnskapen som søkes har nytte for praksis. Dersom man undersøker effekten av inferenstrening, er det relevant å spørre seg om inferenstreningen lar seg gjennomføre innen rammene for vanlig skole. Dersom treningen er svært ressurskrevende, for eksempel at lærertettheten må være svært høy, vil det bli vanskelig å endre den daglige praksis, til tross for viten om at undervisningsformen er gunstig.

Som det ble gjort rede for i avsnitt 2.2.1, kan årsaksretninger mellom ferdigheter av betydning for leseforståelse og den helhetlige leseforståelsen kartlegges også med andre design enn effektstudier, som CAM-design og longitudinelle studier. Disse to designene er imidlertid ikke like godt egnet til å ivareta det pedagogiske perspektivet på forholdet mellom inferens og leseforståelse, sammenlignet med effektstudien. Gjennom effektstudien kan innholdet, i dette tilfellet inferenstreningen, tilpasses de rammer som skolen har. Det praktiske anvendelsespotensialet er dermed et argument for pedagogiske effektstudier.



#### **4.1 Måling av effekt av inferenstrening**

Effekter av inferenstrening kan og bør kartlegges på flere nivåer. Effektene kan deles inn i tre hovedkategorier: Primæreffekter, generaliserte effekter og utsatte effekter. Primæreffekter innebærer effekter av det som det er trent på (den uavhengige variabelen). Dersom man trener deltakerne i å gjøre inferenser, vil den primære effekten kunne måles i deltakernes eventuelle forbedring i å gjøre inferenser. Fra et teoretisk perspektiv er primæreffekter av inferenstreningen, det vil si at treningen faktisk bidrar til at elevene gjør inferenser, avgjørende. Dette fordi inferenstreningen er en operasjonalisering av det teoretiske begrepet inferens. Uten informasjon om primæreffekter, eller dersom primæreffektene uteblir, kan man heller ikke tolke eventuelle generaliserte effekter på for eksempel leseforståelse. Uten primæreffekter kan man ikke bekrefte den teoretiske sammenhengen mellom inferens og leseforståelse. Primæreffekter kan måles på to måter, enten ved å måle effekt med materiell brukt i treningen, eller ved å måle effekt på tilsvarende materiell eller tekster. Sistnevnte innebærer en overføring av effekt og vil heretter bli referert til som primær overført effekt.

Generaliserte effekter innebærer effekter av inferenstreningen målt på andre områder enn de trenede. Den mest nærliggende generaliseringseffekten av inferenstrening er leseforståelse. Det er også mulig å tenke seg generalisering av inferenstrening til for eksempel en generell forbedring av prestasjoner i skolefag hvor lesing er kilde til kunnskap. På en annen side bør generaliseringseffekter være knyttet til lesing spesielt og ikke til skoleferdigheter generelt. Dette fordi en generell, vid effekt av inferenstreningen kan være resultat at en forventningseffekt, eller såkalt Hawthorneeffekt.

Den siste kategorien, utsatte effekter, er et mål på hvorvidt effekter opprettholdes og om det er ytterligere effekt en tid etter at inferenstreningen er avsluttet. Utsatte effekter kan måles både i forhold til primæreffekter og generaliseringseffekter.

I tillegg til å kartlegge effekt med mål på ulike nivåer kan man måle effekt uavhengig av nivå på flere måter. Effekt kan måles ved å sammenligne deltakernes pretestskårer med posttestskårer. Ved å måle effekt på denne måten ser man kun på framgangen til deltakerne som har mottatt påvirkning, man vet ingenting om deltakerne uansett hadde hatt framgang dersom de ikke hadde mottatt påvirkningen. I trening av ferdigheter av betydning for leseforståelse blant skoleelever vil man med et slikt design ikke ta høyde for den naturlige framgangen som elevene uansett vil kunne ha hatt uavhengig av treningen. Man kan ta høyde for den naturlige framgangen ved å inkludere en

kontrollgruppe. Kontrollgruppen kan brukes som sammenligningsgrunnlag for eksperimentgruppen på to måter, enten ved å sammenligne eksperimentgruppens posttestskårer med posttestskårene til en kontrollgruppe, eller ved å sammenligne eksperimentgruppens og kontrollgruppens endring fra pretest til posttest. Den sistnevnte tilnærmingen er å foretrekke siden den tar hensyn til deltakernes ferdigheter før treningen. Kontrollgruppen kan være helt uten påvirkning, ved at den fortsetter sin ordinære undervisning. Kontrollgruppen kan også motta en alternativ påvirkning til den påvirkning eksperimentgruppen får.

Effekt av inferenstrening kan måles ved hjelp av både eksperimentdesignede og ikke-eksperimentdesignede tester. Førstnevnte type test er utviklet av forskerne som designer studien og innholdet i inferenstreningen. Ikke-eksperimentdesignede tester kan være standardiserte prøver eller andre typer tester som forskerne selv ikke har designet spesielt for studien. De ikke-eksperimentdesignede testene blir således en ytre validering. Eksperimentdesignede tester er gunstig for å måle primær- og primær overført effekt fordi dette er mål på det som elevene har øvd på i inferenstreningen. Primær- og primær overført kan vanskelig måles med tester som er designet for andre formål. Generaliserte effekter av inferenstreningen derimot kan med fordel måles ved ikke-eksperimentdesignede tester. Dette fordi disse testene har som mål å undersøke effekter på andre områder enn det trenede.

I en sammenfatning av effektstudier av resiprok undervisning fant Rosenshine og Meister (1994) stor forskjell i effekter på ikke-eksperimentdesignede (standardiserte) tester versus eksperimentdesignede tester. Den gjennomsnittlige effektstørrelsen på ikke-eksperimentdesignede tester kun var .32, hvilket ifølge Cohens (1988) retningslinjer for tolkning av effektstørrelser tilsvarer en liten til medium effekt. To av de ni studiene som den gjennomsnittlige effektstørrelsen for ikke-eksperimentdesignede tester ble beregnet på, var ikke signifikante. Den gjennomsnittlige effektstørrelsen for de eksperimentdesignede testene var betraktelig sterkere, hele .88. Seks av de sju studiene som brukte eksperimentdesignede tester, hadde signifikante resultater.

## **4.2 Eksperimentelle design**

Et fellestrekk ved effektstudier er at man søker svar på problemstillinger knyttet til årsak og virkning, i tillegg blir deltakerne utsatt for en påvirkning. Det er mulig å undersøke forholdet mellom årsak og virkning med ulike varianter av eksperimentelle design (for en oversikt se Shadish, Cook, & Campbell, 2002). Et hovedskille går mellom ekte eksperiment og

kvasieksperiment. De ekte eksperimentene er i motsetning til kvasieksperimentene randomiserte. Manglende randomisering er en klar begrensning ved kvasieksperimentet. Et annet skille mellom de to typene eksperimenter er graden av kontroll. Mens de ekte eksperimentene skjer under svært godt kontrollerte forhold, er graden av kontroll mindre i kvasieksperimentet. Tross disse to begrensningene er kvasieksperimentet et svært vanlig design innenfor den pedagogiske forskningsdisiplinen. Dette fordi det er svært vanskelig å gjennomføre ekte eksperiment med randomiserte utvalg med hensyn til det praktiske. Videre vil studier av mennesker bli preget av mindre kontroll enn ved ekte eksperiment. Kvasieksperimentelle studier kan gjennomføres både med og uten pretester, og med og uten kontrollgruppe, og med kontrollgruppe med eller uten påvirkning. I kvasieksperiment hvor påvirkningen er en form for undervisning, kan man tenke seg at enten både lærer og elever er informert om at de deltar i et eksperiment, at kun lærer er klar over deltakelse, eller at verken elever eller lærer er klar over deltakelse. Validitetsmessige overveielser er nødvendig uansett hvilken type kvasieksperimentelt design som benyttes. Derfor vil validitetsmessig problemstillinger bli diskutert i det følgende.

#### **4.2.1 Validitetsmessige problemstillinger**

Validitet handler om sikkerheten i de slutningene som trekkes i forbindelse med et forskningsspørsmål. Shadish et al. (2002) har utviklet et validitetssystem for å vurdere sikkerheten i kausale undersøkelser. I deres tradisjon betraktes validitet som den antatte sannheten til en slutning (s. 34). Med andre ord forteller validitet oss noe om styrken på de bevis som støtter en slutning eller påstand. Shadish et al. sitt validitetssystem omfatter fire typer validitet: indre validitet, ytre validitet, statistisk validitet og begrepsvaliditet. Under diskuteres de ulike validitetstyper i forhold til studier av inferenstrening med et kvasieksperimentelt design.

##### **4.2.1.1 Statistisk validitet**

Statistisk validitet handler om hvorvidt den målte effekten eller sammenhengen skyldes tilfeldigheter eller er reell. Denne typen validitet berører to typer statistiske slutninger, *om* den antatte årsaken og effekten kovarierer, og *hvor sterkt* årsak og effekt kovarierer (Shadish et al., 2002).

Det er to typer feilslutninger som kan oppstå i forbindelse med hypotesetesting. Den ene kan være at effekten i virkeligheten ikke skyldes treningen, men at det likevel konkluderes med at treningen har effekt. Dette kalles type I feil. Type II feil er når det konkluderes med ingen effekt, når

sannheten er at treningen faktisk har effekt. Sannsynligheten for å begå type I feil beregnes ut fra signifikansnivå. Signifikansnivå er et uttrykk for hvor stor sjansen er for at vi tar feil når vi trekker en slutning. Jo strengere signifikansnivå man velger, desto mindre blir sjansen for å konkludere med effekt når effekten i realiteten ikke er til stede. Samtidig som sjansen for å begå type I feil synker når vi velger et strengt signifikansnivå, øker sjansen for å gjøre type II feil. Det er med andre ord større sjanse for at det faktisk er en effekt selv om vi konkluderer med ingen effekt. I tillegg kan muligheten for å gjøre type I feil ikke utelukkes ved å velge et strengt signifikansnivå. Fordi kontrollen i kvasiekperimentet er begrenset, kan effekten av uavhengig variabel påvirkes av en eller flere tredjevariabler. En vet med andre ord ikke om effekten av uavhengig variabel på avhengig variabel er direkte. Skyldes effekten på inferens og leseforståelse (avhengige variabler) inferenstreningen alene (uavhengig variabel), eller skyldes effekten helt eller delvis en eller flere tredjevariabler?

Tredjevariabler kan påvirke effekten på flere måter. Ved mediering produserer en eller flere tredjevariabler effekt på avhengig variabel gjennom uavhengig variabel (Baron & Kenny, 1986). Wigfield et al. (2008) fant at økt leseforståelse i etterkant av leseforståelsestrening (jevnfør avsnitt 2.3.4) ble mediert av elevenes grad av motivasjon. Dette er et aktuelt problem i kvasiekperimentelle studier av inferens, da det er stor sjanse for at en eventuell effekt av inferenstreningen kan bli påvirket av medierende variabler. Motivasjon kan være en tredjevariabel som medierer effekt ved at inferenstreningen øker elevenes motivasjon, noe som kan virke inn på elevenes prestasjoner på posttestene. Samtidig kan en motivasjonsfremmende inferenstrening også ha en mer direkte effekt på resultatene på posttestene. Dette betyr at effekten av treningen reduseres dersom man kontrollerer for effekten motivasjon har på posttestene. Tredjevariabler kan også fungere som moderator, hvor effekten blir forstyrret på en annen måte enn ved mediering. Moderatorvariabler deler uavhengig variabel inn i undergrupper. Disse undergruppene gir effekt på avhengig variabel (Baron & Kenny, 1986).

Den målte effekten på posttesten kan også påvirkes av supressorvariabler, det vil si variabler som undertrykker effekt. Dersom inferenstreningen tilpasses gjennomsnittseleven, kan inferenstreningen ha negativ eller undertrykkende effekt på elever som befinner seg på ytterpunktene i forhold til leseforståelse. Det er sannsynlig at de svakeste elevene presterer svakere enn de vanligvis gjør utenfor eksperimentsituasjonen fordi de ikke har kognitive ferdigheter til å ta i bruk strategier som det undervises i treningen og at de da mislykkes i å bruke mer krevende strategier enn de vanligvis gjør. Likeledes kan sterke

elever med godt etablerte strategier skåre svakere enn de ellers ville når de blir oppfordret til å bruke strategier eller metoder som er mindre avanserte enn de vanligvis anvender.

#### **4.2.1.2 Indre validitet**

Indre validitet handler om sikkerhet i kausale slutninger, i dette tilfellet om tendensen som måles skyldes inferenstreningen. Kontrollerte forhold gjør at man kan trekke kausale slutninger med stor sikkerhet, noe som gjør at indre validitet i eksperimentelle studier, også i kvasiekperimentet, er sterk. En fare ved eksperimentet er imidlertid at situasjonen som undersøkes kan, nettopp på grunn av høy grad av kontroll, framstå som unaturlig og kunstig. Dette kan ha konsekvenser for ytre validitet, da eksperimentsituasjonen kan bli lite representativ for den populasjonen som forskeren ønsker å generalisere til (Lund, 2002). For å sikre at sammenhenger mellom variablene virkelig er kausale, bør en undersøke om sammenhenger skyldes systematiske faktorer som utgjør en trussel mot indre validitet.

Grunnen til at eksperimentelle design brukes i evaluering av undervisningsmetoder, er at man unngår retningsproblemet ved å manipulere årsaksvariabelen. Med andre ord er det ingen tvil om hva som er årsak og virkning, i og med at treningen (årsaken) er manipulert til å inntreffe før den eventuelle endringen i leseforståelse (virkningen). Dette gjør at den indre validiteten ved kvasiekperimentelle studier er sterk.

Selv om den indre validiteten i kvasiekperimentelle studier med pre- og posttest generelt er sterk, utgjør modning og historie en trussel mot den indre validiteten. Historie innebærer at en hendelse uten tilknytning til undersøkelsen skaper effekten som blir målt i eksperimentet. Dersom en eller flere hendelser oppstår mellom pretest og posttest og berører enten kontrollgruppen eller eksperimentgruppen, svekker dette den indre validiteten. Modning vil si at forsøkspersonene forandrer seg på grunn av biologiske forhold eller forhold knyttet til miljø. Dersom man inkluderer en kontrollgruppe i studien, vil imidlertid barnas eventuelle vekstperiode slå likt ut på både kontrollgruppe og eksperimentgruppe, og differansen mellom gruppene vil således ikke bli berørt. Likeledes vil en hendelse uten tilknytning til undersøkelsen slå likt ut i begge grupper. Etter som faren for modning og historie øker etter hvor forskjellige de to gruppene er, vil det være usikkert hvorvidt eventuelle effekter i mer eller mindre grad skyldes forhold som ikke kan tilskrives eksperimentet, dersom deltakere i eksperimentgruppe og kontrollgruppe ikke er fordelt etter tilfeldighet. Det kan derfor være

problematisk å trekke kausale slutninger dersom eksperimentgruppe og kontrollgruppe er systematisk forskjellige i forkant av inferenstreningen.

I effektstudier er det fare for at eventuelle effekter av påvirkningen delvis eller helt skyldes forventningseffekter, eller Hawthorneeffekt. Hawthorneeffekt innebærer at det å være gjenstand for en studie i seg selv fremkaller effekter. Det at deltakerne i eksperimentgrupper vet at de får en påvirkning, kan påvirke dem til å opptre annerledes enn de vanligvis gjør utenfor eksperiment-situasjonen, for eksempel ved at de enten yter mer fordi de vet at de er med på en undersøkelse, eller ved at de boikotter undersøkelsen. Deltakere i kontrollgrupper kan også oppføre seg annerledes fordi de vet at de sammenlignes med en eksperimentgruppe som får et tiltak som antas å være til eksperimentgruppens fordel. Dette kan påvirke kontrollgruppen til enten å yte mer enn vanlig på tester og oppgaver, eller motsatt, at deltakerne er lite motivert til yte i den grad de ellers ville ha gjort fordi de vet de går glipp av et tiltak som antas å være fordelaktig.

Ved pre- posttestdesign er det en fare for at førstegangs testing kan virke inn på respondentenes besvarelse ved senere testing. Forandringen mellom første og annen gangs måling blir da kunstig, fordi deltakerne har lært noe av pretesten. En slik retesteffekt vil utgjøre en trussel når en opererer med både pre- og posttest (Lund, 2002). I hvilken grad retesteffekt utgjør en trussel mot indre validitet, har sammenheng med hvor lang tid det går mellom pre- og posttest.

#### **4.2.1.3 Begrepsvaliditet**

Begrepsvaliditet handler om slutninger fra empiri til teoretiske begrep, hvilket begrep som måles og hvor godt dette begrepet måles. I forbindelse med vurdering av inferenstrening vil det være viktig å sikre at inferenstreningen slik den blir anvendt i eksperimentet, ligger tett opp mot det teoretiske begrepet (Vellutino & Schatschneider, 2004). Med andre ord er det viktig å undersøke i hvilken grad treningen er implementert for å vurdere om begrepsvaliditeten er ivaretatt.

Implementeringen påvirkes blant annet av hvordan treningen blir administrert. Ved trening blant elever vil det ofte være aktuelt at lærere administrerer treningen. Selv om alle lærerne i prinsippet skal gjennomføre den samme treningen, blir, ifølge Levin (2005), i realiteten hver lærers klasse en egen variant av trening. Dette fordi hver enkelt klasse har ulike betingelser fysisk (for eksempel klasserom og ventilasjon) og organisatorisk (for eksempel hvilken tid på dagen elevene blir utsatt for treningen). I tillegg vil treningen

være preget av læreren som person (undervisningsstil, personlighet, kjønn og så videre).

En annen problematisk side ved implementering er at det kan være vanskelig å manipulere en enkelt komponent ved undervisningen og samtidig holde de andre sidene ved undervisningen konstant (Mayer, 2005). Gjennomføring av inferenstrening vil kunne virke inn på betingelser for læringssituasjonen, som for eksempel hvordan øvrig undervisning organiseres.

#### **4.2.1.4 Ytre validitet**

Ytre validitet handler om hvorvidt en kan generalisere til andre persongrupper og situasjoner. Dette handler om hvor lik undersøkelsen er den virkelighet vi ønsker å trekke slutninger til. Lund (2002) understreker at det derfor er viktig å velge personer og situasjoner til undersøkelsen som er mest mulig lik målet for generalisering.

Hvilke prinsipper som brukes for å finne et utvalg, har først og fremst konsekvenser for generalisering til personer. Utvalget bør ideelt sett speile populasjonen, men praktiske forhold begrenser hvilke prinsipper som kan brukes for å trekke et representativt utvalg. Som tidligere nevnt lar det seg vanskelig gjøre å trekke et representativt utvalg til effektstudier i skolen. Manglende randomisering gjør det problematisk i forhold til å generalisere til persongrupper fordi en i mindre grad får kunnskap om hva som ville ha skjedd dersom de samme deltakerne ikke hadde fått påvirkningen, fordi kontrollgruppen er systematisk forskjellig fra kontrollgruppen.

I forhold til generalisering til situasjoner kan kvaseksperimentet by på problemer på grunn av sin form. Fordi eksperimentsituasjonen er kontrollert, kan den virke kunstig i forhold til virkeligheten. Dersom kontrollen blir for sterk (jvnfør avsnitt 4.2.1.2 om indre validitet), kan det bli vanskelig å generalisere eksperimentsituasjonen til det virkelige liv.

#### **4.2.1.5 Hvordan imøtekomme validitetsmessige problemstillinger?**

En vil aldri kunne garantere seg hundre prosent mot tredjevariabler. Litteraturstudier kan imidlertid bidra til å kartlegge aktuelle tredjevariabler, mediator og moderatoreffekter slik at man kan inkludere dem som kontrollvariabler og dermed styrke den statistiske validiteten.

Indre validitet er sterk i kvaseksperimentelle studier fordi årsaksvariabelen manipuleres. Kontrollen i denne sammenhengen innebærer at elevenes undervisning blir spesielt tilrettelagt. Selv om kontroll er nødvendig for å

ivareta indre validitet, kan for sterk kontroll på eksperimentsituasjonen gjøre generaliseringen vanskelig. Graden av kontroll i eksperimentsituasjonen må derfor balanseres i forholdet indre og ytre validitet. For at inferenstreningen ikke skal oppleves som kunstig for deltakerne, hvilket kan ha innflytelse på ytre validitet, bør treningsmateriellet tilpasses slik at det kan inkluderes i det ordinære pensumet og i den daglige undervisningen. For å redusere sannsynligheten for truslene modning og historie mot den indre validiteten, bør kontrollgruppe inkluderes. Videre kan sannsynligheten for forskjeller mellom gruppene reduseres ved at deltakelse i eksperimentgruppe og kontrollgruppe blir avgjort ved loddtrekning. I tillegg bør tilstrekkelig mange kontrollvariabler inkluderes i studien for å avkrefte eller bekrefte forskjeller mellom gruppene.

Utfordringer knyttet til begrepsvaliditet handler i denne sammenhengen om implementeringen av inferenstreningen, hvilket er avgjørende for hvorvidt man med sikkerhet kan slutte at den effekt som man eventuelt finner, faktisk er resultat av årsaksvariabelen. For å kontrollere implementeringen kan man observere inferenstreningen i praksis. Videre kan også implementeringen av inferenstreningen kartlegges ved at lærerne som gjennomfører treningen fører logg i forbindelse med gjennomføringen. Begrepsvaliditeten kan også bli kontrollert gjennom posttestene. For å sikre at det er det trenede som måles, bør innholdet i en av testene, mål på primæreffekt eller primær overført effekt, ligge tett opp innholdet i inferenstreningen.

Den ytre validiteten vil alltid være et svakt punkt i kvasiexperimenter fordi utvalget ikke er tilfeldig valgt. Selv om det praktiske setter begrensninger for tilfeldig utvelging, kan man likevel tilstrebe og rekruttere skoler som er mest mulig representative med tanke for eksempel innbyggernes økonomi og geografisk beliggenhet.

### **4.3 Ethiske problemstillinger ved pedagogiske effektstudier**

Eksperimentet har vokst fram fra naturvitenskapen. Samfunnsvitenskapen skiller seg fra naturvitenskapen ved at det er mennesker og ikke ting eller objekter som studeres. Eksperimentet må derfor tilpasses før det overføres til et forskningsfelt hvor menneskets handlinger preget av vilje, intensjoner og motiv er målet for forskning. Forskning på barn og unge krever at det tas spesielle hensyn i forhold til etikk. Forskningsetiske retningslinjer krever at forskeren tilpasser metoden i undersøkelsen til den aldersgruppen som deltar i forskning (Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora & Kalleberg, 2006).



I det følgende diskuteres etiske problemstillinger ved pedagogiske effektstudier av inferenstrening. Særlig aktuelt i forhold til studier av skoleelever med et kvasiekperimentelt design med kontrollgruppe uten påvirkning er problemstillinger knyttet til bruk av kontrollgruppe og informert samtykke.

#### **4.3.1 Etikk og bruk av kontrollgruppe**

I flertallet av effektstudier av inferenstrening inkluderes en kontrollgruppe som sammenligning for eksperimentgruppens eventuelle framgang. Å inkludere en kontrollgruppe i en studie av inferenstrening er i seg selv ikke etisk problematisk, så lenge elevene i kontrollgruppen mottar den undervisningen de ville ha fått utenfor studien.

Den etiske forpliktelsen til å gi elever og lærere tilstrekkelig og riktig informasjon om studien de deltar i, samt deres rolle i studien, kan stå i konflikt med validitetsmessige hensyn. Det at eksperimentgruppe eller kontrollgruppe oppfører seg atypisk enten fordi deltakerne vet at de er med i et undervisningsopplegg som er antatt gunstig, eller fordi de vet at de sammenlignes med en gruppe som får et antatt gunstig undervisningsopplegg, kan virke inn på deltakernes prestasjoner. I effektstudier av inferenstrening vil dette fenomenet, Hawthorneeffekt, kunne forekomme blant lærere så vel som elever. Atypisk atferd kan ha konsekvenser for resultatene, og utgjør en trussel mot sikkerheten i de slutninger som trekkes på bakgrunn av dataene.

Dersom en form for pedagogisk praksis er antatt å ha positiv effekt, kan det synes umoralsk å bruke en kontrollgruppe i undersøkelsen som ikke nyter godt av en trening som man tror kan gi positivt utbytte. Det er imidlertid ikke selvsagt at enhver ny undervisningsmetode har positiv effekt – det er nettopp det effektstudier av inferenstrening søker svar på. Konklusjonen av effektstudier er ikke gitt på forhånd. Like fullt vil det ofte være tilfelle at mange ønsker å delta i den gruppen som mottar en pedagogisk praksis som er antatt å ha positiv effekt, derfor hever Reyna (2004) at hvilken gruppe som skal få påvirkning og hvilken som skal fungere som kontroll, av moralske hensyn bør avgjøres ved tilfeldig valg.

#### **4.3.2 Informert samtykke**

Ifølge forskningsetiske retningslinjer bør foreldrene gi samtykke til at barna deltar i forskning når barnet er under 15 år (Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora & Kalleberg, 2006). Ideelt bør elever og foreldre, uten press utenifra, avgjøre om de vil delta i forskning etter å ha fått tilstrekkelig informasjon om prosjektet og dets hensikt. I praksis kan

dette idealet være vanskelig å opprettholde. I noen situasjoner vil deltakerne i større eller mindre grad føle seg presset til å delta i forskning. I forskning hvor barn og ungdom er deltakere, bør man særlig være bevisst hvilken innflytelse autoritetspersoner kan ha på barnas deltakelse. I skolesammenheng vil lærer eller rektor kunne anbefale eller ytre at det er et sterkt ønske om deltakelse. I effektstudier som omfatter hele klasser, vil elever og foreldre også kunne føle seg presset til å delta fordi det kan berøre hele skoleklasser dersom noen reserverer seg fra å delta. Det kan eksempelvis skape press på enkeltelever eller foreldre dersom påvirkningen planlegges gjennomført i en hel klasse.

### **4.3.3 Hvordan imøtekomme etiske problemstillinger?**

Etiske hensyn må tas både i forhold til valg av måleinstrumenter og måten data hentes inn på. Prøver og spørreskjema som benyttes i forskning på mennesker generelt, og barn spesielt, må være tilpasset alder og forutsetninger. Videre bør data hentes inn i deltakernes vante omgivelser. For elever vil klasserommet være den naturlige arena. Å studere mennesket gjennom et kvasiekseptimentelt design innebærer ikke nødvendigvis å redusere mennesket til et objekt dersom man tar menneskelige hensyn og moderasjoner. Konklusjonene bør deretter ikke være absolutte, men modererte, hvilket også må prege begrepsbruken i rapportering fra effektstudier av inferenstrening.

Deltakelse i eksperimentgruppe og kontrollgruppe bør avgjøres ved loddrekning fordi det er sannsynlig at enkelte deltakere har preferanser i forhold til om de ønsker å delta i eksperimentgruppe eller kontrollgruppe. Dersom man derimot *har* empiriske holdepunkter for at inferenstreningen er gunstig, må man forholde seg denne kunnskapen i forhold til kontrollgruppen. Den mest innlysende måten å behandle slik kunnskap på vil være å gi kontrollgruppen det samme materiell og den samme metode som eksperimentgruppen etter at studien er gjennomført.

Forpliktelsen til å informere deltakerne om eksperimentet kan ikke forbigås av hensyn til at informasjonen til deltakerne kan føre til Hawthorneeffekter. Det kan kontrolleres for Hawthorneeffekter ved for eksempel å teste deltakerne på andre områder enn det trenede. Dersom deltakerne uavhengig av gruppe ikke har en generell forbedring på områder som ikke angår det trenede, er dette en indikasjon på fravær av Hawthorneeffekter.

I effektstudier hvor data hentes inn på flere tidspunkter, er det særlig viktig at potensielle deltakere får tilstrekkelig informasjon om studiens hensikt og omfang. Når deltakere gir en aktiv bekreftelse på at de ønsker å delta, er

sjansen større for at deltakerne har satt seg inn i hva studien innebærer enn ved passivt samtykke. Elever og foreldre må også få en reell mulighet til å reservere seg fra deltakelse. Som det ble gjort rede for over, kan studier hvor hele klasser mottar inferenstrening, føre til press til deltakelse. Dersom inferenstreningen ligger innenfor de rammer som læreplanen gir, bør lærere og skoleledelse kunne gjennomføre inferenstrening som en del av den ordinære undervisningen uten foreldrenes samtykke. Elever og foreldre vil fortsatt kunne reservere seg mot å delta i datainnsamlingen, selv om inferenstreningen blir gjennomført. Elevene og foreldre vil da være friere til å reservere seg mot å delta på datainnsamling.

#### **4.4 Oppsummering: Verdien av effektstudier av inferenstrening**

Ekte eksperiment er en god metode for å påvise årsaksforbindelser, men praktiske hensyn gjør det vanskelig å gjennomføre i studier der hele klasser deltar. Kvasieksperimentet ivaretar i stor grad den kontroll som kausale problemstillinger krever, dersom validiteten er tilstrekkelig. Som det er gjort rede for over, kan mange validitetsmessige utfordringer ved kvasieksperimentet imøtekommes ved teoretisk forarbeid og god planlegging.

Når kvasieksperimentelle design, så vel som andre design, brukes i pedagogisk forskning, må forskeren ta høyde for at det er mennesker som er gjenstand for studiene. Resultatene kan dermed ikke bli like absolutte som for eksempel i naturvitenskapen. Resultatene er ikke mindre verd av den grunn, siden pedagogisk forskning bør ligge til grunn for og bidra til utvikling i den pedagogisk praksis.

Av de ulike varianter av kvasieksperimentet er studier med både pre- og posttester samt kontrollgruppe best egnet til å måle effekter av inferenstrening, dersom motivasjonen for studien er både teoretisk og pedagogisk. Dette fordi en med kontrollgruppe kan redusere trusler mot validiteten som for eksempel modning og historie. Ved å inkludere pretester kan man også både kontrollere om eksperimentgruppe og kontrollgruppe er sammenlignbare og kontrollere for elevenes ferdigheter i utgangspunktet før oppstart av inferenstreningen. I tillegg bør effekter av inferenstrening måles med både eksperimentdesignede og ikke-eksperimentdesignede tester. Eksperimentdesignede tester er nødvendig av leseteoretiske hensyn. For å tolke generaliserte effekter må man vite om elevene faktisk blir bedre i det de har trent på. Ikke-eksperimentdesignede tester er nødvendige for å måle generaliseringseffekter til for eksempel leseforståelse. Det er også en styrke dersom effekten også måles etter avsluttet trening.

Selv om kvasieksperimentelle design har begrensinger i forhold til generalisering, er det et design som er svært godt egnet til å vinne kunnskap om årsaksforbindelser i det pedagogiske forskningsfeltet. Videre er kvasieksperimentet både praktisk og etisk gjennomførbart. Oppsummert er kvasieksperimentet et solid design som er vel egnet i pedagogisk forskning, forutsatt at forskeren befinner seg innenfor en tradisjon som aksepterer å studere effekt når mennesker er involvert.



## 5. Metaanalyse av inferenstrening i tidligere effektstudier

### 5.1 Mål og hypoteser

I kapittel 2 ble det redegjort for inferens som en ferdighet av stor betydning for leseforståelse. På grunn av den rolle inferens spiller i den helhetlige leseforståelsen, er det i løpet av de siste 30 årene blitt utviklet og forsket på flere undervisningsopplegg som har til hensikt å særlig fremme barns evne til å gjøre inferenser. I dette kapitlet oppsummeres resultater fra effektstudier av inferenstrening i en metaanalyse.

Hensikten med en metaanalyse er å sammenligne forskning innen et avgrenset område, både for å kunne generalisere og for å avdekke konflikter i forskningen. Resultater fra metaanalyser kan angi retninger for praksis og danne grunnlag for ny teori, samt gi utgangspunkt for nye studier. Denne metaanalysen har til hensikt å undersøke effekter av inferenstrening, samt peke ut områder hvor det er behov for ytterligere forskning. Som det er gjort rede for tidligere, er det en korrelasjon mellom leseforståelse og ferdigheter i inferens. Korrelasjoner bekrefter imidlertid kun sammenhenger, man kan ikke trekke konklusjoner om årsaksforbindelser på bakgrunn av korrelasjoner. Leseteoretisk er effektstudier av inferenstrening interessant for å påvise årsaksforbindelsen mellom inferens og leseforståelse. Pedagogisk er inferenstrening interessant for å finne en måte å fremme barns evne til å gjøre inferens på spesielt, og deres leseforståelse generelt. Den overordnede problemstillingen som metaanalysen søker svar på, er følgende:

*Hvordan er årsaksforbindelsen mellom inferens og leseforståelse?*

I tillegg ble det formulert følgende underproblemstillinger:

- 1) *Hva er effekten av inferenstrening for inferens spesielt og for leseforståelse generelt?*
- 2) *Opprettholdes eventuelle effekter etter at inferenstreningen er avsluttet?*
- 3) *Hva karakteriserer inferenstrening med sterke effekter?*

Den første problemstillingen angår primæreffekter og generaliserte effekter av inferenstrening som omtalt i avsnitt 4.1, mens problemstilling 2 angår

langtidseffekter både for primæreffekter og generaliserte effekter. Problemstilling 3 søker svar på hva som karakteriserer inferenstrening med sterke effekter. For å finne svar på problemstilling 3 vil studiene bli kategorisert og sammenlignet med hensyn til følgende elementer: For det første hvilke inferenstyper som trenes. Inferenstyper i studiene vil bli kategorisert i henhold til taksonomien i figur 1. For det andre vil komponenter i inferenstreningen i de ulike studiene bli identifisert. For det tredje vil det bli gjort en kategorisering av organiseringen av inferenstreningen for å kunne sammenligne studiene med hensyn til hvem som administrerte treningen, undervisningsrammer og inferenstreningens intensitet og varighet. Til slutt vil studiene bli sammenlignet med hensyn til hvilke elevgrupper som mottok inferenstrening og for hvilke elever treningen var mest effektiv.

## 5.2 Utvalg

Artikler publisert i perioden 1980 til 2008 ble søkt opp via de elektroniske datasøkebasene ERIC og ISI Web of Science. Studiene ble identifisert ved å søke med kombinasjoner av nøkkelordene *inference*, *reading comprehension*, *training*, *teaching*, *intervention* og *instruction*. Relevante artikler ble videre gjennomført for ytterligere referanser. Kun studier av barn i barneskolealder ble inkludert i utvalget. Studier som inkluderte deltakere med fysiske begrensninger, for eksempel døve barn, ble ekskludert, da dette ville ha innvirkninger på alle variabler i studien, og det ville være vanskelig å sammenligne slike studier med resten av utvalget. Alle studiene i utvalget hadde en form for instruksjon i inferens, enten inferenstrening generelt eller spesifikke inferenstyper definert av forskerne. Studier som karakteriserte treningen som "inferenstrening" (*inference training*), "inferens bevissthets-trening" (*inference awareness training*) eller "inferal spørsmålsgenerering" (*inferential questioning*) eller som på annen måte beskrev at deltakerne fikk trening i inferens, ble inkludert. Eksperimentbetingelsene i studiene i utvalget vil heretter bli referert til som inferenstrening. Inferens og/eller leseforståelse var avhengige variabler i alle studiene i utvalget. Studier som kombinerte inferenstrening med andre forståelsesfremmende elementer ble ikke inkludert i utvalget, da dette ville gjøre det vanskelig å se om en eventuell framgang i leseforståelse skyldtes inferenstrening eller den alternative treningen. Søkene i databasene viste at det var mange studier som kombinerte trening i flere ferdigheter.

Til sammen ti studier ble inkludert i utvalget, se tabell 1 for en oversikt. Åtte av de ti studiene ble publisert på 1980-tallet, kun en studie ble publisert etter 1993.

*Tabell 1: Oversikt over studiene i utvalget*

Studie nr	Forfattere	Publiseringsår	Land
1	Hansen	1981	USA
2	Hansen og Pearson	1983	USA
3	Carr, Dewitz og Patberg	1983	USA
4	Holmes	1985	USA
5	Dewitz, Carr og Patberg	1987	USA
6	Sundbye	1987	USA
7	Yuill og Oakhill	1988	Storbritannia
8	Reutzel og Hollingsworth	1988	USA
9	Winne, Graham og Porck	1993	Canada
10	McGee og Johnson	2003	Storbritannia

## **5.2.1 Beskrivelse av studiene i utvalget**

### **5.2.1.1 Deltakere**

De ti studiene inkluderte elever i andre til sjette klasse. Samtlige studier ble gjennomført blant engelskspråklige barn. Totalt antall deltakere i studiene varierte fra 19 til 101, se tabell 2 for en oversikt. Gjennomsnittlig antall deltakere i de ti studiene var 39. I det følgende refereres det til studiens nummer, ikke artikkelforfatter, jevnfør tabell 1.

Deltakerne, til sammen 454 elever, hadde leseferdigheter på forskjellig nivå. Tre av studiene (studie 2, 7 og 10) sammenlignet effekten av inferenstreningen blant sterke og svake lesere. I studie 2 ble sterke og svake lesere definert ut fra skårer på Stanford Achievement Test og lærervurderinger. Sterke og svake lesere ble i studie 7 klassifisert etter ordforrådsskårer på Gates-MacGinitie reading test, og skårer i ordavkoding og forståelse på testen Neale Analysis of Reading Ability. I studie 10 ble svake lesere definert som elever med gjennomsnittlige, eller over gjennomsnittlige skårer på ordavkoding, men med skårer under kronologisk alder på forståelse på testen Neale Analysis of Reading Ability. De sterke leserne i studie 10 hadde gjennomsnittlige eller over gjennomsnittlige skårer på ordavkoding.

Studie 5 og 8 skilte mellom elever med leseferdigheter under gjennomsnittet, gjennomsnittlige leseferdigheter og leseferdigheter over gjennomsnittet. I studie 5 ble de tre ferdighetsnivåene definert ut fra leseskårer på Iowa Test of Basic Skills. De tre ferdighetsnivåene ble i studie 8 definert ut fra skårer på Gates-MacGinitie Reading Test (Level C, Form 1).



Studie 4 inkluderte kun elever med generelt svake leseferdigheter, hvilket ble definert som leseskårer som lå to år eller mer under klassenivå på California Achievement Test.

Studie 9 inkluderte kun elever med generelt svake ferdigheter, dette var elever som fikk spesialundervisning. Elevene var etter kriterier som ikke ble oppgitt i studien, diagnostisert som barn med lærevansker av det lokale skole-distriktet.

Deltakerne i studie 1 hadde leseferdigheter på eller like over klassenivå. Elevenes leseferdigheter ble definert ut fra lærervurdering og kartlegging foretatt av skolens leselærer. Hva kartleggingen besto i, ble ikke oppgitt i studien. Deltakerne i studie 3 var elever i ordinære skoleklasser. Studien oppga ingen ytterligere informasjon om elevenes leseferdigheter ved oppstart av inferenstreningen. Elevene som deltok i studie 6 hadde 90 % eller mer rett på ordavkodingsoppgaver på testen Informal Reading Inventory.

Tabell 2: Oversikt over deltakerne i studiene

Studie nr	Alder	Deltakernes leseferdigheter	N
1	2. klasse	Elever med leseferdigheter på eller like over klassenivå	24
2	4. klasse	Sterke og svake lesere	40
3	6. klasse	Elever i ordinære klasser	75
4	4. og 5. klasse	Svake lesere	24
5	5. klasse	Elever med leseferdighet over og under gjennomsnittlig og gjennomsnittlig	101
6	3. klasse	Elever med gode ordavkodingsferdigheter	36
7	7 til 8 år	Gode og svake leseforståelsesferdigheter	19 <sup>1</sup>
8	3. klasse	Elever med leseferdighet over og under gjennomsnittlig og gjennomsnittlig	71
9	3. til 5. klasse	Elever med lærevansker	24
10	6,5 til 9,92 år	Sterke og svake lesere	40

---

<sup>1</sup> Denne studien besto av til sammen 52 elever. I denne metaanalysen ble en av kontrollgruppene "comprehension exercise" bestående av 14 elever utelatt. I tillegg ble 19 elever med gode leseferdigheter utelatt fra metaanalysen fordi det ikke var mulig å estimere standardavvik for denne gruppen.

### **5.2.1.2 Kategorier av inferens**

Studiene som inngår i denne sammenfatningen, brukte forskjellige definisjoner av inferens, som var mer eller mindre sammenfallende. Ikke alle studiene definerte typer eller kategorier av inferens. For eksempel brukte studie 3 det overordnede begrepet inferens med en presisering av hva de la i begrepet. For å lettere kunne sammenligne studiene vil inferens i studiene bli kategorisert i forhold til inferenstypene som ble presentert i kapittel 3.2. Siden noen av studiene ikke ga klare definisjoner på begrepet inferens og/eller eksempler på inferens, er kategoriseringen av inferenstyper ikke absolutt, kategoriseringen er snarere et redskap for å kunne sammenligne de ulike formene for inferenstrening.

Studie 1 brukte termen "slot filling" inferens om å forstå ideer i teksten basert på leserens langtidsminne. Studie 1 nevner ikke eksplisitt at leseren gjennom "slot filling" inferens danner situasjonsmodeller, men likevel er det mest nærliggende å plassere inferenstypen som undersøkes i studie 1 i kategorien utfyllende inferens. Studie 2 er basert på studie 1, og plasseres derfor i samme kategori selv om det ikke går eksplisitt fram av studie 2 hvilken type inferens som det fokuseres på.

I Studie 3 ble inferens definert som ikke uttalt informasjon ("unstated information") og logiske forbindelser ("logical connections") i teksten. Denne definisjonen gir lite informasjon om hvor mye kunnskap som kreves av leseren for at han eller hun skal kunne gjøre inferenser om ikke uttalt informasjon. Definisjonen ligger dog nærmest kategorien brobyggende inferens, både logisk og kunnskapsbasert i figur 1. Studie 5 ga forholdsvis lite informasjon om hvilke(n) inferenstype(r) som inngikk i undersøkelsen, ut over at de definerte inferens som slutninger på bakgrunn av kunnskap. Studie 5 var en oppfølgingsstudie av studie 3, dermed er det rimelig å kategorisere inferenstypene på samme måte, som brobyggende inferens.

Studie 4 fokuserte på fire typer inferensspørsmål i forbindelse med lesing av tekst, spørsmål som krever at leseren identifiserer person ("person"), handling ("action"), egenskaper ("attribute") og objekt ("object") i teksten. Disse inferensspørsmålene krevde at deltakerne utdypet innholdet i teksten ved hjelp av tidligere kunnskap. Av eksemplene på tekst og spørsmål som var gjengitt i studien, gikk det tydelig fram at inferensspørsmålene krevde at leseren stilte opp en situasjonsmodell. De fire inferensspørsmålene klassifiseres derfor som utfyllende inferens.

Studie 6 brukte termene "slot filling" og "text connecting" med referanse til Warren, Nicholas og Trabasso (1979). Warren et al. har på bakgrunn av tidligere klassifiseringer av inferens identifisert to hovedfunksjoner til inferenser. For det første å fylle inn manglende "hull" med informasjon i teksten, og for det andre å knytte sammen grunnleggende hendelser i teksten. Den førstnevnte funksjonen til inferens handler om "slot filling", og den sistnevnte om "text-connecting". Ifølge Warren et al. krever "slot filling" gjenkalling av informasjon ("information retrieval"), noe som de hevder er mer krevende enn "text-connecting" inferens (s. 47). Selv om verken forfatterne av studie 6 eller Warren et al. direkte knytter "slot filling" inferens til situasjonsmodeller, er det dog utfyllende inferens som er den mest nærliggende kategorien å plassere inferenstypen i. Den andre funksjonen av inferens som inngår i inferenstreningen i studie 6, "text-connecting", handler om å skape sammenheng i teksten. Denne typen inferens har klare fellestrekk med brobyggende inferens i taksonomien i figur 1. Warren et al. sin definisjon av "text-connecting" som en funksjon av inferens, omfatter både det logiske og det kunnskapsbaserte perspektivet i brobyggende inferens.

Studie 7 fokuserte på "lexical" inferens og "prediction". Disse to typene inferens ble ikke eksplisitt definert av forfatterne, men det ble oppgitt eksempler på hvordan deltakerne ble trent i de to typene inferens. Begrepet leksikal inferens ble brukt om den delen av inferenstreningen som innebar at deltakerne fikk presentert setninger, hvorpå de ble bedt om å velge ut et ord som de skulle gi utdypende informasjon om. Deretter skulle deltakerne knytte de ulike ordene sammen for å skape en overordnet forståelse av avsnittet ordene sto i. På bakgrunn av denne informasjonen ble leksikal inferens vurdert å ligge nærmest brobyggende, kunnskapsbasert inferens i inferens-taksonomien som ble presentert i kapittel 3. Deltakerne i studie 7 øvde også på å gjøre forutsigelser ("predictions") ved å lese fortellinger hvor enkelte setninger var blanket ut. Deltakerne fikk i oppgave å gjette innholdet i setningene som var blanket ut på bakgrunn av teksten rundt, hvilket også med en viss forsiktighet kan karakteriseres som brobyggende inferens. I studie 10 fikk deltakerne trening i leksikal og foregripende inferens etter samme prosedyre som i studie 7, derfor ble inferenstypene i studie 10 kategorisert på samme måte som i studie 7.

I studie 8 ble det skilt det mellom ti ulike inferenstyper ("location", "agent", "time", "action", "instrument", "category", "object", "cause/effect", "problem/solution", "feeling/attitude") som deltakerne i fikk opplæring i, både gjennom forklaringer fra lærer og ved arbeid med nøkkelord i tilknytning til de spesifikke inferenstypene. De ti inferenstypene var basert på Johnson og

Johnson (1986), som gir eksempler på alle ti typene, som med en samlebetegnelse kan karakteriseres som utfyllende inferens.

Studie 9 skilte mellom inferens basert på to typer kunnskap, deklarativ og prosedural. Inferens på bakgrunn av deklarativ kunnskap innebærer ifølge forfatterne slutninger ut fra leserens skjema, det vil si hans eller hennes kunnskap om tema for teksten. Inferenser på bakgrunn av prosedural kunnskap innebærer å kjenne til regler, prosedyrer og logikk. Inferens basert på prosedural kunnskap karakteriseres som "text-based inferences" av forfatterne av studie 9. Denne typen inferens er i liten grad avhengig av tidligere skjematisk kunnskap og blir definert som "*inferences that depend fundamentally on procedural knowledge and less so on prior schematic or world knowledge*" (Winne et al., 1993, s. 54). I inferenstreningen var det kun inferenser på bakgrunn av prosedural kunnskap, "text-connecting" inferens, som inngikk. "Text-connecting" inferens ble kategorisert som brobyggende, logisk inferens i forhold til taksonomien i figur 1.

Tabell 3: Oversikt over inferenstyper i studiene

---

Studie	Inferenstype(r)
1	Utfyllende inferens
2	Utfyllende inferens
3	Brobyggende inferens (logisk og kunnskapsbasert)
4	Utfyllende inferens
5	Brobyggende inferens (logisk og kunnskapsbasert)
6	Brobyggende inferens (logisk og kunnskapsbasert)
7	Brobyggende (kunnskapsbasert)
8	Utfyllende inferens
9	Brobyggende inferens (logisk)
10	Brobyggende (kunnskapsbasert)

---

### 5.2.1.3 Grupper

De 10 studiene i utvalget besto av 18 eksperimentgrupper som fikk inferenstrening og 10 kontrollgrupper (hvorav en ble ekskludert fra metaanalysen). Se tabell 4 for en oversikt over gruppene. I dette avsnittet gjøres det først rede for de ulike gruppene, deretter gis det en kort oversikt over innholdet i undervisningen i de ulike eksperimentgruppene i alle 10 studier. Det vil også bli redegjort for studienes motivasjon og begrunnelser for inferenstrening.

*Metaanalyse av inferenstrening i tidlige effektstudier*

*Tabell 4: Oversikt over eksperimentgrupper, subgrupper og kontrollgrupper*

Studie	Eksperimentbetingelse				Kontrollbetingelse	
	Eksperiment-gruppe 1	Eksperiment-gruppe 2	Subgruppe 1	Subgruppe 2	Kontrollgruppe med påvirkning	Kontrollgruppe uten påvirkning
1	Question	Strategy	-	-	Trening i fonologisk bevissthet og ordforråd	-
2	Question-strategy	-	-	-	-	Samme tekster
3	Structured overview-cloze	-	Cloze	-	-	Supplerende oppgaver
4	Strategy plus material	-	Material only	Strategy only	-	Samme tekster
5	Structured overview-cloze	-	Cloze	Structured overview	-	Supplerende oppgaver
6	Question and text segmentation	Explicit text	-	-	-	Supplerende oppgaver
7	Inference training	-	-	-	Decoding	-
8	Generative-reciprocal procedure	Basal inference training	-	-	-	Ordinær undervisning
9	Induktiv feedback	Eksplisitt feedback	-	-	-	-
10	Inference training	-	-	-	Comprehension exercises	-

Effekter av inferenstreningen ble målt i forhold til tre typer grupper: Eksperimentgrupper, subgrupper og kontrollgrupper. Eksperimentgruppene mottok den treningen som studien hadde som hovedmål å undersøke. Noen studier hadde to eksperimentgrupper med to forskjellige typer inferenstrening. Subgruppene fikk trening i en eller flere av delkomponentene i inferenstreningen som eksperimentgruppen mottok. Studiene i utvalget hadde to typer kontrollgrupper. For det første kontrollgrupper uten eksperimentundervisning. Disse kontrollgruppene fikk enten ordinær undervisning, tilleggsoppgaver, eller leste de samme tekstene som eksperimentgruppen, men uten instruksjon eller oppgaver. For det andre kontrollgrupper med eksperimentundervisning. Disse gruppene fikk en form for trening som ikke hadde noe med inferens å gjøre, for eksempel trening i avkodning.

Studie 1 hadde to eksperimentgrupper som fikk inferenstrening og en kontrollgruppe. 10 fortellinger ble brukt som utgangspunkt for inferenstreningen i begge eksperimentgruppene og for undervisningen i kontrollgruppen. Kontrollgruppen fikk tradisjonell undervisning i fortellingene med en blanding av faktaspørsmål og inferensspørsmål. De to eksperimentgruppene fikk to ulike undervisningstilnærminger til de 10 tekstene. Formålet med inferenstreningen til eksperimentgruppe 1, *question group*, var å fremme spontane inferenser gjennom å trene på å besvare inferensspørsmål fra lærer. Forfatteren av studie 1 hevdet at svake prestasjoner i å gjøre inferenser kan skyldes at elevene mangler trening med henvisning til tidligere undersøkelser som har vist at lærerne i liten grad stiller inferensspørsmål. Videre hevdes det at ved å fokusere på inferens i de spørsmål de stiller barna, vil barna kunne endre tankemønsteret fra å huske innholdet til å tolke innholdet. Eksperimentgruppe 2 i studie 1 ble betegnet som *strategy group* eller *strategy method*. Hensikten med treningen som denne gruppen fikk, var å vise betydningen av tidligere kunnskap for forståelsen, samt å gjøre barna bevisste på at de kunne gjøre inferenser mellom tidligere kunnskap og ny kunnskap ervervet gjennom tekst. Forholdet mellom ny og eksisterende kunnskap, og hvordan forståelse krever at leseren knytter ny informasjon til det han eller hun allerede har kunnskaper om, ble konkretisert for barna ved hjelp av symbolske papirstrimler. Grått papir skulle representere den kunnskapen elevene allerede hadde, og striper i farget papir skulle symbolisere den nye kunnskapen. Tre viktige ideer fra teksten ble plukket ut, som barna skulle jobbe med. Først ble barna stilt spørsmål som var relatert til deres tidligere kunnskap. Etter en muntlig gjennomgang skrev barna opp sine egne erfaringer knyttet til tekstinnholdet på det grå arket som representerte tidligere kunnskap. Deretter skulle barna lage hypoteser om hvorvidt noe liknede kunne skje i den historien de nettopp hadde lest, og

skrive dette på det fargede papiret. I tillegg til inferenstreningen fikk de to eksperimentgruppene og kontrollgruppen ordforrådstrening og aktiviteter knyttet til fonologi.

Studie 2 var en oppfølging av studie 1, og kombinerte prinsippene i inferenstreningen i *question* og *strategy* fra studie 1. Studie 2 hadde som mål å evaluere og videreutvikle studie 1 på tre punkter. For det første å undersøke om undervisningsopplegget kunne gjennomføres av en lærer, i motsetning til studie 1, hvor inferenstreningen ble gjennomført av en forsker. For det andre søkte studie 2 svar på om inferenstreningen kunne anvendes på eldre elever enn i studie 1, og for det tredje om det var forskjell mellom hvordan sterke og svake elever profitterte på undervisningen. Kontrollgruppen leste de samme tekstene som elevene i eksperimentgruppen og svarte på spørsmål til teksten.

Motivasjonen for studie 3 var å undersøke om det å trene på en spesifikk strategi i resonnering kunne bidra til å utvikle ferdigheter i inferens og leseforståelse. Videre hadde studien som mål at deltakerne skulle bli i stand til å overføre det lærte til nye tekster. Forfatterne av studie 3 mente det også var et poeng å bruke fagtekster i eksperimentbetingelsen, i motsetning til foregående studier som hadde trent inferens i forbindelse med lesing av narrative tekster. Inferenstreningen i studie 3 besto av tre prosedyrer. For det første "structured overview", eller strukturert overblikk på norsk, som var et kart over hovedideene i teksten ordnet hierarkisk. Bruken av strukturert overblikk ble i studie 3 omtalt som en teknikk for å aktivere bakgrunnskunnskap hos elevene. Den andre prosedyren i inferenstreningen var "cloze", eller luketekster på norsk. Luketekster er tekster der enkeltord eller setninger er blanket ut, og hvor oppgaven er å finne ord eller setninger som mangler. Luketekstene hadde til hensikt å hjelpe elevene med å relatere tidligere kunnskap med informasjon fra teksten. Den tredje prosedyren i inferenstreningen var en "self-monitoring checklist", det vil si en sjekkliste for selvmonitorering som hadde til hensikt å hjelpe elevene med å internalisere de kognitive prosedyrene som er nødvendige for å anvende en lært strategi eller teknikk i en ny situasjon. Sjekklisten inneholdt spørsmål som elevene skulle stille seg selv mens de løste luketekstoppgaver og svarte på inferensspørsmål, som for eksempel "gir svaret mening?". Eksperimentbetingelsen til eksperimentgruppe 1 i studie 3 besto av strukturert overblikk kombinert med luketekster og sjekkliste for selvmonitorering, mens eksperimentbetingelsen til eksperimentgruppe 2 besto av luketekster og sjekkliste for selvmonitorering. Kontroll elevene fikk supplerende oppgaver som å lese samfunnsfaglige tidsskrifter og oppgaver knyttet til kart.

Studie 4, i likhet med studie 1, begrunnet behovet for inferenstrening med at elevene mangler erfaring i å gjøre inferenser fordi lærere sjelden stiller inferenskrevende spørsmål. Studie 4 ble gjennomført blant elever med svak leseforståelse. Forfatteren av studie 4 hevdet at barn med svak leseforståelse i liten grad benytter sin tidligere kunnskap i møte med tekst, videre gjør de færre assosiasjoner som kan knyttes opp mot teksten de leser. Videre hevdet det i studie 4 at barn med svak leseforståelse i mindre grad enn sterke lesere bruker logisk resonnering der hvor inferens er påkrevd. I studie 4 ble inferenstreningen gjennomført i tre eksperimentgrupper, *strategy plus materials*, *strategy only* og *materials only*. I tillegg hadde studien en kontrollgruppe. Noen elementer av inferenstreningen var felles for alle fire grupper. For det første leste alle gruppene tekster høyt i klassen, for det andre deltok elevene i diskusjoner hvor elevene øvde på å matche svar og spørsmål til teksten. Eksperimentbetingelsen *strategy only* innebar at læreren forklarte hvordan han/hun kom fram til svaret på et inferensspørsmål ved å lage forskjellige mulige svar til spørsmålet, og ved å identifisere nøkkelinformasjon fra teksten, før han/hun formulerte et endelig svar. Gruppen *materials only* leste tekster og løste oppgaver som var rangert fra enkle til vanskelige, basert på fire faktorer. For det første hvor svaret til spørsmålet var plassert. For det andre bredde, det vil si i hvilken grad informasjonen som var nødvendig for å besvare oppgaven var å finne i én setning eller var spredt utover flere setninger. For det tredje type respons, viser til i hvilken grad oppgaven krevde ja/nei svar, flervalg eller eget produsert svar. For det fjerde i hvilken grad elevene hadde kjennskap til tema. Gruppen *strategy plus materials* innebar at elevene mottok begge av de ovennevnte betingelsene.

Forfatterne av studie 5 hevder at grunnen til at en del barn ikke gjør inferenser når de leser, skyldes manglende bakgrunnskunnskap, eller de ikke er oppmerksomme på relevant informasjon i teksten som kan bidra til en grundigere forståelse av teksten. Inferenstreningen i studie 5 besto av de samme tre prosedyrene som i studie 3; i tillegg understreket forfatterne i studie 5 at instruksjonen i inferenstreningen skulle følge følgende prinsipper: For det første at elevene skulle få nøye forklaringer om hensikten med prosedyrene. For det andre skulle de tre prosedyrene i treningen formidles gjennom modellering, hvilket innebar at læreren viste hvordan prosedyren skulle praktiseres og ga elevene feedback. Og til slutt skulle læreren oppfordre til selvstendig overføring av prosedyren til situasjoner ut over den lærte. Kontrollelevne i studie 5 fikk supplerende oppgaver av samme type som i studie 3.



Målet med studie 6 var å utvikle en systematisk prosedyre for å utdype innholdet i fortellinger ved å gjøre innholdet i tekster mer eksplisitt og detaljert. Studien testet ut to måter å gjøre tekstinnholdet mer eksplisitt på. For det første ved å skrive om historiene slik at de ble mer detaljerte. For det andre ved å stille barna inferensspørsmål for dermed å gjøre historiene mer eksplisitte gjennom verbal kommunikasjon. Hypotesene i studie 6 var for det første at eksplisitt tekst ville bedre elevenes skåre på stillelesing, interesse for teksten og deres evne til å gjenfortelle og svare på spørsmål fra teksten. For det andre at effekten av skriftlig versus muntlig utdyping av teksten ikke ville være forskjellig. To eksperimentgrupper var inkludert i studien. En gruppe fikk betingelsen *explicit text versions* som innebar at elevene leste tekster som var omskrevet slik at de var detaljerte og eksplisitte. Den andre eksperimentgruppen, mottok betingelsen *questions and text segmentation*. Denne gruppen fikk inferensspørsmål fra lærer underveis mens de leste fortellinger. Fortellingene som elevene leste, ble delt inn i mindre enheter som elevene fikk spørsmål til. Elevene i kontrollgruppen leste de samme tekstene som elevene i de øvrige to gruppene (den opprinnelige versjonen av tekstene) som individuell stillelesing.

Inferenstreningen i studie 7 hadde som mål å synliggjøre overfor elevene hvordan mening kan trekkes ut av teksten, samt å få elevene til å forstå hvordan hvert enkelt ord hjelper leseren å bygge opp tekstens mening. Inferenstreningen besto av tre komponenter, leksikal inferens, stille spørsmål til teksten og predikering. Leksikal inferens ble trent ved at barna leste setninger som de skulle velge ut et ord fra. Deretter ble barna bedt om å gi utdypende informasjon om ordene. Deretter ble samme teknikk benyttet på korte tekstavsnitt. Her ble barna bedt om å finne nøkkelord for så å binde nøkkelordene sammen. Elevene skulle også stille kontrollspørsmål angående hvem, hvor og hvorfor til fortellinger og tekstavsnitt som de leste. Hensikten med å stille spørsmål til teksten var å gjøre elevene oppmerksomme på ulike inferenstyper, samt at spørsmålene skulle bidra til monitorering av egen forståelse. Predikeringskomponenten besto av oppgaver med luketekster hvor deler av teksten var blanket ut. Hensikten med luketekstene var å gjøre elevene oppmerksomme på betydningen av inferens for forståelse. I studie 7 ble det også inkludert en kontrollgruppe som fikk trening i leseforståelse generelt, men som ikke spesifikt arbeidet med inferenser. Denne betingelsen, "comprehension exercise", innebar blant annet at barna stilte spørsmål til teksten. I tillegg ble det også inkludert en kontrollgruppe som fikk trening i avkoding. I denne metaanalysen brukes gruppen som har fått avkodingstrening som kontrollgruppe, da det er rimelig å anta at avkodings-  
treningen har hatt liten innvirkning på gruppens evne til å gjøre inferenser.

Forfatterne av studie 8 hevder at barns problemer med å gjøre inferenser kan spores tilbake til to kilder. For det første at barna har manglende bakgrunnskunnskaper, og for det andre at barna mangler de nødvendige strategier for å kunne analysere den relevante informasjonen. Dette er utgangspunktet for innholdet i de to eksperimentbetingelsene, *generative-reciprocal procedure* og *basal inference training*. *Generative-reciprocal procedure* ble gjennomført i økter bestående av fem steg. Hver økt begynte med en introduksjon hvor læreren introduserte nye inferenstyper. Neste steg innebar at læreren redegjorde for hensikten med økten. Det tredje steget innebar at læreren ga eksempler på en bestemt type inferens og forklarte hvordan inferenstypen i fokus for økten skulle hjelpe elevene å bli bedre lesere. Deretter fikk elevene i oppgave å plukke ut nøkkelord fra en tekst som var relatert til den bestemte inferenstypen. Øktene ble avsluttet med en lærerstyrt diskusjon om inferensene og listene over nøkkelord som elevene hadde laget. Inferenstreningen i *basal inference training* gikk ut på at elevene arbeidet med tekster hvor de ble eksponert for de samme ti inferenstypene som elevene i gruppen *generative-reciprocal procedure*. Elevene som mottok *basal inference training*, arbeidet også med luketekster der enkeltord var blanket ut, men det går ikke fram av studien hva formålet med luketekstene var. Elevene i gruppen *basal inference training* trente også på å svare på skrevne inferensspørsmål til tekst. Kontrollgruppen i studie 8 mottok ordinær undervisning.

Studie 9 fokuserer på betydningen av instruksjon for å lære lesesvake elever å gjøre inferenser. Studien hadde to eksperimentgrupper som begge gjennomførte inferenstrening bestående av de samme stegene, men gruppene fikk forskjellig veiledning eller feedback, *inductive feedback* og *explicit feedback*. Stegene i inferenstreningen innebar at læreren/veilederen leste en tekst høyt mens barnet fulgte med på en kopi av teksten. Deretter fjernet veilederen teksten og stilte forståelsesspørsmål som eleven svarte på. Så fikk eleven teksten tilbake og veilederen gjentok spørsmålet og det eleven hadde svart, og eleven fikk i oppgave så streke under setninger og ord i teksten som hadde ledet ham/henne til svaret. De to eksperimentgruppene fikk forskjellig tilbakemelding. *Inductive feedback* innebar at veileder ga tilbakemelding på om svaret på forståelsesspørsmålet var rett eller galt. Hvis svaret var feil, informerte veilederen om hva som var rett svar og streket under ord og setninger som ledet fram til rett svar. Hvis svaret var rett, noterte veilederen seg det og oppmuntret barnet til å jobbe videre. Et sentralt trekk ved denne formen for tilbakemelding var at barnet selv måtte utlede prosessene som er involvert i å gjøre inferenser. *Explicit feedback* innebar at veileder ga

tilbakemelding på barnets svar på forståelsesspørsmålene. Uansett om svaret var rett eller galt, ga veilederen forklaringer og demonstrerte hvordan anvende en regel og hvordan gjøre bruk av viktige fakta i et avsnitt for å gjøre inferens. Veileder forklarte rollen til hver enkelt informasjon i teksten, også hva som var irrelevant og hvorfor. Studie 9 hadde ingen kontrollgruppe.

Studie 10 gjennomførte inferenstrening i en eksperimentgruppe, *inference training*, som fikk tilsvarende påvirkning som eksperimentgruppen i studie 7. Denne studien hadde også en kontrollgruppe, "comprehension exercise", som fikk tilsvarende påvirkning som kontrollgruppen med samme betegnelse i studie 7.

#### 5.2.1.4 Komponenter i inferenstreningen

Formålet med dette avsnittet er å identifisere ulike komponenter i inferenstreningen i de ti studiene. Dette for å kunne sammenligne de ulike eksperimentgruppene. Aktivitetene i eksperimentgruppene kan klassifiseres i fire komponenter, *å skape bevissthet om inferens, monitorering, aktivisering av relevant kunnskap og å skape forbindelser*. I tillegg inkluderte noen av eksperimentgruppene aktiviteter som ikke kunne klassifiseres i de fire ovennevnte komponentene. Komponenter i inferenstreningen er oppsummert i tabell 5 under.

Komponenten å skape bevissthet om inferens handler om å gjøre deltakerne oppmerksomme på at inferens er nødvendig for å oppnå forståelse for teksten. Å gjøre deltakerne bevisste på inferens er selvfølgelig selve poenget med inferenstrening, men det var likevel mulig å identifisere aktiviteter i inferenstreningen i de ulike studier og eksperimentgrupper som særlig hadde til hensikt å synliggjøre den avgjørende rolle inferens spiller for den helhetlige forståelse av teksten. Å skape bevissthet om inferens ble for eksempel gjort ved at den som administrerte inferenstreningen, ga eksplisitte forklaringer og modellerte hvordan man resonnerer seg fram til riktig svar på inferensspørsmål. Å skape bevissthet om inferens inngikk i samtlige tre eksperimentgrupper i studie 5, i begge eksperimentgrupper i studie 8, og i gruppen som mottok *strategy method* i studie 4. I studie 7 og 10 ble luketekster brukt for å gjøre elevene bevisste på den betydningen inferens hadde for den overordnede leseforståelsen, således kan inferenstreningen i disse to studiene også klassifiseres innen komponenten skape bevissthet om inferens. Denne komponenten var også sentral i inferenstreningen i *generative-reciprocal procedure* i studie 8. I *generative-reciprocal procedure* ga læreren både eksempler på forskjellige typer inferens og forklarte hvordan det å gjøre en bestemt type inferens ville gjøre elevene til bedre lesere. I

studie 9 ble bevissthet om inferens understreket på to måter. For det første ved at den som administrerte inferenstreningen veiledet elevene, demonstrerte hvordan man kan resonnerer seg fram til rett svar på inferensspørsmål. For det andre ble bevissthet om inferens fremmet ved at elevene fikk i oppgave å streke under ord og setninger i teksten som ledet til rett svar på inferensspørsmålet.

Komponenten monitorering innebærer å kontrollere sin egen forståelse under og i etterkant av lesing. Flere av studiene i utvalget brukte monitorering som et ledd i inferenstreningen på flere måter. For det første ved at elevene stilte seg selv kontrollspørsmål til teksten, og for det andre gjennom sjekklister for forståelse. I studie 7 og 10 inngikk kontrollspørsmål som en del av inferenstreningen. I studie 3 inngikk sjekklister for forståelse i begge eksperimentgrupper, mens sjekklister inngikk i to av de tre eksperimentgruppene i studie 5.

Leseforståelse handler om å knytte kunnskap og erfaringer om verden til det innhold som beskrives i teksten. Forståelse avhenger av at leseren gjør inferenser om teksten med utgangspunkt i sin kunnskap om verden utenfor teksten. Flere av de ti studiene inkluderte komponenten aktivisering av relevant kunnskap. I *strategy method* i studie 1 og i studie 2 ble relevant kunnskap aktivert gjennom at elevene skrev ned sine kunnskaper om tekstens tema før lesing av tekst. I gruppene som fikk strukturert overblikk i studie 3 og 5, ble relevante kunnskaper aktivert gjennom tankekart. I studie 7 og 10 aktiverte elevene relevant kunnskap gjennom å assosiere rundt utvalgte ord i teksten.

Komponenten å skape forbindelser innebærer både det å skape forbindelser mellom relevant kunnskap og teksten, og forbindelser mellom ulike deler av teksten. Førstnevnte forbindelse ble vektlagt i *strategy method* i studie 1 og 2 ved å symbolsk integrere relevant kunnskap med opplysninger i teksten og notere ny og tidligere kunnskap på papirstrimler. Luketekstene i studie 3 og 5 hadde til hensikt å forbinde relevant kunnskap og informasjon fra teksten.

Å stille eller svare på inferensspørsmål til teksten inngikk i samtlige studier i utvalget i forskjellige former. Inferensspørsmål ble brukt i forbindelse med alle de fire ovennevnte komponentene i inferenstreningen. Derfor var det vanskelig å definere det å stille eller svare på inferensspørsmål som en egen komponent i inferenstrening. Innholdet i inferenstreningen i eksperimentgruppen *question group* i studie 1 falt imidlertid ikke inn under noen av de ovennevnte komponentene i inferenstrening. Innholdet i *question group* i studie 1 besto i at læreren stilte inferensspørsmål til elevene etter at de hadde

lest en tekst. Studie 1 viser til tidligere undersøkelser som har funnet at lærere i liten grad stiller inferensspørsmål. Påvirkningen var derfor å stille inferensspørsmål i stedet for faktaspørsmål. De to eksperimentgruppene i studie 6 falt heller ikke naturlig inn under noen av de ovennevnte komponentene. Den ene eksperimentgruppen i denne studien, *question and text segmentation*, innebar å stille inferensspørsmål til elevene etter at de hadde lest en tekst, for på denne måten å gjøre teksten mer eksplisitt for elevene. Den andre eksperimentgruppen i studie 6, *explicit text*, innebar at de samme tekstene som elevene i *explicit text*-gruppen leste, ble skrevet om slik at de ble mer detaljerte og eksplisitte. Eksperimentgruppen i studie 8 som mottok *basal inference training*, svarte skriftlig på inferensspørsmål. Også inferenstreningen i *material only* i studie 4 besto i å svare på inferensspørsmål.

Tabell 5: Oppsummering av kjennetegn ved studiene

Studie	Eksperimentgruppe	Komponenter i treningen
1	Question	-
	Strategy method	B+F
2	Question-strategy	B+F
3	Structured overview-cloze	M+A+F
	Cloze	M+F
4	Strategier- materiell	B
	Material only	-
	Strategy only	B
5	Structured overview-cloze	M+A+F+B
	Cloze	M+F+B
	Structured overview	A+B
6	Question and text segmentation	-
	Explicit text	-
7	Inference training	B+A
8	Generative-reciprocal procedure	B+F
	Basal inference training	-
9	Induktiv feedback	B
	Eksplisitt feedback	B
10	Inference training	B+A

Komponenter i treningen: B= å skape bevissthet om inferens, M= monitorering, A= aktivering av relevant kunnskap, F= å skape forbindelser

### 5.2.1.5 Organisering av inferenstreningen

Inferenstreningen ble i fire av studiene gjennomført av forskere og i fire studier av lærere. I studie 6 ble halvparten av inferenstreningen gjennomført av lærer og halvparten av forsker. Inferenstreningen i studie 9 ble gjennomført

av studenter. I tre av studiene ble inferenstreningen gjennomført i skoleklasser, ingen av disse studiene oppga eksakt antall elever per klasse. Tre studier gjennomførte treningen i grupper på mellom 8 og 12 elever, mens i tre av studiene ble treningen gjennomført i grupper på 3 til 5 elever. I studie 9 ble treningen gjennomført som individuell veiledning. Varigheten på inferenstreningen varierte fra 6 økter à 20 minutter over tre uker til flere måneder. I studie 1, 2, 3 og 6 var det kun oppgitt informasjon om over hvor lang tid inferenstreningen strakk seg, det var uklart hvor mange og hvor lange økter som ble gjennomført. Hvordan inferenstreningen ble organisert i de ti studiene, er oppsummert i tabell 6.

Tabell 6: Oppsummering av organisatoriske trekk ved inferenstreningen i studiene i utvalget

Studie nr	Administrering av treningen	Organiseringen av	Varighet
1	Forsker	Grupper à 8 elever	40 dager
2	Lærer	Grupper à 10 elever	10 uker
3	Lærer	Skoleklasser	8 uker
4	Forsker	Grupper à 3 elever	8 økter à 20 minutter
5	Lærer	Skoleklasser	5 økter à 20 minutter
6	Lærer/forsker	Grupper à 12 elever	3 måneder
7	Forsker	Grupper à 3 til 5 elever	7 økter over 3,5 uker
8	Lærer	Skoleklasser	19 økter à 50 minutter
9	Studenter	Individuell veiledning	9 økter over 5 uker
10	Forsker	Grupper à 5 elever	6 økter à 20 minutter over 3 uker

## 5.3 Metode

### 5.3.1 Måling av effekt av inferenstreningen

Som det ble gjort rede for i avsnitt 4.1, er det fire hovedkategorier av effekter som er interessante for å vurdere hvorvidt inferenstrening har vært effektivt, primæreffekter, primær overførte effekter, generaliserte effekter og utsatt effekt. I dette avsnittet redegjøres det for hvilke typer mål på effekt som ble brukt i de ti studiene. Primær effekt ble målt ved prøver i materialet som ble brukt i inferenstreningen, primær overført effekt ble målt ved prøver med tilsvarende (men ikke samme) tekster og spørsmål som brukt i inferenstreningen. Utsatt effekt ble målt ved mål på primær overført effekt. Videre redegjøres det for hvorvidt testene som ble brukt til å måle effekt, var eksperimentdesignede eller ikke-eksperimentdesignede.

Flere av studiene inkluderte delskårer for eksempel litterale (bokstavelige) spørsmål eller spørsmål knyttet til strategibruk i tillegg til inferensspørsmål. I disse tilfellene vil det kun bli referert til skårer på inferensspørsmål. Det vil ikke bli referert til tester som ikke angår inferens spesielt eller leseforståelse generelt, for eksempel leseinteresse, stillelesingsferdigheter og avkodingsferdigheter.

Studie 1 målte både primæreffekt og primær overført effekt med eksperimentdesignede tester. I testen av primær overført effekt ble det oppgitt skårer både for tekster hvor temaet var kjent og for tekster der temaet var ukjent. Generalisert effekt ble i studie 1 målt ved en standardisert prøve i leseforståelse, altså en ikke-eksperimentdesigned test. På samtlige prøver ble det kontrollert for skårer i leseforståelse ved pretest, målt med en eksperimentdesigned prøve.

I studie 2 ble primæreffekt kartlagt ved en eksperimentdesigned test. En ikke-eksperimentdesigned test ble brukt til å måle generalisert effekt. Dette var tekster med spørsmål hentet fra et lesemateriell som ikke var brukt i inferenstreningen.

Studie 3 målte primæreffekt, primær overført og utsatt primær overført effekt. Primæreffekt ble testet tre ganger i løpet av inferenstreningen. I denne sammenfatningen er kun posttesten som ble gjennomført etter avsluttet inferenstrening, tatt med i beregningen av effekter. Målet på primær overført effekt var også en test i det trenede, men på annet materiell enn det som ble brukt i treningen. Utsatt test av primær overført effekt ble gjennomført 6 uker etter avsluttet inferenstrening. Alle tre testene var eksperimentdesigned. Samtlige prøver var todelt med både litteral (bokstavelig) forståelse og tekstimplisitt forståelse. I denne sammenfatningen blir kun skårer på tekstimplisitt forståelse tatt i betraktning. I de oppgitte gjennomsnittsskårene på posttesten og overføringstestene i studie 3 ble det kontrollert for skårer på en eksperimentdesigned pretest.

I studie 4 ble primær overført effekt målt ved eksperimentdesigned test, generalisert effekt ble målt ved en ikke-eksperimentdesigned, standardisert prøve i leseforståelse. Det ble kontrollert for pretestskårer for begge prøvene.

Studie 5 testet primæreffekt, primær overført effekt og utsatt primær overført effekt (6 uker og 4 måneder etter endt inferenstrening) med eksperimentdesignede prøver. Primæreffekt ble målt på tre tidspunkter i løpet av inferenstreningen, i denne sammenfatningen er kun posttesten som ble

gjennomført etter avsluttet inferenstrening, tatt med i beregningen av effekter. Utsatt effekt etter 4 måneder ble ikke tatt med i beregning av effekt, da det ikke var opplysninger om standardavvik, og det lot seg ikke gjøre å estimere standardavvik for denne prøven.

Studie 6 testet effekt på tre av tekstene som inngikk i treningen. Det var oppgitt flere mål på effekt, elevens rangering av interesse for fortellingen, skårer i gjenfortelling av fortellingen, spørsmål knyttet til sentrale hendelser i fortellingen, spørsmål knyttet til mål i fortellingen, for eksempel en persons mål ved en spesifikk handling, rangering av stillelesing og rangering av forståelse. I denne metaanalysen ble kun beregnet effektstørrelser for rangering av forståelse. I studie 6 ble eksperimentdesignede mål på primæreffekt brukt.

I studie 7 ble generalisert effekt målt ved en ikke-eksperimentdesignet, standardisert prøve i leseforståelse før og etter inferenstreningen. Den oppgitte effekten var endringsskåre fra pretest til posttest.

Fire eksperimentdesignede tester ble brukt til å måle effekter av inferenstreningen i studie 8. Primær overført effekt målt med en test hvor de inferenstypene som elevene i begge eksperimentbetingelsene hadde blitt eksponert for, inngikk. Generalisert effekt ble målt ved to prøver, en luketekst og tekster med spørsmål som både krevde at elevene hentet informasjon direkte ut av teksten og spørsmål som krevde at elevene trakk slutninger ut over det som teksten eksplisitt uttrykte. Studie 8 oppga delskårer for begge typer spørsmål, i metaanalysen ble effektstørrelser beregnet for kun delskårene for sistnevnte spørsmålstype. Både luketekstene og tekstene med spørsmål hadde til hensikt å måle leseforståelse. Selv om arbeid med luketekster inngikk i *basal inference training*, ble ikke luketekstprøven karakterisert som et mål på primæreffekt, fordi forfatterne understreket at prøven var en overføringstest. I tillegg ble generalisert utsatt effekt målt en måned etter avsluttet inferenstrening ved tekster med spørsmål tilsvarende målet på generalisert effekt. Samtlige prøver var utviklet av forskerne selv, og kan derfor karakteriseres som eksperimentdesignede.

I studie 9 ble primær effekt målt ved en eksperimentdesignet prøve. Prøven ble gjennomført som både pretest og posttest. Siden studie 9 ikke hadde kontrollgruppe, ble effektstørrelsen beregnet på de to eksperimentgruppens endring fra pretest til posttest.



Studie 10 målte generalisert effekt av inferenstreningen med en standardisert, ikke-eksperimentdesignet prøve i leseforståelse.

Ikke-eksperimentdesignede tester ble kun brukt til å måle generalisert effekt. Eksperimentdesignede tester ble brukt til å måle både primær-, primær overført og generalisert effekt. Hvilke mål på effekt som inngikk i de ti studiene er oppsummert i tabellen under.

Tabell 7: Oversikt over mål på effekt i studiene

Studie	Eksperimentdesignede tester			Ikke-eksperimentdesignede tester	
	Primær-effekt	Primær overført effekt	Generalisert effekt	Utsatt effekt	Generalisert effekt
1	X	X			X
2	X				X
3	X	X		X	
4		X			X
5	X	X		X	
6	X				
7					X
8		X	X	X	
9	X				
10					X

### 5.3.2 Beregning av effektstørrelser

Cohen's  $d$  ble beregnet for å få ett felles mål på effekt, og dermed et grunnlag for sammenligning av studiene. Cohen (1988) definerer  $d$  som differansen mellom gjennomsnittene til eksperimentgruppe og kontrollgruppe ( $M_1 - M_2$ ), dividert på samlet standardavvik. Samlet standardavvik ble beregnet ved hjelp av følgende formel hvor  $\sigma_1$  og  $\sigma_2$  referer til henholdsvis eksperimentgruppens og kontrollgruppens standardavvik (Rosenthal & Rosnow, 2008, s. 364):

$$\sigma_{\text{pooled}} = \sqrt{(\sigma_1^2 + \sigma_2^2) / 2}$$

Effektstørrelsens styrke beskrives i tråd med Cohens (1988) definisjoner av små = 0.2, medium = 0.5 og store = 0.8 effektstørrelser (s. 25-26).

Det var ikke oppgitt informasjon om standardavvik i studie 5 og 7, derfor ble standardavvikene i disse studiene estimert på bakgrunn av t-verdier for tohalet test, ut fra oppgitte signifikansnivåer. I studie 5 ble det ikke funnet signifikante effekter av målet på utsatt effekt etter 4 måneder, og signifikansnivå for denne prøven var ikke oppgitt. Derfor var det ikke mulig å estimere standardavvik for prøven og dermed heller ikke mulig å beregne

effektstørrelser. I studie 7 var det kun oppgitt signifikansnivå for gruppen med svake ferdigheter i leseforståelse, det var dermed kun mulig å estimere signifikansnivå for denne gruppen.

Hvilke typer skårer som var oppgitt i studiene varierte. Studie 2, 5, 6 og 8 oppga kun posttestskårer, mens studie 1, 3 og 4 oppga posttestskårer hvor det var kontrollert for de enkelte elevens skårer på pretest. Studie 7 oppga endringsskårer, det vil si forskjellen mellom pretest og posttest. Skårene som var oppgitt i studie 7, var ikke endring i råskårer, men endring i lesealder. Studie 9 hadde ikke kontrollgruppe, her ble effektstørrelse beregnet på bakgrunn av pre- og posttestresultater. Studie 10 hadde ingen ren kontrollgruppe; i studien ble gruppen som hadde deltatt i inferenstrening sammenlignet med en gruppe som har deltatt i leseforståelsesundervisning ("comprehension exercise"). I studie 10 var det ikke oppgitt gjennomsnittsskårer for gruppene. Derimot ble det oppgitt opplysninger om F-verdier fra variansanalyse. Det var imidlertid ikke mulig å skille mellom effekter for sterke og svake lesere. For studie 10 ble effektstørrelsen,  $d$ , beregnet på bakgrunn av F test ved hjelp av Hays (1994, s. 411) formel.

$$d = \sqrt{F(1/n_1 + 1/n_2)}$$

I noen av studiene ble kun opplysninger om det totale antall deltakere oppgitt, i disse tilfellene ble det antatt at det var likt antall personer i gruppene, og gjennomsnittlig antall deltakere ble brukt i beregninger.

Gjennomsnittlige effektstørrelser ble kalkulert ved å vekte hver effektstørrelse med dens inverse varians ved hjelp av Lipsey og Wilsons (2001, s. 114) formel. Symbolene ES og  $w$  henviser til henholdsvis effektstørrelse og effektstørrelsens inverse varians.

$$\overline{ES} = \frac{\sum (w_i ES_i)}{\sum w_i}$$

I metaanalyser strebes det etter å identifisere alle relevante studier til utvalget. Det er aldri mulig å være sikker på hvor mange relevante studier som faktisk finnes, fordi upubliserte studier kan være vanskelige å oppspore. Dette er et problem i metaanalyser fordi det er større sannsynlighet for at publiserte studier har flere signifikante resultater enn upubliserte studier (Kramer & Rosenthal, 1999; Lipsey & Wilson, 2001). Dette problemet ble avdekket i en metastudie av Lipsey og Wilson (1993) hvor utvalget besto av både publiserte og upubliserte studier. Lipsey og Wilson fant at den gjennomsnittlige vektete effektstørrelsen var større for publiserte enn for upubliserte studier. Siden

publiserte studier er mer tilgjengelige, er det fare for "utvalgsbias", det vil si at utvalget er systematisk feil. For å gi et anslag over hvor stor sjansen er for at man trekker en uriktig konklusjon som følge av denne systematiske feilen ved utvalget, er det blitt utviklet en formel som estimerer hvor mange upubliserte med ikke-signifikant resultat som må til for å redusere effekten på utvalget av publiserte studier til 0, kalt "fail-safe N" (Lipsey & Wilson, 2001, s. 165-166). I formelen referer  $k$  til antallet studier i den gjennomsnittlige effektstørrelsen.  $c$  referer til kriteriet for nivået i effektstørrelsen, det vil si signifikansnivået.

$$k_0 = k(\overline{ES}_k / \overline{ES}_c - 1)$$

Fail-safe N vil i det følgende bli oppgitt i tabeller hvor gjennomsnittlig vektete effektstørrelser er oppgitt. Fail-safe N vil bli regnet med utgangspunkt med signifikansnivå  $p < .05$ .

## 5.4 Resultater

### 5.4.1 Primær-, primær overført og generalisert effekt

I de ti studiene var det oppgitt resultater på både primær, primær overført og generalisert effekt. Studie 2, 5 og 8 hadde oppgitt skårer for elever med ulike ferdighetsnivåer i lesing. Studie 5 og 8 hadde i tillegg opplyst om totalskårer, effektstørrelsene i tabellen under ble beregnet på totalskårer. Studie 9 hadde oppgitt ti delskårer per gruppe for målet på primæreffekt, effektstørrelsene i tabellen er gjennomsnittet av effektstørrelser for de ti delskårene. Studie 6 hadde tre delskårer per gruppe, også her er effektstørrelsene i tabellen gjennomsnittet. Studie 1 oppga to effektstørrelser for primær overført effekt, en for tekster der temaene var kjente og en for tekster der temaet var ukjent. I studie 8 var det også oppgitt to delskårer for generalisert effekt. I tabellen under er de gjennomsnittlige effektstørrelsene oppgitt.

I tabell 8 er effektstørrelser (Cohen's  $d$ ) for samtlige studier på de tre nivåene av målt effekt oppgitt. Ikke alle studiene oppga effekter for alle tre nivåene, derfor inneholder tabellen mange manglende verdier.

De primære effektene av inferenstreningen var gjennomgående moderate til sterke, med unntak av for gruppen med sterke lesere i studie 2, hvor effekten var svært liten. Det var stor variasjon i de primær overførte effektene. Studie 5 rapporterte svært høye effekter, mens både studie 1 og 8 rapporterte at kontrollgruppen hadde overgått en av eksperimentgruppene, og dermed hadde negativ effekt. Studie 7 hadde en særdeles sterk effekt på målet på generalisert

effekt. Her er det verd å merke seg at effekten kun ble beregnet for gruppen som hadde svake ferdigheter i leseforståelse av grunner redegjort for i avsnitt 5.3.2. Dette innebærer at deltakerne som det ble beregnet effektstørrelse for, hadde stort potensial for framgang. De generaliserte effektene som ble rapportert, var generelt sterke med unntak av gruppen *basal inference training* i studie 8, som også på dette effektmålet hadde negativ effektstørrelse, altså mindre framgang enn kontrollgruppen.

Tabell 8: Oppsummering av effektstørrelser (Cohen's *d*) for primær-, primær overført og generalisert effekt

Studie	Betingelse	Primær-effekt	Primær overført effekt	Generalisert effekt
1	Question	1.41	0.76	1.57
	Strategy	1.43	0.26	0.79
2	Question-strategy (sterke lesere)	0.11	-	1.12
	Question-strategy (svake lesere)	1.52	-	1.03
3	Structured overview-cloze	2.01	0.51	-
	Cloze	2.05	0.70	-
4	Strategy plus materials	-	2.01	1.16
	Materials only	-	0.72	2.11
	Strategy only	-	0.33	0.81
5	Structured overview-cloze	2.01	1.90	-
	Cloze	1.98	3.62	-
	Structured overview	0.78	2.42	-
6	Question and text segmentation	1.59	-	-
	Explicit text	1.04	-	-
7	Inference training (svake lesere)	-	-	5.42
8	Generative-reciprocal procedure	-	1.63	0.65
	Basal inference training	-	-0.02	0.07
9	Inductive feedback	0.47	-	-
	Explicit feedback	0.62	-	-
10	Inference training	-	-	0.71
<i>Vektet gjennomsnittlig effektstørrelse</i>		<i>1.39</i>	<i>1.07</i>	<i>0.69</i>
<i>Fail-safe N</i>		<i>268</i>	<i>204</i>	<i>128</i>

Studie 3, 4 og 5 inkluderte subgrupper som mottok deler av eksperiment-undervisningen. I alle disse tre studiene var det eksempler på at en eller begge subgruppene overgikk eksperimentgruppen i effekter på ett eller to av målene. Forfatterne av studie 3 forklarte dette med at gruppene i utgangspunktet ikke matchet på ferdigheter i leseforståelse. I studie 4 var det oppgitt standardavvik og justerte gjennomsnittskårer. Det var svært liten forskjell i justerte gjennomsnittskårer mellom gruppene *material plus strategy* og *materials only* på målet på generalisert effekt. Gruppen *material plus strategy* hadde et

betraktelig høyere standardavvik, noe som gjorde at effektstørrelsen ble høyere for gruppen *materials only* enn for *material plus strategy*. I diskusjonen av gjennomsnittsskårene i studie 4 antydes det at det var innholdet i *materials only* som ga effekt. I studie 5 ble det ikke direkte kommentert at subgruppene hadde bedre resultater på primær overført effekt enn eksperimentgruppen, men forfatterne tolket resultatene som at innholdet i cloze-prosedyren fungerte særlig bra med tanke på å bedre elevenes evne til å svare på både inferensspørsmål og litterale (bokstavelige spørsmål). Den andre prosedyren, structured overview, hadde ifølge forfatterne ikke blitt gjennomført på en hensiktsmessig måte, og prosedyren derfor ikke hadde den forventede effekt.

De vektete gjennomsnittlige effektstørrelsene var sterke for alle tre mål på effekt. Effekten var sterkest for mål på primæreffekt og noe mindre for primær overført effekt og generalisert effekt.

#### **5.4.2 Effekter knyttet til kjennetegn ved studiene**

I dette avsnittet presenteres resultater som belyser problemstilling 3 om hva som karakteriserer inferenstrening med sterk effekt. Studiene blir sammenlignet med hensyn til inferenstyper, aktiviteter i inferenstreningen og hvordan treningen ble organisert og gjennomført

##### **5.4.2.1 Effekter for inferenstyper**

De ti studiene ble kategorisert i forhold til hvilke inferenstyper som ble trent. Fire av studiene trente utfyllende inferens, og seks studier trente brobyggende inferens. Vektete gjennomsnittlige effektstørrelser ble beregnet for de to typene inferens. Som det går fram av tabellen under, var effektene jevnt høyere for studiene som trente brobyggende inferens. Det ser med andre ord ut til at det kan være større effekter av å trene brobyggende inferens enn utfyllende inferens. Dette resultatet vil bli ytterligere kommentert og diskutert i avsnitt 5.5.

Tabell 9: Vektete gjennomsnittlige effektstørrelser for studier av brobyggende inferens og utfyllende inferens<sup>2</sup>

	Brobyggende inferens		Utfyllende inferens	
	Studie	Vektet gjennomsnittlig effektstørrelse	Studie	Vektet gjennomsnittlig effektstørrelse
Primær effekt	3, 5, 6	1.47 (85.2)	1, 2	1.02 (38.8)
Primær overført effekt	3, 5	1.48 (57.2)	1, 4, 8	0.66 (36.6)
Generalisert effekt	7, 10	1.16 (44.4)	1, 2, 4, 8	0.62 (45.6)

#### 5.4.2.2 Effekter for komponenter i inferenstreningen

Det var ikke mulig å beregne gjennomsnittlige effektstørrelser for de ulike komponenter i inferenstreningen i studiene, eller på annen måte identifisere aktiviteter som var særlig effektive å trene for å bedre elevenes inferens og leseforståelse. Dette fordi det var stor overlapp i komponentene i de fleste betingelsene.

#### 5.4.2.3 Effekter knyttet til hvem som gjennomførte inferenstreningen

Som det går fram av tabell 6, ble inferenstreningen i fire av studiene gjennomført av lærer og fire av studiene av forsker. I de siste to studiene ble inferenstreningen henholdsvis gjennomført av studenter og delvis av forsker og lærer. For de fire studiene hvor treningen ble gjennomført av forsker, var de vektete gjennomsnittlige effektstørrelsene for primæreffekter, primær overført effekt og generalisert effekt henholdsvis 1.42, 0.63 og 1.20. For de studiene hvor lærer gjennomførte inferenstreningen, var de vektete gjennomsnittlige effektstørrelsene for primæreffekter, primær overført effekt og generalisert effekt henholdsvis 1.54, 0.64 og 0.46. Med andre ord var det svært liten forskjell i primæreffekt og primæreffekter mellom elever hvor lærer hadde undervist i inferens og elever hvor forskere sto for undervisningen. Med hensyn til generaliserte effekter var det betraktelig større forskjell. Dette skyldes den særlig sterke effekten i studie 7.

#### 5.4.2.4 Effekter knyttet til gruppestørrelse og treningens varighet

Som det går fram av tabell 6, ble inferenstreningen gjennomført både i hele klasser, individuelt og i grupper av ulik størrelse. På grunn av at kun to studier gjennomførte inferenstrening i like store grupper, ble det ikke beregnet gjennomsnittlige vektete effektstørrelser basert på gruppestørrelse inferenstreningen ble gjennomført i. Ut fra en kvalitativ sammenligning av effektstørrelsene i tabell 8 var det ingen ting som tydet på at gruppestørrelse hadde betydning for effekt. Det var heller ingen klare tendenser til at studiene

<sup>2</sup> Fail-safe N er oppgitt i parentes

med flest eller lengst økter med inferenstrening eller med lengst treningsperiode var mer effektive enn studier der inferenstreningen strakk seg over kortere tid. Her må det tas i betraktning at den oppgitte informasjonen om treningsens varighet og intensitet var noe begrenset i enkelte av studiene.

#### 5.4.2.5 Effekter knyttet til deltakernes leseferdigheter

Flere av studiene oppga delskårer for sterke, gjennomsnittlige og svake lesere; i tillegg ble noen av studiene gjennomført kun for elever som ble karakterisert som svake lesere (se tabell 2 for en oversikt over leseferdigheter hos deltakerne). Det ble ikke beregnet gjennomsnittlige effektstørrelser for elever på ulike ferdighetsnivåer. Det var to grunner til dette. For det første var kategoriseringen av ferdighetsnivå i studiene såpass forskjellig at det ikke ga mening å sammenligne effekter for elever på ulike ferdighetsnivå i lesing. For det andre er det sannsynlig at studiene tilpasset inferenstreningen etter ferdighetsnivået til elevene.

#### 5.4.3 Langtidseffekter

Tre av de ti studiene målte effekter med en utsatt posttest 4 til 6 uker etter at inferenstreningen var avsluttet. De tre studiene som gjennomførte utsatt posttest, målte kun primær overført effekt. Ingen av studiene målte generalisert effekt med utsatt posttest. Effektstørrelser for samtlige eksperimentgrupper i disse tre studiene er gjengitt i tabell 10 under. Styrken på effektstørrelsene for de sju eksperimentgruppene var svært varierende fra både negative og ubetydelige effekter i henholdsvis *structured overview* i studie 5 og *basal inference training* i studie 8 til sterk effekt for gruppen *generative-reciprocal procedure* i studie 8. Begge de to eksperimentgruppene i studie 3 skåret høyere på den utsatte testen enn på testen som ble gjennomført umiddelbart etter avsluttet inferenstrening. Eksperimentgruppene i studie 5 og 8 hadde lavere effekter på utsatt prøve av primær overført effekt enn på den umiddelbare prøven. Den ene eksperimentgruppen i studie 5, *structured overview*, hadde negativ effekt på den utsatte prøven av primær overføring. Dette kan tolkes som at innholdet i *structured overview* ikke har effekt når den ikke gjennomføres sammen med *cloze*. Som det går fram av tabellen under, var den vektete gjennomsnittlige effektstørrelsen for primær overføring målt ved utsatt test moderat, og lavere enn den vektete gjennomsnittlige effektstørrelsen for primær overføring målt umiddelbart etter gjennomført inferenstrening (jvnfør tabell 8).

Tabell 10: Effektstørrelser for primær overføring målt med utsatt prøve

Studie	Eksperimentbetingelse	Testtidspunkt etter avsluttet trening	Primær overført effekt
3	Structured overview/cloze	6 uker	0.96
	Cloze	6 uker	0.89
5	Structured overview/cloze	6 uker	0.04
	Cloze	6 uker	0.82
	Structured overview	6 uker	-1.08
8	Generative- reciprocal procedure	4 uker	1.25
	Basal inference training	4 uker	0.07
Vektet gjennomsnittlig effektstørrelse			0.41
Fail-safe N			21.6

## 5.5 Diskusjon

### 5.5.1 Effekter av inferenstrening på fire nivåer

I alt 10 studier av inferenstrening blant skolebarn ble identifisert og inkludert i denne metaanalysen. Effekter av inferenstreningen ble målt på fire nivåer, primær effekt, primær overført effekt og generalisert effekt. I dette avsnittet diskuteres den generelle effekten av inferenstrening for inferens og leseforståelse (jevnfør problemstilling 1), samt effekter etter avsluttet inferenstrening (jevnfør problemstilling 2).

Den generelle tendensen i de ti studiene var at effekt målt med test på primæreffekt var sterkst, noe som ikke er overraskende siden oppgavene i testen på primæreffekt ble målt med materiell som ble brukt i inferenstreningen. Effekten var svakest for mål på generalisert effekt. Med andre ord var de vektete gjennomsnittlige effektstørrelsene svakere desto fjernere de var fra innholdet i inferenstreningen. Fem eksperimentgrupper (begge eksperimentgruppene i studie 1, gruppen med sterke lesere i studie 2, gruppene *materials only* og *strategy only* i studie 4) fulgte imidlertid ikke dette mønsteret. Disse gruppene hadde sterkere effektstørrelser på målene på generalisert effekt enn på målene på primæreffekt og/eller overført primæreffekt. Siden målene på generalisert effekt var prøver i leseforståelse, hadde treningen i disse eksperimentbetingelsene sterkere effekt på leseforståelse enn på inferens. Grunnen til dette kan være at forhold som ikke angikk inferenstreningen spilte inn. Det kan også skyldes at inferens er en ferdighet som virker inn på andre områder som igjen har innflytelse på leseforståelse. En tredje forklaring kan være at ett enkelt element i inferenstreningen fikk stor innflytelse på det generaliserte målet, slik at prøvene i primæreffekt og primær overført effekt kun delvis berørte dette



elementet, mens det samme elementet fikk større utslag på testen i generalisert effekt.

Siden målene på primæreffekter og primær overført effekt var tester i inferens, og målene på generaliserte effekter var testene i leseforståelse, blir svaret på problemstilling 1 at inferenstrening hadde effekt både på inferens og leseforståelse.

Både primæreffekter og primær overført effekt ble målt med eksperimentdesignede tester i samtlige studier. Generaliserte effekter ble derimot målt med ikke-eksperimentdesignede tester i ni av de ti studiene. Siden effekten var lavere for målene på generaliserte effekter enn for de to andre målene på effekt, er det tilsynelatende en tendens til større effekter på eksperimentdesignede tester enn på ikke-eksperimentdesignede tester, noe som er i tråd med Rosenshine og Meister (1994) i sin sammenfatning av effektstudier av resiprok undervisning (se avsnitt 4.1). En alternativ, og kanskje mer sannsynlig, forklaring på hvorfor den gjennomsnittlige vektede effektstørrelsen var mindre for mål på generalisert effekt enn de to øvrige, er at målet på den generaliserte effekten var innholdsmessig fjernere fra innholdet i inferenstreningen enn målene på primæreffekt og primær overført effekt.

Den vektede gjennomsnittlige effektstørrelsen for primær overføring var lavere ved den utsatte prøven enn umiddelbart etter avsluttet inferenstrening. Altså var tendensen i de tre studiene som målte utsatt effekt, at effekten av inferenstreningen avtok etter at inferenstreningen var avsluttet. I studie 3 var imidlertid effektene større på den utsatte prøven enn umiddelbart etter inferenstreningen. Dette kan bety at effekten av inferenstreningen først slår ut en stund etter at elevene har jobbet med inferenser. Siden alle prøvene i denne studien var eksperimentdesignet, kan dette resultatet skyldes at vanskegraden på den umiddelbare og den utsatte testen av primær overført effekt ikke var lik. En annen forklaring kan knyttes til hvem som gjennomførte inferenstreningen. I studie 3 ble inferenstreningen gjennomført av elevenes lærer. Det er sannsynlig at læreren fortsatte å anvende prinsippene fra inferenstreningen også etter at eksperimentperioden var avsluttet.

### **5.5.2 Effekter for studier med ulike kjennetegn**

De ti studiene ble kategorisert i forhold til inferenstyper, komponenter i treningen, elevenes leseferdigheter og organisatoriske trekk som hvem som gjennomførte undervisningen, treningens varighet og intensitet og størrelsen på gruppene treningen ble gjennomført i. Dette for å svare på problemstilling

3 om hva som kjennetegner inferenstrening med sterk effekt. I dette avsnittet diskuteres resultater knyttet til disse kategoriseringene.

Studier som fokuserte på brobyggende inferens hadde generelt sterkere effekter enn studier som trente utfyllende inferens. Det er rimelig å gå ut fra at inferenstreningen i de respektive studiene var tilpasset den inferenstypen som de hadde til hensikt å trene elevene i. Det at effektene var sterkest for studier som trente brobyggende inferens, betyr ikke nødvendigvis at den pedagogiske verdien er størst ved å trene brobyggende inferens. Dersom en lavere effektstørrelse for utfyllende inferens er uttrykk for at denne inferenstypen er vanskeligere å trene, vil det kanskje være særlig viktig å gi elever undervisning i utfyllende inferens.

Det var ikke mulig å sammenligne effekten av studier med ulike komponenter i inferenstreningen på grunn av stor overlapp i studiene. Det var med andre ord klare likheter ved innholdet i inferenstreningen i de ti studiene på tross av at de trente to typer inferens. De komponentene som ble identifisert, kan således betraktes som aktiviteter som fremmer inferens generelt, men denne metaanalysen gir ikke svar på hvilke komponenter som er mest effektive å trene for å bedre elevene inferens og leseforståelse.

Generelt var tendensen at organiseringen av inferenstreningen ikke hadde stor betydning for effektene. Hvem som gjennomførte inferenstreningen eller over hvor lang tid treningen strakk seg, hadde ikke tydelig innvirkning på effektene. Om inferenstreningen ble gjennomført i store eller små grupper så heller ikke ut til å spille noen avgjørende rolle for effekten av inferenstreningen. Dette gikk særlig fram av studie 9, hvor inferenstreningen foregikk i hel klasse. Studien hadde to eksperimentgrupper, og resultatet var svært forskjellig i de to gruppene. Mens det var sterke effekter for den ene gruppen, *generative-reciprocal procedure*, hadde den andre eksperimentgruppen, *basic inference training*, svært svak effekt, faktisk hadde denne gruppen mindre framgang enn kontrollgruppen på to av de tre målene på effekt. En rimelig tolkning av disse resultatene er at det ikke er størrelsen på gruppene som er avgjørende for effekten av inferenstrening, noe som er i samsvar med hva Rosenshine og Meister (1994) fant i sin metaanalyse av resiprok undervisning.

Denne metaanalysen ga ikke klart svar på for hvilke elever, med hensyn til leseferdigheter, som inferenstrening er mest effektiv for. Studiene som kun undersøkte effekten av inferenstrening blant elever med samme ferdighetsnivå (som studie 1, 4 og 9), tilpasset trolig ferdighetsnivået til deltakerne, mens

andre studier som inkluderte elever på ulike ferdighetsnivåer (for eksempel studie 2 og 5), sannsynligvis var tilpasset gjennomsnittseleven. Derfor var det ikke mulig å sammenligne for hvilke elever inferenstreningen var mest effektiv.

### **5.5.3 Begrensinger ved metaanalysen**

Utgangspunktet for denne metaanalysen var ti studier av inferenstrening. Dette er et forholdsvis lavt antall studier, like fullt kan samme framgangsmåte ifølge Kramer og Rosenthal (1999) brukes for metaanalyser med små utvalg som for metaanalyser med store utvalg. Et generelt problem ved metaanalyser er at man aldri kan vite hvor mange upubliserte studier som er gjennomført. Ifølge Kramer og Rosenthal er dette problemet av særlig betydning for metaanalyser med få studier. I denne studien ble det imidlertid regnet ut et estimat for hvor mange studier med ikke-signifikante resultater som må til for å redusere effekten av studiene i utvalget til 0. Antallet studier med ikke-signifikant resultat som må til for å effektene målt på de tre nivåene, var forholdsvis betydelige, 268, 204 og 128 for henholdsvis primæreffekt, primær overført effekt og generalisert effekt. Kun tre studier målte effekt med utsatt prøve, antallet studier med ikke-signifikante resultater som må til for å forkaste den vektete gjennomsnittlige effektstørrelsen på utsatt prøve, var derfor betraktelig mindre, fail-safe N ble estimert til 21.6. Med andre ord er usikkerheten hva langtidseffekter av inferenstrening angår, større enn for den umiddelbare effekten.

De ti studiene i utvalget i denne metaanalysen har klare fellestrekk i og med at samtlige av studiene undersøkte effekten av inferenstrening blant skolebarn, og at inferens og/eller leseforståelse var uavhengig variabel i alle studiene. Samtidig er studiene designet og gjennomført av forskjellige forskere under ulike omstendigheter over en tidsperiode på over tjue år. Hypotesene og bakgrunnen for de 10 studiene er heller ikke de samme. Spørsmålet som melder seg i denne metaanalysen, som i de fleste andre metaanalyser, er derfor i hvilken grad studiene i utvalget er sammenlignbare. Problemstillingene som ble presentert i avsnitt 5.1 åpnet for sammenligninger på flere områder, hvorav noen sammenligninger ble utelatt. Av grunner som allerede er nevnt, var det ikke hensiktsmessig å sammenligne studiene verken med hensyn til hvilke komponenter i inferenstreningen som hadde mer eller mindre effekt, eller med hensyn til ferdighetsnivå i lesing. Metaanalysen gir derfor begrenset informasjon om hvem inferenstrening er effektivt for. Riktignok sammenlignet flere av studiene elever på ulike ferdighetsnivå eller gjennomførte studien blant elever på samme nivå i lesing, men definisjoner av ferdighetsnivå på tvers av studiene var ikke sidestilte. Det er uansett ikke

mulig å få et reelt bilde av hvilke elever, med hensyn til ferdighetsnivå, som hadde mest utbytte av inferenstreningen. Dette fordi resultatene som var oppgitt i studiene var basert på råskårer. Resultater basert på råskårer tar ikke høyde for at de elevene som har lav skåre ved pretest har større mulighet for framgang ved posttest enn elever som har høy skåre ved pretest.

Når det gjelder sammenligning av hvilke inferenstyper som var mest effektive å trene, ble det foretatt en sammenligning av studiene, og vektete gjennomsnittlige effektstørrelser ble beregnet. Som det går fram i avsnitt 5.2.1.2 var det noe usikkerhet i forhold til bruken av inferensbegrepet i de ti studiene. Det var forskjell i hvilke begreper som ble brukt og hvilken betydning de la i begrepene. Videre var det et problem at ikke alle studiene ga klare definisjoner og/eller eksempler på hvordan de definerte inferens. På tross av disse problemene ble studiene kategorisert etter taksonomien i figur 1. Det er derfor usikkerhet knyttet til sammenligningen av effekter for studiene av brobyggende versus utfyllende inferens.

En sammenligning som er mindre problematisk i denne metaanalysen, var sammenligningen av effekter på ulike nivå (primær-, primær overført og generalisert effekt), samt effekter på eksperimentdesignede versus ikke-eksperimentdesignede mål. Dette fordi samtlige studier oppga tilstrekkelige opplysninger om måleinstrumentene.

I denne studien var det ikke samme type skårer som var oppgitt i de ulike studiene. Noen studier oppga for eksempel resultater hvor det var kontrollert pretestskårer, andre studier oppga kun posttestskårer. Som det går fram av tabell 8, var det flere studier som ikke oppga informasjon om enten primær-, primær overført effekter eller generaliserte effekter. Studie 7 og 10 oppga kun generaliserte effekter målt med en standardisert prøve i leseforståelse. Dette er problematisk siden det ikke forligger data på hvorvidt elevene faktisk har hatt en forbedring i det de trente på. Man vet da ikke om det er inferenstreningen som er årsaken til bedringen i leseforståelse. Omvendt rapporterte studie 3 og 5 ingen generaliserte effekter, og studie 6 og 9 rapporterte verken primær overførte eller generaliserte effekter av inferenstrening. Informasjon om overføringsverdien av inferenstreningen til andre områder enn eksperiment-situasjonen er derfor begrenset i forhold til disse fire studiene.

#### **5.5.4 Konklusjon**

Denne metaanalysen har vist at inferenstrening hadde effekt på inferens og leseforståelse, og effekten for inferens avtok etter at treningen var avsluttet. Effektene var sterkest for studier som trente brobyggende inferens.

Organisatoriske trekk ved inferenstreningen, for eksempel hvem som gjennomførte treningen og hvor store grupper treningen ble gjennomført i, hadde lite betydning for effekten av inferenstrening. Metaanalysen viste videre at inferenstreningen i de ti studiene besto av mange av de samme komponentene. Metaanalysen kunne imidlertid ikke identifisere elevgrupper, med hensyn til leseferdigheter, som profiterte mer av inferenstrening enn andre.

Det var overraskende få studier som ble inkludert i utvalget. Sammenlignet med for eksempel studier som undersøker effekten av å trene fonologisk bevissthet for avkoding, er studier som undersøker effekter av inferenstrening for leseforståelse i klart mindretall. Det var også overraskende få studier av nyere dato. Det ser ut til at inferenstreningen hadde sin glansperiode på 1980-tallet. Hvorfor så få studier av inferenstrening av nyere dato? Studier av undervisning som har til hensikt å bedre elevers leseforståelse har ikke opphørt, men det kan se ut til at treningen har fått et annet innhold. De siste årene har det blitt publisert mange studier hvor flere ferdigheter i leseforståelse trenes samtidig. Man kan forvente høyere effekter på leseforståelse av studier som trener flere ferdigheter samtidig, men disse studiene kan ikke gi kunnskap om årsaksforhold mellom enkelte ferdigheter og leseforståelse.

Metaanalysen har bidratt med forslag til framtidige studier av inferenstrening, både metodisk og innholdsmessig. Effektstudier bør måle effekt ved pre- og posttest for å sikre at gruppene er like før inferenstreningen starter. For å kunne fastslå om det er elever med sterke, middels eller svake leseferdigheter som har mest utbytte av treningen, bør det ved måling av individuell framgang kontrolleres for at den enkelte elevs mulighet for framgang er den samme. Videre bør effekt av inferenstreningen undersøkes på flere nivåer ved både å måle effekten av inferens ved mål på primær og/eller primær overført effekt, i tillegg til å måle generalisert effekt på leseforståelse. Effekter av inferenstreningen bør også måles en tid etter at inferenstreningen er avsluttet. Som det har gått fram av denne studien, er inferens langt i fra et ensbetydende begrep. Framtidlige studier bør derfor sørge for gi leseren en fullstendig definisjon av både det teoretiske begrepet og operasjonaliseringen av begrepet gjennom beskrivelse av aktiviteter i inferenstreningen.

Det generelle inntrykket av de ti studiene i denne metaanalysen var at argumentet eller motivasjonen for å gjennomføre studier av inferenstrening først og fremst var pedagogisk. De ti studiene som er blitt sammenfattet i dette kapittelet, begrunnet inferenstreningen med troen på et forbedrings-

potensial hos elevene med hensyn til den dypere leseforståelsen. Videre ble det i flere av studiene hevdet at den dypere leseforståelsen er blitt neglisjert i undervisningen. Gjennom denne metaanalysen er det blitt identifisert komponenter og aktiviteter som er hensiktsmessige å undervise i for å bedre elevenes evne til å gjøre inferens spesielt og bedre deres leseforståelse generelt. Hvilke av disse komponentene som gir best effekt ble dog ikke avdekket gjennom metaanalysen.



## **6. Betydningen av situasjonsmodeller for inferens – en forundersøkelse**

### **6.1 Bakgrunn**

Metaanalysen viste generelt sterke effekter av inferenstrening blant skolebarn, noe som gjør det sannsynlig at man vil kunne finne tilsvarende effekter med samme type trening blant norske elever. Likevel avdekket metaanalysen flere begrensninger i tidligere effektstudier av inferenstrening. Derfor ble det i forkant av effektstudien av inferenstrening, som i sin helhet er beskrevet i kapittel 7, gjennomført en forundersøkelse. Forundersøkelsen inkluderte ikke trening over tid, men ble gjennomført med ett prøvetidspunkt hvor deltakerne leste tekster og svarte på spørsmål under to ulike betingelser, eksperiment- og kontrollbetingelse. Med andre ord hadde forundersøkelsen et within-subject design.

Forundersøkelsen var et ledd i utviklingen av undervisningsmaterieell og tester til effektstudien av inferenstrening. Særlig hadde forundersøkelsen til hensikt å eksponere elevene for tekster og oppgaver hvor det var påkrevd å gjøre inferenser. Dette for vinne kunnskap om hva slags hjelp til å gjøre inferenser elevene hadde behov for, og som det dermed var hensiktsmessig å inkludere i inferenstreningen. Eksperimentbetingelsen innebar en manipulering med tekster og oppgaver som hadde til hensikt å hjelpe elevene med å gjøre utfyllende inferenser. Undersøkelsen kan karakteriseres som en mikro effektstudie, i den forstand at effekten ble målt på de samme tekstene som hjelpen ble gitt til. Målet med forundersøkelsen var å fange opp eventuelle effekter av eksperimentbetingelsen ved å sammenligne skårene til elevene under de to betingelsene.

### **6.2 Mål og problemstillinger**

Forundersøkelsen prøver ut en hypotese om at situasjonsmodeller har betydning for i hvilken grad leseren gjør inferenser. Mer konkret søker forundersøkelsen svar på om hjelp til å konstruere situasjonsmodeller har innflytelse på evne til gjøre utfyllende inferenser hos elever på niende trinn. Utfyllende inferenser blir definert i tråd med definisjonen i kapittel 3. Med situasjonsmodeller menes leserens forestillinger av et scenario eller en situasjon beskrevet i en tekst.

For å undersøke betydningen av situasjonsmodeller for utfyllende inferens ble deltakerne i forundersøkelsen eksponert for en eksperimentbetingelse og en



kontrollbetingelse. Eksperimentbetingelsen innebar at elevene svarte på instruksjonsspørsmål som hadde til hensikt å hjelpe dem med å etablere situasjonsmodeller om innholdet i tekstene. Kontrollbetingelsen besto i å lese de samme tekstene, men uten instruksjonsspørsmålene. Elevene skulle både under eksperimentbetingelsen og kontrollbetingelsen svare på spørsmål fra teksten som krevde at de gjorde utfyllende inferenser på bakgrunn av situasjonsmodeller. Etter som deltakerne under eksperimentbetingelsen ble eksponert for instruksjonsspørsmålene, var forventningen at elevene skulle stille opp situasjonsmodeller. Hensikten var da å se hvor stor betydning situasjonsmodellene har for å gjøre inferenser. I tillegg skulle elever i begge betingelser svare på faktaspørsmål som krevde at de hentet ut informasjon som sto eksplisitt i teksten. I forkant av forundersøkelsen ble det formulert følgende problemstillinger:

*4) Er det en generell effekt av hjelp til å etablere situasjonsmodeller på elevenes svar på inferensspørsmål?*

*5) Er det en spesifikk effekt av hjelp til å etablere situasjonsmodeller?*

*6) Har betingelse betydning for elevenes skåre på faktaspørsmål?*

*7) Har betingelse betydning for elevenes rapportering av vanskelighetsgrad på inferensspørsmål, faktaspørsmål eller tekst?*

Problemstilling 4 søker svar på hva den generelle effekten av hjelp til å etablere situasjonsmodeller via instruksjonsspørsmål er på skåre på inferensspørsmål. Denne problemstillingen vil bli undersøkt ved å sammenligne totalskåre på inferensspørsmålene under de to betingelsene. Problemstilling 5 dreier seg om den spesifikke effekten av instruksjonsspørsmålene, med andre ord om det er en tendens til at elever som har svart riktig på et instruksjonsspørsmål i større grad svarer riktig på det tilhørende inferensspørsmålet enn øvrige inferensspørsmål. Problemstilling 6 søker svar på hvorvidt hjelp til å etablere situasjonsmodeller har betydning for elevenes skåre på faktaspørsmål, det vil si spørsmål som krever at elevene henter ut informasjon som står eksplisitt i teksten. Problemstilling 7 søker svar på hvorvidt elevene opplever vanskelighetsgraden på tekst og spørsmål forskjellig under de to betingelsene.

## 6.3 Metode

### 6.3.1 Design

I forbindelse med forundersøkelsen ble det konstruert seks inferenstekster som ble delt inn i to oppgavesett, oppgavesett A og oppgavesett B. Elevene som deltok i studien, ble fordelt på to grupper, gruppe 1 og gruppe 2. Gruppe 1 fikk oppgavesett A som eksperimentbetingelse, og sett B som kontrollbetingelse. Omvendt fikk gruppe 2 oppgavesett A som kontrollbetingelse og sett B som eksperimentbetingelse. På denne måten inngikk alle elever i begge betingelser (se tabell 11). Tekstene ble rullert slik at halvparten av elevene i begge gruppene fikk eksperimentbetingelsen først og kontrollbetingelsen til slutt, og den andre halvparten fikk kontrollbetingelsen først og eksperimentbetingelsen til slutt.

Tabell 11: Oversikt over eksperimentbetingelse og kontrollbetingelse

Tekster	Eksperimentbetingelse	Kontrollbetingelse
Sett A	Gruppe 1	Gruppe 2
Sett B	Gruppe 2	Gruppe 1

Begrunnelsen for dette designet er først og fremst praktisk. Et av målene ved forundersøkelsen var å teste ut materiell til inferenstreningen. For å få testet ut mest mulig materiell ble to sett med tekster brukt. Det var ikke ønskelig å bruke for mye tid på forundersøkelsen, derfor var det nødvendig å begrense antallet deltakere, samtidig som de planlagte analysene krevde et visst deltakerantall.

### 6.3.2 Deltakere

Deltakerne var 84 elever fra fire klasser på niende trinn. Elevene kom fra to skoler i to kommuner i Rogaland. 44 av deltakerne var gutter og 37 jenter, tre personer hadde ikke oppgitt kjønn. Både gruppe 1 og gruppe 2 hadde elever fra begge skoler. Gruppe 1 besto av 41 elever (20 jenter og 20 gutter, 1 elev hadde ikke oppgitt kjønn) og gruppe 2 besto av 43 elever (17 jenter og 24 gutter, 2 elever hadde ikke oppgitt kjønn). Elevene og deres foresatte ble informert om målet og omfanget av undersøkelsen per brev som ble distribuert av skolen. I brevet ble det også informert om muligheten for å reservere seg fra undersøkelsen.

### **6.3.3 Måleinstrumenter**

#### **6.3.3.1 Screening prøve**

Elevenes leseforståelse og avkodningsferdigheter ble målt med delprøver fra en screening prøve (Nasjonalt læremiddelsenter, 1997) for niende klasse. Prøven ble standardisert med et utvalg på 1202 åttendeklasseelever i 1996. Da tiårig grunnskole ble innført i Norge, ble prøven tatt i bruk på niende klasse uten ny standardisering. Prøven har en "kritisk grense" for hver delprøve. Elever som skårer under den kritiske grensen, karakteriseres som en risikogruppe. Prøven har normtabeller med "risikogruppe", "mellomgruppe" og gruppen som får "alt rett".

Leseforståelse ble målt med to tekstoppgaver. Den ene teksten, "Tårer lagt i jorden", er en fiksjonstekst. Den andre teksten, "Kroppens klokke", kan karakteriseres om en nonfiksjonstekst, denne teksten inneholder mange faktaopplysninger. Begge delprøvene besto av flervalgsoppgaver hvor elevene ble bedt om å krysse av ett av flere mulige svaralternativer. 14 var høyest oppnåelig poengsum på hver av delprøvene. For fiksjonsteksten var kritisk grense satt til 7 feil. 20.1 % av utvalget i standardiseringen hadde 7 feil eller mer og ble dermed kategorisert som risikogruppe. Mellomgruppen med mellom 1 og 6 feil utgjorde 56.8 %, mens 23.1 % hadde alt riktig. Standardiseringsutvalgets gjennomsnittlig skåre på fiksjonsteksten var 10.6 riktige. For nonfiksjonsteksten var kritisk grense også satt til 7 feil eller mer, 17.1 % av standardiseringsutvalget hadde 7 feil eller mer og ble dermed kategorisert som risikogruppe. Mellomgruppen med mellom 1 og 6 feil utgjorde 64.9 %, mens 16.6 % hadde svart riktig på alle 14 spørsmål. Standardiseringsutvalgets gjennomsnittlig skåre på nonfiksjonsteksten var 10.4 riktige.

I utvalget i forundersøkelsen skåret 28.6 % av elevene i risikogruppen på fiksjonsteksten, mens 50 % av elevene skåret i mellomgruppen, og 21.4 % fikk alt rett. Gjennomsnittlig skåre på fiksjonsteksten i forundersøkelsen var 9.60. På nonfiksjonsteksten skåret 23.8 % av elevene i risikogruppen, 58.3 % i mellomgruppen, mens 17.9 fikk alt rett. Den gjennomsnittlige skåren til elevene i forundersøkelsen var 9.83. I forundersøkelsen var korrelasjonen mellom de to tekstene så høy (.75) at de ble slått sammen til en samlevariabel, heretter referert til som leseforståelse.

På ordavkodingsprøven fikk elevene se et bilde, ved siden av bildet var det flere alternative skrivemåter til ord som beskrev bildet, elevene skulle krysse av for rett skrivemåte. Eksempelvis var det bilde av et par strikkepinner, ved

siden av bildet skulle elevene krysse av for rett skrivemåte hvor svaralternativene var følgende: sdrikkepinner – strikkepinnre – strikkegenser – strikkepinner. Ordavkodingsprøven hadde 32 item, og 32 var høyeste oppnåelige poengsum. Kritisk grense var satt til 9 feil eller mer på ordavkodingsprøven, 18.8 % av standardiseringsutvalget hadde 9 feil eller mer og ble dermed kategorisert som risikogruppe. Mellomgruppen med mellom 1 og 8 feil utgjorde 55.5 %, mens 25.7 % hadde svart riktig på alle de 32 oppgavene. Standardiseringsutvalgets gjennomsnittlige skåre på ordavkodning var 27.6 riktige.

13.1 % av elevene i forundersøkelsen hadde skårer som tilsvarte risikogruppen på ordavkodning, 53.6 % skåret i mellomgruppen og 33.3 % av elevene fikk alt rett på ordavkodning. Forundersøkelsens utvalg hadde en gjennomsnittlig skåre på 28.36 på ordavkodning.

### **6.3.3.2 Ordforråd**

Ordforråd ble målt med to prøver, semantisk og fonologisk ordforrådsprøve. Prøvene hadde til hensikt å måle elevenes passive ordforråd, og har henholdsvis fonologiske og semantiske distraktorer. To danske ordforrådsprøver (<http://laes.hum.ku.dk>) ble brukt som modell for prøvene. Prøvene ble laget for denne undersøkelsen og er derfor ikke standardisert, denne undersøkelsen var snarere en utprøving av prøvene. Deltakerne fikk lest opp et ord eller en forklaring, deretter ble de bedt om å krysse av for hvilket av tre alternative ordene som betydde det samme som det første ordet eller forklaringen. Begge prøvene hadde 30 item, og 30 var høyeste oppnåelige poengsum for hver av prøvene. Da prøvene ble prøvd ut i forundersøkelsen, hadde utvalget en høyreskjev fordeling på begge prøvene. Etter skala analyser av prøvene, ble til sammen 14 item tatt ut. Det var to grunner til at item ble tatt ut. For det første at alternativene viste seg å være for dårlige, hvilket innebar at elevene kun hadde brukt ett eller to av de tre alternativene. For det andre lav item total korrelasjon, hvilket innebærer at itemet har lite til felles med øvrige item. Etter at itemene ble slettet, besto den fonologiske prøven av 22 item og den semantiske av 24 item. Cronbach's alfa var .78 for den fonologiske ordforrådsprøven og .79 for den semantiske prøven. Etter skalaanalysene ble utvalget nærmere en normalfordeling på begge ordforrådsprøvene. Korrelasjonen mellom de to prøvene var høy, .75, derfor ble de to variablene slått sammen. Variablen ordforråd er samlet skåre på de to prøvene.

### 6.3.3.3 Inferenstekster

To oppgavesett med til sammen seks inferenstekster inngikk forundersøkelsen. Oppgavesett A besto av tekstene 1, 2 og 3, oppgavesett B besto av tekstene 4, 5 og 6. Tekst 5 og 2 kan karakteriseres som fiksjonstekster, mens de øvrige fire tekstene kan karakteriseres som nonfiksjonstekster. De to oppgavesettene ble administrert under to betingelser, eksperimentbetingelse og kontrollbetingelse. Instruksjonen til eksperimentbetingelsen var: "Les det første avsnittet og svar på to spørsmål (instruksjonsspørsmål). Les deretter videre og svar på spørsmålene til slutt." Kontrollbetingelsen innebar at elevene leste teksten uten hjelp til å konstruere situasjonsmodeller. Instruksjonen kontrollbetingelsen var: "Les teksten nøye to ganger og svar på spørsmålene til slutt." De to spørsmålene som elevene i eksperimentbetingelsen svarte på etter å ha lest første avsnitt, blir heretter referert til som instruksjonsspørsmål. Teksten under er eksempel på første avsnitt i tekst 6. Denne teksten med spørsmål er gjengitt i sin helhet i appendiks sammen med instruksjon til forundersøkelsen.

#### *Norske emigranter*

*På 1800- og 1900-tallet utvandret mer enn 800 000 nordmenn til Amerika. Den første skuten som fraktet norske emigranter heiste seil i 1825. De første emigrantskipene var relativt små, senere foregikk transporten av emigranter med større skip som fullriggere og barker. Overfarten til Amerika kunne være en påkjenning for passasjerene. Mangel på vann begrenset mulighetene for renhold, toalettene om bord var ofte veldig urenlige. Passasjerene sov svært tett, og lopper og lus spredde seg raskt fra køye til køye.*

Elevene som fikk teksten "Norske emigranter" (tekst 6) som eksperimentbetingelse fikk følgende to instruksjonsspørsmål. Instruksjonsspørsmålene skulle hjelpe de med å etablere situasjonsmodeller om transportmidlene som emigrantene benyttet til overfarten til Amerika og om forholdene om bord i transportmidlene:

*Ordene seil, fullrigger og bark forteller deg om skutene som ble brukt til å frakte emigranter til Amerika. Har du en fellesbetegnelse på skutene?*

*Ordene begrenset mulighetene for renhold, urenlige, lopper og lus forteller om hvorfor reisen kunne være en påkjenning for passasjerene. Hva besto denne påkjenningen i?*

De to instruksjonsspørsmålene skulle sette elevene på sporet av at emigrantene kom seg til Amerika ved hjelp av seilskuter og at det var uhygieniske eller urenlige forhold om bord på skutene. Dette var nødvendig

for å forstå resten av teksten og for å kunne svare på inferensspørsmålene i etterkant av teksten. Etter å ha svart på de to instruksjonsspørsmålene, leste elevene i eksperimentbetingelsen resten av teksten. Elevene i kontrollbetingelsen leste første og påfølgende avsnitt sammenhengende.

*Det var vanlig at emigrantene hadde med, og tilberedte mat selv. Maten måtte tåle lagring så lenge reisen varte. Den vanligste provianten besto derfor av saltet eller tørket mat som for eksempel spekekjøtt og flatbrød. Emigrantene ble også oppfordret til å ha med gryter, bestikk og servise. Daglige rasjoner med vann og ved var vanligvis inkludert i billetten.*

*Under reisen var passasjerene opptatt av daglige gjøremål, det kunne være matlaging, strikking, lek eller spill. Søndager var det vanlig at kapteinen holdt gudstjeneste på dekket. Dårlig vær satte imidlertid begrensninger for aktivitetene.*

*Emigrantene fryktet storm og dårlig vær, skjønt de ville helst også unngå lange perioder med stille vær. Uvær gjorde selvfølgelig reisen ubehaglig, men i de første femti år som utvandringen til Amerika pågikk, var det faktisk ingen som omkom i forlis eller havari. Likevel var det en del emigranter som døde før de nådde Amerika.*

Til hver av tekstene var det seks spørsmål som alle elevene, både under eksperimentbetingelsen og kontrollbetingelsen, svarte på. Det første av de seks spørsmålene var: "Hvor vanskelig synes du teksten var?". Her skulle elevene angi vanskelighetsgrad ved å sette en strek på en linje fra lett til vanskelig. Denne linjen var ikke nummeret, men ble skåret på en skala fra 1 til 10 i etterkant. Spørsmål 2 og 3 var spørsmål om innholdet i teksten som krevde at elevene gjorde utfyllende inferenser. Disse to spørsmålene krevde at elevene hadde konstruert situasjonsmodeller på bakgrunn av det første avsnittet elevene i eksperimentbetingelsen hadde fått hjelp til gjennom instruksjonsspørsmålene. Begge disse spørsmålene var åpne. Følgende to inferensspørsmål ble gitt til teksten "Norske emigranter":

- 2. Hvorfor ville emigrantene unngå stille vær?*
- 3. Hva tror du var årsaken til at del emigranter døde før de nådde Amerika?*

For å kunne svare riktig på inferensspørsmål 2, at emigrantene ville unngå stille vær fordi da tok reisen lengre tid, måtte elevene ha forestillinger om transportmiddelet, seilskutene. Det første instruksjonsspørsmålet hadde til hensikt å hjelpe elevene med å danne nettopp denne forestillingen. På samme måte er situasjonsmodellen om de uhygieniske forholdene om bord nødvendig for å gjøre inferensen om at en del emigranter døde av sykdommer før de nådde Amerika. Inferensspørsmål 3 var altså knyttet til det andre instruksjonsspørsmålet. Det fjerde spørsmålet var en rangering av hvor

vanskelig elevene syntes de to inferensspørsmålene 2 og 3 var, på samme måte som i spørsmål 1.

Spørsmål 5 var et faktaspørsmål om innholdet i teksten. Spørsmålet krevde ikke at elevene gjorde inferens, det var kun nødvendig å huske informasjon som var eksplisitt uttrykt i teksten. Dette var også et åpent spørsmål.

*5. Hvem pleide å holde gudstjeneste om bord i emigrantskipene?*

Faktaspørsmålene ble inkludert for undersøke om en eventuell effekt av eksperimentbetingelsen kun var knyttet til elevenes skåre på inferensspørsmål, eller om eksperimentbetingelsen også hadde betydning for i hvilken grad elevene klarte å hente ut informasjon som står eksplisitt i teksten.

Det siste spørsmålet var rangering av vanskelighetsgrad på spørsmål 5, på samme måte som i spørsmål 1 og 4. Alle de 6 spørsmålene ble gitt på en separat side, elevene hadde ikke lov til å bla tilbake til teksten mens de svarte på spørsmål.

Ett av spørsmålene i tekst 2 i oppgavesett A ble utelatt fra analysene fordi det i etterkant av forundersøkelsen gikk tydelig fram at elevene kunne besvare spørsmålet riktig uten å bruke informasjonen fra det tilhørende instruksjonsspørsmålet. Dette spørsmålet inngikk heller ikke i totalskårene til oppgavesett A. Fordi oppgavesett A inneholdt ett spørsmål mindre enn oppgavesett B, ble begge totalskårene omregnet til procenter.

#### **6.3.4 Gjennomføring**

Elevene ble testet i grupper på mellom 7 og 12 elever med prøver i ordforråd, avkoding og leseforståelse. I tillegg fikk elevene to oppgavesett med til sammen seks inferenstekster med spørsmål til, som beskrevet i avsnittet over. Alle elevene i gruppe 1 og gruppe 2 leste alle de til sammen seks tekstene, men gruppene leste tekstene under eksperimentbetingelse og kontrollbetingelse, se tabell 11 for en oversikt.

Data fra forundersøkelsen ble analysert med versjon 15 av programvaren Statistical Package for the Social Sciences (SPSS).

## 6.4 Resultater

### 6.4.1 Beskrivende analyser

I dette avsnittet presenteres først deskriptiv statistikk for alle variabler både for utvalget samlet og for gruppe 1 og gruppe 2. Deretter vil det bli gjort rede for korrelasjoner mellom variablene. Dette for å gi oversikt over datamaterialet og forutsetninger for videre analyser.

#### 6.4.1.1 Deskriptive analyser

I tabell 12 er gjennomsnitt, standardavvik, skewness og kurtosis oppgitt for gruppe 1 og gruppe 2. Forskjellene mellom de to gruppene ble beregnet ved hjelp av t-test for uavhengige grupper. Størrelsen på forskjellene mellom gruppene blir i tabellen under uttrykt via kvadrert eta, som ble beregnet ved hjelp av følgende formel (Pallant, 2005, s. 208):

$$\eta^2 = \frac{t^2}{t^2 + (N1 + N2 - 2)}$$

Ifølge Cohens (1988, s. 286-287) retningslinjer for tolkning av kvadrert eta, kan  $\eta^2 = .0099$  betraktes som liten,  $\eta^2 = .00588$  som middels og  $\eta^2 = .1379$  som sterk effekt. Ifølge Byrne (1998) kan skårer som har skewness mindre enn 2.00 og kurtosis mindre enn 7.00 betraktes som normalfordelte. Videre kan skårer med skewness mellom 2.00 og 3.00 og kurtosis mellom 7.00 og 21.00 karakteriseres som moderat ikke-normale, mens skårer med skewness over 3.00 og kurtosis over 21.00 betraktes som ekstremt ikke-normale. Både leseforståelse og ordavkoding var noe høyreskjeve i begge grupper, men i henhold til Byrnes retningslinjer for tolkning av skewness og kurtosis var alle variablene normalfordelte i begge gruppene.

Det var en del missing på inferenstekstene, mens det bare var noen få missing på øvrige prøver. Missing ble erstattet med 0. Også på ordforrådsprøvene ble ikke-besvarte oppgaver skåret som 0. En elev i gruppe 2 hadde ikke svart på noen av oppgavene på de to ordforrådsprøvene, denne eleven ble definert som missing i samlevariabelen ordforråd. I tillegg til ubesvarte oppgaver ga noen få elever nonsenssvar på en eller flere oppgaver. Med nonsenssvar menes respons som ikke er svar på oppgaven, men kommentarer til undersøkelsen, kommentarer som uttrykker at deltakeren ikke kan eller vil svare på oppgaven, eller andre kommentarer eller tegninger som ikke har relevans for undersøkelsen eller oppgavene. Denne typen respons ble skåret som 0 poeng.



Noen elever tegnet/klusset eller skrev inn kommentarer i tillegg til å svare på oppgavene.

Det var signifikant forskjell mellom gruppene på variabelen ordavkodning i gruppe 2 sin favør. Denne forskjellen kan karakteriseres som moderat. Det var ingen signifikante forskjeller mellom gruppene på ordforråd eller leseforståelse. Det var heller ingen signifikante forskjeller mellom gruppene med hensyn til samleskårer på inferensspørsmål eller faktaspørsmål i de to oppgavesettene med inferenstekster.

Elever i begge grupper rapporterte at tekstene i sett A var vanskeligere enn tekstene i sett B, det var ikke signifikant forskjell mellom gruppenes rangering av tekstenes vanskelighetsgrad, verken i oppgavesett A eller B. Videre rapporterte begge grupper at både inferensspørsmålene og faktaspørsmålene i oppgavesett A var vanskeligere enn i oppgavesett B. Heller ikke med hensyn til vanskelighetsgrad på inferensspørsmål og faktaspørsmål var det signifikant forskjell mellom gruppene. Da hele utvalget ble analysert samlet, viste t-test for avhengige grupper at elevene rapporterte signifikant høyere vanskegrad på både tekstene  $t(83)=2.35$ ,  $p=.021$ ,  $\eta^2=.06$ , faktaspørsmålene,  $t(83)=3.99$ ,  $p<.001$ ,  $\eta^2=.16$ , og inferensspørsmålene,  $t(83)=3.08$ ,  $p=.003$ ,  $\eta^2=.10$ , i oppgavesett A enn oppgavesett B. Elevene rapporterte også at inferensspørsmålene var signifikant vanskeligere enn faktaspørsmålene både i oppgavesett A,  $t(83)=9.65$ ,  $p<.001$ ,  $\eta^2=.53$ , og oppgavesett B,  $t(83)=7.56$ ,  $p<.001$ ,  $\eta^2=.41$ . Analysene ble gjentatt med Wilcoxon test for matchede par, som bekreftet resultatene fra t-testene.

Betydningen av situasjonsmodeller for inferens – en forundersøkelse

Tabell 12: Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler for gruppe 1 og gruppe 2, samt forskjeller mellom gruppene

Mål	Gruppe 1 (N=41)				Gruppe 2 (N=43)				$\eta^2$	P
	M	SD	Skew.	Krt.	M	SD	Skew.	Krt.		
Ordavkoding	27.32	5.0	-1.09	0.34	29.35	3.11	-1.27	1.54	=.06	=.029
Ordforråd <sup>3</sup>	30.59	6.54	0.11	-0.98	29.74	8.70	-0.22	-0.91	<.01	=.618
Leseforståelse	19.41	6.93	-0.66	-0.27	19.44	7.57	-0.95	0.07	<.01	=.391
Inferensspørsmål, sett A <sup>4</sup>	39.51	28.10	0.22	-0.94	37.21	29.15	0.45	-0.65	<.01	=.714
Inferensspørsmål, sett B <sup>4</sup>	34.55	27.48	0.48	-0.51	41.09	24.49	0.03	-0.76	=.02	=.253
Faktaspørsmål, sett A	1.85	1.01	-0.30	-1.10	1.93	0.910	-0.45	-0.58	<.01	=.716
Faktaspørsmål, sett B	2.15	0.937	-0.69	-0.68	2.23	0.841	-0.73	-0.46	<.01	=.658
Vanskelighetsgrad inferensspørsmål, sett A	14.43	5.89	-0.45	0.72	15.40	6.03	-0.23	-0.71	<.01	=.458
Vanskelighetsgrad inferensspørsmål, sett B	13.12	5.89	0.61	0.40	11.10	5.14	-0.36	0.71	=.01	=.098
Vanskelighetsgrad faktaspørsmål, sett A	8.11	5.52	0.59	-0.24	10.21	6.02	0.82	0.92	=.03	=.100
Vanskelighetsgrad faktaspørsmål, sett B	6.71	5.78	1.93	5.18	7.60	5.81	0.68	-0.18	=.03	=.485
Vanskelighetsgrad tekst, sett A	11.07	5.67	0.14	-0.41	11.45	6.49	0.13	-0.13	<.01	=.773
Vanskelighetsgrad tekst, sett B	9.77	5.84	1.53	2.92	9.93	5.34	0.18	0.08	<.01	=.895

<sup>3</sup> N=42 for gruppe 2 på ordforråd

<sup>4</sup> Skåren på inferensspørsmålene er prosent riktige.

Kjønnsforskjeller ble kartlagt ved hjelp av t-test for uavhengige grupper. Det var ingen signifikante kjønnsforskjeller med hensyn til ordforråd, ordavkoding eller leseforståelse. I forhold til inferensspørsmålene til inferenstekstene rapporterte jentene (N=37, M=13.22, SD=5.71) signifikant høyere vanskegrad på tekstene til oppgavesett A enn guttene (N=44, M=9.70, SD=6.13),  $t(79)=2.65$ ,  $p=.010$ ,  $\eta^2=.08$ . På øvrige oppgaver til inferenstekstene var det ingen signifikante kjønnsforskjeller.

Tabell 13 viser prosentvis hvor mange elever i gruppe 1 og gruppe 2 som fikk riktig og feil svar på instruksjonsspørsmål til tekstene i henholdsvis oppgavesett A og B. I tillegg viser tabellen prosent av missing for hvert instruksjonsspørsmål. Gruppe 2 skåret elevene generelt riktig på flere oppgaver enn elevene i gruppe 1.

Tabell 13: Gruppens frekvens i prosent for instruksjonsspørsmålene for instruksjonsspørsmålene til tekstene i oppgavesett A og B

		Gruppe 1					Gruppe 2		
		Frekvens i prosent					Frekvens i prosent		
Tekst	Spm.	Feil	Riktig	Missing	Tekst	Spm.	Feil	Riktig	Missing
1	I	61.00	22.0	17.00	4	I	37.20	58.10	4.70
1	II	65.90	19.50	14.60	4	II	29.30	61.00	9.70
2	I	31.70	56.10	12.20	5	I	27.90	69.80	2.30
2	II	48.80	41.50	9.70	5	II	4.70	95.30	0.00
3	I	51.20	41.50	7.30	6	I	62.80	27.90	9.30
3	II	29.30	61.00	9.70	6	II	34.90	58.10	7.00

#### 6.4.1.2 Korrelasjoner mellom variablene

Forholdet mellom variablene ble undersøkt med Pearson produkt moment korrelasjon (Pearson's r). Styrken på Pearson's r tolkes i henhold til Cohens (1988, s. 79-80) retningslinjer, hvor .10 tilsvarer liten korrelasjon, .30 tilsvarer middels korrelasjon og .50 sterk korrelasjon. Korrelasjoner mellom ordforråd, ordavkoding, leseforståelse og inferensspørsmål og faktaspørsmål fra inferenstekstene er oppsummert i tabell 14.

Prøvene i ordforråd, ordavkoding og leseforståelse korrelerte sterkt. Det var også en sterk korrelasjon mellom inferensspørsmål i oppgavesett A og B. Inferensspørsmål i både oppgavesett A og B korrelerte sterkt med både leseforståelse og ordforråd, og moderat med ordavkoding. Faktaspørsmål og inferensspørsmål korrelerte sterkt, mens korrelasjonene mellom faktaspørsmål og øvrige variabler var moderat.

Tabell 14: Oversikt over korrelasjoner

	1	2	3	4	5	6	7
1 Ordforråd	-						
2 Ordavkoding	.48	-					
3 Leseforståelse	.75	.64	-				
4 Inferensspørsmål sett A	.61	.37	.49	-			
5 Inferensspørsmål sett B	.49	.33	.52	.56	-		
6 Faktaspørsmål sett A	.42	.25	.32	.53	.46	-	
7 Faktaspørsmål sett B	.31	.26	.34	.48	.63	.39	-

Signifikansnivået var  $p < .05$  for alle variabler

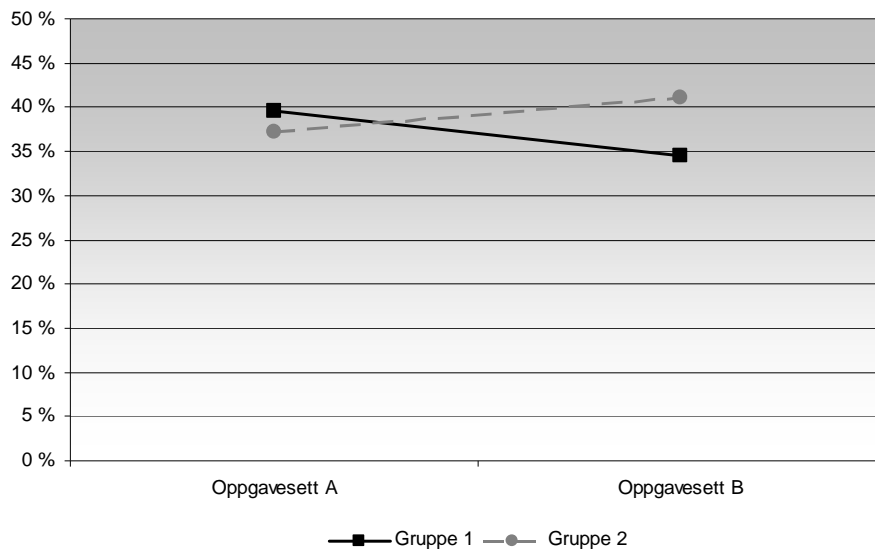
Korrelasjonen mellom totalskåre på instruksjonsspørsmålene og totalskåre på inferensspørsmålene i oppgavesett A og B ble undersøkt ved Pearson's  $r$ . Analysene ble gjort separat for gruppe 1 og for gruppe 2. For gruppe 1 var korrelasjonen mellom skåre på instruksjonsspørsmålene og inferensspørsmålene i oppgavesett A (eksperimentbetingelse) .72, mens korrelasjonen mellom skåre på instruksjonsspørsmålene og inferensspørsmålene i oppgavesett B (kontrollbetingelse) var noe lavere, .42. Analysene ble gjentatt med Spearman's  $\rho$ , som viste tilsvarende resultat. Da det ved hjelp av partiell korrelasjonsanalyse ble kontrollert for leseforståelse i korrelasjonen mellom instruksjonsspørsmål og inferensspørsmål, sank korrelasjonen til .63 for skåre på instruksjonsspørsmål og inferensskåre i oppgavesett A, og .26 for skåre på instruksjonsspørsmål og inferensskåre i oppgavesett B. For gruppe 2 var korrelasjonen mellom skåre på instruksjonsspørsmålene og inferensspørsmålene i oppgavesett A (kontrollbetingelsen) .66, mens korrelasjonen mellom skåre på instruksjonsspørsmålene og inferensspørsmålene i oppgavesett B (eksperimentbetingelse) var nesten tilsvarende, .60. Også her viste Spearman's  $\rho$  tilsvarende resultat. Da det ble kontrollert for leseforståelse, sank korrelasjonen mellom skåre på instruksjonsspørsmål og inferensskåre i oppgavesett A til .51. Korrelasjonen mellom instruksjonsspørsmål og inferensskåre i oppgavesett B sank til .34 da det ble kontrollert for leseforståelse.

#### 6.4.2 Effekter av eksperimentbetingelser

I dette avsnittet redegjøres det for de to gruppens resultater på inferensspørsmålene under de to betingelsene. For det første redegjøres det for den generelle effekten av eksperimentbetingelsen på inferensspørsmålene. Deretter undersøkes betydningen av elevenes respons på instruksjonsspørsmålene for inferensspørsmålene. Til slutt gjøres det rede for de to gruppens skårer på faktaspørsmål under de to betingelsene.

#### 6.4.2.1 Generell effekt av eksperimentbetingelsen på inferensspørsmålene?

Som det gikk fram av tabell 12, var det ikke signifikante forskjeller mellom gruppe 1 og gruppe 2 på samleskårer for inferensspørsmål verken i oppgavesett A eller B. Kvadrert eta var svært lav for begge oppgavesettene. Effekter av eksperimentbetingelsen ble også beregnet ved hjelp av Cohen's  $d$ , jevnfør avsnitt 5.3.2. Cohen's  $d$  ble beregnet til ubetydelige .08 og .25 for henholdsvis gruppe 1, som hadde oppgavesett A som eksperimentbetingelse og gruppe 2 som hadde oppgavesett B som eksperimentbetingelse. Figur 2 viser de to gruppenes resultater på inferensspørsmålene i oppgavesett A og B.



Figur 2. Gruppens samlede skåre på inferensspørsmål i oppgavesett A og B

T-test for avhengige grupper ble gjennomført for hver av de to gruppene for å undersøke om gruppens gjennomsnitt var forskjellige under eksperimentbetingelsen og kontrollbetingelsen. For gruppe 1 viste t-test at gjennomsnittsskåren ikke var signifikant høyere ved eksperimentbetingelsen (oppgavesett A) enn ved kontrollbetingelsen (oppgavesett B),  $t(40)=1.127$ ,  $p=.266$ ,  $\eta^2=.03$ . Gruppe 2 sin skåre under eksperimentbetingelsen (oppgavesett B) var heller ikke signifikant høyere enn under kontrollbetingelsen (oppgavesett A),  $t(42)=-1.137$ ,  $p=.262$ ,  $\eta^2=.03$ . Resultatene ble bekreftet av Wilcoxon test for matchede par.

#### 6.4.2.2 Spesifikk effekt av instruksjonsspørsmålene for skåre på inferensspørsmålene?

For å undersøke om det var en spesifikk effekt av instruksjonsspørsmålene på inferensspørsmålene til samme tekst, ble det beregnet korrelasjoner ved hjelp av Pearson's  $r$ . Analysene ble gjentatt med Spearman's  $\rho$ , som viste tilsvarende resultat. Samlet skåre på de to instruksjonsspørsmålene til hver tekst ble beregnet. Det ble også regnet ut samlet skåre på de to inferensspørsmålene til hver tekst, med unntak for tekst 2 hvor det ene inferensspørsmålet ble tatt ut. Deretter ble totalskåre av de to instruksjonsspørsmålene for hver tekst korrelert med totalskåre på de to inferensspørsmålene for hver tekst. Dette for å undersøke om korrelasjonen mellom instruksjonsspørsmål og inferensspørsmål til samme tekst var sterkere enn korrelasjoner mellom instruksjonsspørsmål og inferensspørsmål på tvers av tekster. Korrelasjonsanalysene er oppsummert i tabell 15, korrelasjoner mellom instruksjonsspørsmål og inferensspørsmål i samme tekst er uthevet.

Tabell 15: Korrelasjon mellom instruksjonsspørsmål og inferensspørsmål

Inferensspørsmål	Instruksjonsspørsmål					
	Gruppe 1 (N=41)			Gruppe 2 (N=43)		
	Tekst 1	Tekst 2	Tekst 3	Tekst 4	Tekst 5	Tekst 6
Tekst 1	<b>.51**</b>	.40*	.44**	.51**	.09	.56**
Tekst 2	.21	<b>.52**</b>	.44**	.25	-.07	.43**
Tekst 3	.41**	.44**	<b>.54**</b>	.23	.14	.49**
Tekst 4	.30	.40*	.20	<b>.60**</b>	.13	.52**
Tekst 5	.06	.07	.11	-.09	<b>-.04</b>	.06
Tekst 6	.45**	.21	.28	.35*	.14	<b>.50**</b>

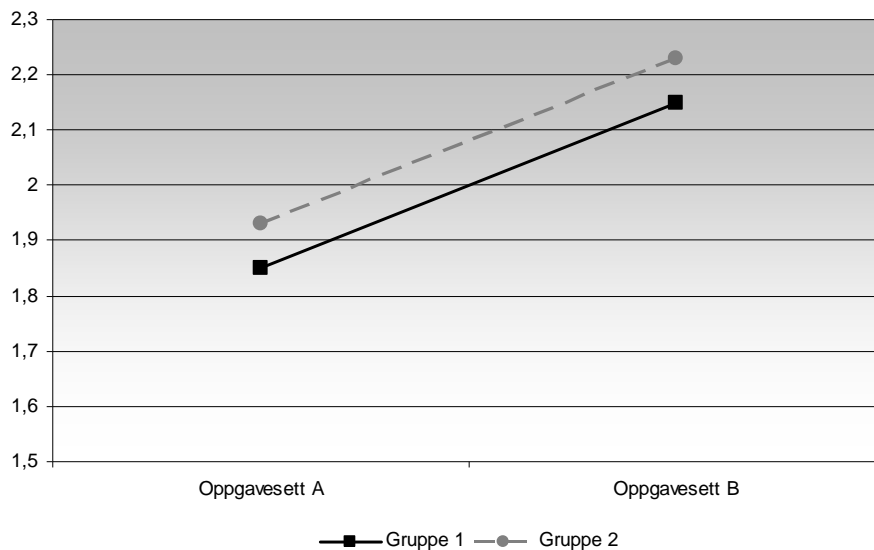
\*\* Korrelasjonen er signifikant på .01 nivå

\* Korrelasjonen er signifikant på .05 nivå

For gruppe 1 var korrelasjonene mellom instruksjonsspørsmål og inferensspørsmål i samme tekst noe høyere enn korrelasjonene på tvers av tekster. I gruppe 2 var det også sterk korrelasjon mellom instruksjonsspørsmål og inferensspørsmål i tekst 4 og 6. Tekst 5 skilte seg derimot ut fra øvrige tekster ved at det ikke var signifikant korrelasjon mellom instruksjonsspørsmålene og inferensspørsmålene i denne teksten. Frekvensanalyser av spørsmålene til tekst 5 viste store tak- og gulveffekter. I alt 69.8 % av elevene i gruppe 2 hadde svart rett på det første instruksjonsspørsmålet, mens hele 95.3 % hadde svart riktig på instruksjonsspørsmål 2. På inferensspørsmålene hadde kun 9.3 % av elevene i gruppe 2 svart riktig på det ene spørsmålet, mens 53.7 % hadde svart riktig på det andre inferensspørsmålet. Det var heller ingen signifikante korrelasjoner mellom inferensspørsmålene i tekst 5 og instruksjonsspørsmål i noen av de andre tekstene.

#### 6.4.2.3 Effekt av eksperimentbetingelsen på faktaspørsmålene?

Som det går fram av tabell 12, var det ikke signifikante forskjeller mellom de to gruppene med hensyn til faktaspørsmål til tekstene i de to oppgavesettene. Med andre ord var det ingen effekt av eksperimentbetingelsen på elevenes skårer på faktaspørsmålene. Dette går tydelig fram av figur 3 under som viser gruppenes resultater på faktaspørsmålene i oppgavesett A og B. Begge grupper hadde høyest skåre på faktaspørsmålene til oppgavesett B.



Figur 3. Gruppens samlede skårer på faktaspørsmål i oppgavesett A og B

## 6.5 Diskusjon

### 6.5.1 Problemstillinger og resultater

To av problemstillingene som denne forundersøkelsen søkte svar på, var hvorvidt det var generell og spesifikk effekt av å gi elever hjelp til å etablere situasjonsmodeller for skåre på inferensspørsmål. Resultatene viste at det ikke var en generell effekt av hjelpen til å etablere situasjonsmodeller på skåre på inferensspørsmål. Det var heller ingen tydelig tendens til spesifikk effekt av instruksjonsspørsmålene.

Nullfunnet i denne undersøkelsen reiser spørsmålstegn ved materialet som ble brukt. For å kunne få effekter av hjelp til å etablere situasjonsmodeller må elevene ha et reelt behov for hjelp. Dette betyr at tekstene og inferensspørsmålene må være så vanskelige at elevene faktisk trenger hjelp til

å skape situasjonsmodeller. Derfor kan de ikke-signifikante resultatene skyldes at tekstene ikke var vanskelige nok, og at elevene derfor ikke hadde problemer å etablere situasjonsmodeller uten hjelp fra instruksjonsspørsmålene. Det var mange elever som lot instruksjonsspørsmålene stå ubesvart. På enkelte item var det opp til 17 % missing. Fordi instruksjonsspørsmålene sto på samme side som overskriften og første avsnitt av teksten, er det er lite trolig at missing skyldes at elevene lot oppgavene stå ubesvart fordi de hadde bladd over siden med spørsmålene, eller av annen grunn ikke hadde oppdaget dem. Dersom elever lot et eller flere instruksjonsspørsmål stå ubesvart fordi spørsmålene var for vanskelige, ville det også ha vært vanskelig å svare på inferensspørsmålene. Dette fordi inferensspørsmålene nettopp krever at leseren har etablert den situasjonsmodellen som instruksjonsspørsmålet har til hensikt å utløse. Den høye andelen missing gjør at man ikke vet hvor mange av elevene som faktisk stilte opp situasjonsmodeller.

Hensikten med instruksjonsspørsmålene var å hjelpe elevene med å etablere situasjonsmodeller som var nødvendig for å finne svar på inferensspørsmålene. Tekst 5 fungerte ikke etter intensjonen. Som det går fram av tabell 15, var det ingen signifikant korrelasjon mellom instruksjonsspørsmålene og inferensspørsmålene i denne teksten. Selv om en stor andel av elevene hadde svart riktig på begge instruksjonsspørsmålene, var det svært få elever som hadde svart riktig på det ene av de to inferensspørsmålene til tekst 5. Dette kan skyldes flere forhold, for eksempel at instruksjonsspørsmålene ikke var tilstrekkelige for å etablere den nødvendige situasjonsmodellen, eller at inferensspørsmålet var for vanskelig i den forstand at det ikke var nært nok knyttet til situasjonsmodellen.

Ubetydelige effekter av eksperimentbetingelsen kan skyldes at instruksjonsspørsmålene ikke var formulert på en hensiktsmessig måte. Et alternativ til åpne instruksjonsspørsmål kunne vært flervalgsspørsmål. Denne tilnærmingen kunne trolig hjelpe elevene med å etablere situasjonsmodeller. Problemet med flervalgsspørsmål er at man ikke kan vite om deltakerne virkelig bruker instruksjonsspørsmålene som hjelp til å etablere modeller, eller om det er svaralternativene som er utslagsgivende for at deltakeren skaper situasjonsmodellen. I så fall vil svaralternativet kun være til hjelp i prøvesituasjonen, og ikke være en trening for deltakeren til å skape situasjonsmodeller i reelle lesesituasjoner. Problemet vil være tilsvarende ved å bruke illustrasjoner som hjelp til å etablere situasjonsmodeller.



Ubetydelige effekter av eksperimentbetingelsen kan også skyldes at instruksjonsspørsmål ikke var rette måten å gi hjelp til å etablere situasjonsmodeller på. Et alternativ til instruksjonsspørsmål kunne vært å framstille deler av innholdet i teksten i en grafisk modell. En slik grafisk modell må hjelpe elevene med å identifisere relevant eksplisitt informasjon i teksten som er nødvendig for å danne situasjonsmodeller, samtidig må modellen også synliggjøre at teksten inneholder implisitt informasjon. Videre må den grafiske modellen understreke sammenhengen mellom situasjonsmodellen og inferenser som kan gjøres ved å integrere situasjonsmodellen og andre deler av teksten.

Forundersøkelsen ble gjennomført som en engangssamling. Elevene svarte på instruksjonsspørsmålene, men fikk ikke trening i å svare på denne typen spørsmål eller forklaringer på hvordan instruksjonsspørsmålene skulle gjøre det lettere å forstå teksten og svare på inferensspørsmålene. Det er sannsynlig at elevene trengte mer trening på å svare på denne typen spørsmål og mer forståelse for hensikten med instruksjonsspørsmålene, og at dette er grunnen til at effektene uteble.

Et annet ankepunkt ved denne forundersøkelsen var vanskegraden på materialet som ble brukt. Elevene rapporterte forskjellig vanskegrad på tekster og spørsmål i de to oppgavesettene uavhengig av gruppe. Videre var det flere elever som hadde riktige svar på instruksjonsspørsmålene i gruppe 2 enn i gruppe 1. Dette kan tolkes på to måter. For det første kan dette bety at elevene i gruppe 2 var flinkere til å danne situasjonsmodeller enn elevene i gruppe 1. For det andre kan det være sannsynlig at instruksjonsspørsmålene i oppgavesett A, som gruppe 1 svarte på, var vanskeligere enn instruksjonsspørsmålene i oppgavesett B. Dette gjør det vanskelig å sammenligne både betingelsene og de to gruppene. Siden det ikke var effekt av instruksjonsspørsmålene verken når oppgavesett A eller B ble brukt som eksperimentbetingelse, er det ingen grunn til å tro at verken forskjell i vanskegrad i oppgavesettene eller forskjell mellom gruppene var årsaken til det ikke-signifikante funnet.

Svaret på problemstilling 6 blir at det ikke var signifikant effekt av eksperimentbetingelsen på elevenes skåre på faktaspørsmål. Dette var ikke overraskende siden det ikke var effekt på inferensspørsmålene. Problemstilling 7 søkte svar på hvorvidt betingelse har betydning for elevenes rapportering av vanskegrad på tekst og spørsmål. Resultatene viste at det ikke var forskjell i elevenes rapportering under de to betingelsene. Dette betyr at elevene verken syntes teksten og spørsmålene ble lettere eller vanskeligere på

grunn av instruksjonsspørsmålene, hvilket er nok en indikasjon på at eksperimentbetingelsen ikke fungerte etter intensjonen.

### **6.5.2 Erfaringer fra forundersøkelsen**

Forundersøkelsen har gitt erfaringer med tanke på å utarbeide materiell til bruk i undervisningen av inferens i skolen. For det første må elevene ha et reelt behov for hjelp til å gjøre inferenser. Dersom elevene må bruke tid på å jobbe med innholdet i teksten, både det som står eksplisitt og det som er implisitt, har de sjansen til å reflektere over hva de gjør når de forstår eller ikke forstår en tekst. Videre vil de også kunne rette fokus mot hvilken informasjon i teksten som er nøkkelen til å forstå teksten. Dette krever at tekstene er så vanskelige at elevene faktisk trenger hjelp.

Ikke bare må tekstene som skal brukes til å trene på inferenser være utfordrende, og hjelpen til å gjøre inferenser må også være egnet. Denne undersøkelsen prøvde ut effekten av instruksjonsspørsmål som hadde til hensikt å utløse en situasjonsmodell på inferensspørsmål. I foregående avsnitt ble grafiske modeller skissert som et alternativ til instruksjonsspørsmålene. Dette må være grafiske modeller som har fokus på inferensen, men som samtidig understreker behovet for å danne situasjonsmodeller. Videre må elevene også få hjelp i tilstrekkelig omfang. Antakeligvis ga forundersøkelsen ikke elevene tilstrekkelig innblikk i hensikten med hjelpen, det vil si instruksjonsspørsmålene de fikk. Trolig var instruksjonsspørsmålene heller ikke tilstrekkelig for å gjøre elevene oppmerksomme på at de selv må tilføye teksten kunnskap til opplysninger i teksten for å få en sammenhengende forståelse.

Denne forundersøkelsen har også synliggjort utfordringer knyttet til måling av effekt. Prøvene i avkoding og leseforståelse som ble brukt i denne undersøkelsen, var noe høyreskjeve, hvilket innebærer takeffekter eller tendenser til takeffekter. I denne undersøkelsen ble variablene leseforståelse og avkoding kun brukt som kontrollvariabler. I en mer omfattende studie av inferens vil det være naturlig å inkludere leseforståelse som et mål på effekt. I tillegg bør effekten måles på flere tidspunkter. Da blir det avgjørende med at målet på leseforståelse er tilstrekkelig vanskelig, slik at det er rom for framgang, noe som er nødvendig for å kunne måle eventuelle effekter.

Forundersøkelsen har også avdekket utfordringer knyttet til utvalg og alder på deltakere. Det var flere tilfeller av nonsenssvar i denne undersøkelsen. Hva slike nonsenssvar er uttrykk for, er ikke lett å vite, men det er sikkert at slike svar ikke gir informasjon om effekter av en gitt betingelse. I forbindelse med

debatten rundt Norges resultater på PISA-undersøkelsene stiller Sjøberg (2008) spørsmålsmålstegn ved hvorvidt norske 15-åring yter sitt beste på undersøkelser som for eksempel PISA. Han viser til at elevene ikke får karakterer eller annen tilbakemelding på resultatene på undersøkelsen. Deltakerne i forundersøkelsen var omtrent på samme alder som deltakerne i PISA, og heller ikke her ble det gitt tilbakemelding på elevenes prestasjoner. Derfor kan det også i forhold til forundersøkelsen være tvil om hvorvidt deltakerne bidro med full innsats. Kanskje er niendeklassinger en lite gunstig aldersgruppe for denne typen undersøkelser? Dersom undersøkelser av denne typen i større grad inkluderes i den ordinære undervisningen, kan det være at elevene i større grad opplever den anvendte tiden som fruktbar, noe som kan øke sjansen for at elevene yter sitt beste.

## **7. Effekter av inferenstrening blant norske sjetteklassinger**

### **7.1 Mål og problemstillinger**

Som det går fram av kapittel 5, har flere tidligere studier vist varierende effekt av inferenstrening. Ingen av disse studiene ble gjennomført i Norge eller Skandinavia, og dermed kan ikke resultatene uten videre overføres til norske forhold. Denne studien har til hensikt å undersøke effekter av inferenstrening blant norske elever. Målet og motivasjonen for studien er både teoretisk og pedagogisk. Teoretisk søker studien å undersøke forholdet mellom inferens og leseforståelse. Ved å benytte et kvasiekperimentelt design kan studien bidra til å bekrefte årsaksretningen i forholdet mellom inferens og leseforståelse. Studien fokuserer på utfyllende inferenser og tar utgangspunkt i definisjonen av utfyllende inferenser i kapittel 3. Å gjøre inferenser innebærer prosesser som foregår mentalt og som det ikke er lett å sette ord på. Å lære andre hvordan man gjør inferenser når en leser, er derfor en utfordring. Denne utfordringen er utgangspunktet for studiens pedagogiske mål, nemlig å utvikle materiell til bruk i inferenstrening. Det ble utviklet materiell til åtte økter inferenstrening à ca. 20 minutter som ble prøvd ut i en eksperimentgruppe høsten 2008.

Undersøkelsen ble gjennomført blant elever på sjette trinn. Dette trinnet ble valgt ut fra erfaringer fra forundersøkelsen og ut fra inferenstreningens innhold. I forundersøkelsen var det flere elever som svarte med nonsenssvar på oppgavene. Denne erfaringen tatt i betraktning syntes det mer hensiktsmessig å legge effektstudien til et lavere klassetrinn i håp om at yngre elever i mindre grad ville oppgi nonsenssvar på prøvene. På en annen side var det ikke hensiktsmessig å inkludere for unge barn fordi studien fokuserer på utfyllende inferens, som stiller krav til leserens evne til å tenke abstrakt i en slik grad som ikke kan forventes av de yngste elevene. En annen erfaring fra forundersøkelsen som ble særlig vektlagt i inferenstreningen, var elevenes bevissthet om at de selv må bidra med egen kunnskap for å gjøre inferenser på bakgrunn av opplysninger i teksten. Dette var et moment som forundersøkelsen ikke kunne vurdere.

I forkant av effektstudien ble det formulert følgende problemstillinger for å konkretisere de overordnede målene:

8) *Hvilken effekt har inferenstrening via grafiske modeller på elevenes:*

*a) evne til å gjøre inferenser?*

*b) leseforståelse?*

9) *Hva skjer med eventuelle effekter etter at inferenstreningen er avsluttet?*

10) *Er det forskjell i eventuell endring i inferens og leseforståelse hos:*

*a) elever med sterk, gjennomsnittlig og svak leseforståelse?*

*b) elever med gode, gjennomsnittlige og svake ferdigheter i inferens?*

*c) gutter og jenter?*

*d) tospråklige og enspråklige?*

11) *I hvilken grad kan ordforråd, ordavkodning og nonverbale ferdigheter forklare variasjon i individuell endring i inferens og leseforståelse?*

Problemstilling 8 søkte svar på hvorvidt det var effekt av inferenstrening på elevenes evne til å gjøre inferenser og deres leseforståelse. Effekten ble målt ved å sammenligne eksperimentgruppens endring i skårer før og etter inferenstreningen med en kontrollgruppe. Å undersøke effekter av inferenstreningen på inferens er avgjørende for å kunne slå fast om elevene faktisk er blitt bedre i det de har trent på. Det er i forhold til inferens at man kan forvente de største effektene. For å finne svar på årsaksforbindelsen mellom leseforståelse og inferens ble også generaliserte effekter av inferenstreningen på leseforståelse undersøkt. Problemstilling 9 fokuserer på langtidseffekter av inferenstreningen.

Som det er gjort rede for i avsnitt 2.2.5, har barn med svak leseforståelse tross adekvate avkodingsferdigheter særlig problemer med å gjøre inferenser. Derfor vil studien også fokusere på hvordan elever som i utgangspunktet har ulike ferdigheter i inferens og leseforståelse profitterer på treningen, jevnfør problemstilling 10 a og b. Den norske PIRLS-undersøkelsen (Solheim et al., 2007) viste at det var kjønnsforskjeller i lesing blant norske fjerdeklassinger, derfor var det naturlig å undersøke eventuelle kjønnsforskjeller også i denne studien. I effektstudien er eventuelle kjønnsforskjeller interessante på to måter. For det første å undersøke om en eventuell effekt av inferenstreningen er lik for både gutter og jenter, eller med andre ord om gutter og jenter profitterer likt på treningen. For å studere dette sammenlignes gutter i eksperimentgruppen med gutter i kontrollgruppen og jenter i eksperimentgruppen med jenter i kontrollgruppen. For det andre er det interessant å undersøke hvorvidt inferenstreningen utjevner eventuelle kjønnsforskjeller. Dette kan undersøkes ved å studere skårene til gutter og

jenter i eksperimentgruppen før og etter inferenstrening. PIRLS-undersøkelsen viste også at minoritetsspråklige elever generelt skåret lavere på lesing enn jevnaldrende majoritetsspråklige (Ogle, Begnum, & Solheim, 2008). Derfor var det interessant å undersøke om det også var forskjeller med hensyn til språklig bakgrunn blant elevene i effektstudien, og om det var forskjell i effekt hos enspråklige og tospråklige elever.

En eventuell endring i inferens og leseforståelse kan skyldes andre ferdigheter og betingelser av betydning for leseforståelse. I denne studien undersøkes betydningen av ordforråd, ordavkodning og nonverbale ferdigheter for individuell endring i inferens og leseforståelse. Hvorfor nettopp disse tre variablene ble valgt, blir nærmere begrunnet i avsnittet om kontrollvariabler i avsnitt 7.3.3.2.

## **7.2 Undervisning i inferens**

For å kunne trene inferens ble et materiell utviklet med hensikt å forbedre sjettetrinns elever ferdigheter i å gjøre inferenser når de leser. Materialet hadde spesiell fokus på utfyllende inferens. Inferenstreningen hadde til hensikt å synliggjøre hva inferens er, samt at elevene skulle få øve på å gjøre inferenser med tekster som ble laget spesielt til dette formålet. Inferenstreningen var lagt opp som 8 leksjoner à 20 minutter. De 20 minuttene skulle være ren undervisningstid, eksklusivt tid som eventuelt gikk med til for eksempel å komme til ro etter friminutt. De 8 leksjonene skulle gjennomføres i løpet av en periode på fire til fem uker. I de to påfølgende avsnittene gjøres det rede for henholdsvis arbeidsformen og materialet som ble benyttet i inferenstreningen.

### **7.2.1 Arbeidsform og rammer for inferenstreningen**

For å finne fram til en hensiktsmessig arbeidsform og passende rammer for inferenstreningen ble tidligere studier som har vist positiv effekt av undervisning i leseforståelse, gjennomgått. De siste tiårene er det utviklet flere undervisningsmetoder og programmer som har til hensikt å fremme barns leseforståelse. Mange av metodene forsøker å bevisstgjøre elevene i forhold til hva det innebærer å lese med forståelse, og hvordan de kan overvåke sin egen leseforståelse. Flere av undervisningsmetodene har sitt opphav i resiprok undervisning som ble utviklet av Palincsar og Brown (1984) tidlig på 1980-tallet. Karakteristisk for resiprok undervisning er at en samtaleleder guider en gruppe elever gjennom fire strategier for leseforståelse, oppsummering, spørsmålsgenerering, oppklaring og foregripelse. Samtalelederen begynner med å oppsummere tekstens hovedpunkter, deretter stiller samtalelederen elevene spørsmål om innholdet i en tekst som gruppen har lest. Etter hvert

stiller også elevene seg selv og andre elever spørsmål knyttet til teksten. Videre følger oppklaring av eventuelle misforståelser og vanskelige ord. Den siste strategien, foregripelse, handler om at barna skal gjøre antakelser om hva som kommer til å skje videre i teksten. Oppgaven til samtalelederen, som gjerne er en lærer, er å fungere som modell og se til at alle i elevgruppen inkluderes i samtalen. Resiprok undervisning er inspirert av Vygotskys tanker om sonen for den nærmeste utvikling. Eleven deltar først som tilskuer, og får gradvis mer ansvar.

Mot slutten av 1980-tallet utviklet Pressley og hans kolleger et undervisningsprogram i leseforståelse som ble kalt Transactional Strategies Instruction (TSI) (se for eksempel Pressley, 1998; Pressley et al., 1992). TSI har flere fellestrekk med resiprok undervisning, for eksempel at lærer modellerer og forklarer bruk av lesestrategier. Pressley og kolleger understreker likevel at det er klare forskjeller mellom de to tilnærmingene til instruksjon i leseforståelse. Undervisningens struktur og rigidhet slik den kommer til uttrykk gjennom lærer, er den største forskjellen mellom de to tilnærmingene. I TSI har lærer en mindre synlig rolle, og undervisningsformen har mindre faste rammer enn i resiprok undervisning.

Guthrie og Wigfield med kolleger har også utviklet en undervisningsform som har til hensikt å fremme leseforståelse (se for eksempel Guthrie, Wigfield, Barbosa et al., 2004; Guthrie, Wigfield, & Perencevich, 2004). Denne undervisningsformen, Concept-Oriented Reading Instruction (CORI), har i likhet med de to ovennevnte undervisningsformene fokus på strategier for leseforståelse og formidlingen av disse. Guthrie og Wigfield med kolleger framhever i tillegg at "reading engagement", eller leseengasjement, er en betingelse for tilegnelse av strategier og leseforståelse. Målet med CORI er å skape engasjerte lesere, hvor engasjerte lesere forstås som kognitivt kompetente ved at de leser strategisk. Videre er de er motiverte og ønsker å lære, er kunnskapsdrevet og sosialt interaktive. Hovedtrekkene i CORI-undervisningen er å fremme motivasjon gjennom bruk av interessante tekster, bruk av konkrete og følbare erfaringer, og ved å organisere samarbeid mellom elevene. Videre gis det eksplisitt undervisning gjennom modellering av strategier som for eksempel aktivering av bakgrunnskunnskap, stille spørsmål, søke informasjon og grafisk organisering.

Effekter av de ovennevnte undervisningsformene er godt dokumentert, både når det gjelder resiprok undervisning (for en oversikt se Palincsar, 2003; Rosenshine & Meister, 1994), TSI (se for eksempel V. Anderson, 1992; Brown, Pressley, van Meter, & Schuder, 1996) og CORI (se for eksempel

Guthrie, Wigfield, Barbosa et al., 2004; Wigfield et al., 2008). Andreassen (2008) har kombinert prinsipper fra alle de tre ovennevnte undervisningsformene i det han kaller eksplisitt leseforståelsesundervisning (ELU). ELU ble prøvd ut blant norske femteklassinger i en periode på et halvt år. Andreassen fant ingen effekter av ELU på leseforståelse, men hevder årsakene til at effektene uteble blant annet kan skyldes at prøvene som ble brukt til å måle leseforståelse, ikke var følsomme nok til å fange opp effekter. Videre hevder Andreassen at fravær av effekter også kan skyldes at enkelte kjerneprinsipper i ELU ikke ble godt nok implementert i klassene.

Inferenstreningen som prøves ut i denne studien, bygger på noen av de samme prinsippene som de ovennevnte undervisningsformene. Det mest sentrale er å gjøre elevene bevisste på hva leseforståelse og inferens er og hvordan man oppnår leseforståelse.

Undervisningsoppleggene som er beskrevet over, er prøvd ut med både lærere og forskere som instruktører. Rosenshine og Meister (1994) fant i en sammenfatning av 16 studier av resiprok undervisning at det ikke spilte noen rolle for effekt hvorvidt undervisningen ble gjennomført av lærer eller forsker. I denne undersøkelsen ble inferenstreningen gjennomført av de respektive norsklærerne til klassene som deltok. Dette ble vurdert som mest hensiktsmessig med tanke på det pedagogiske målet om å utvikle materiell til bruk i leseundervisning.

I utprøving av tidligere undervisningsopplegg i leseforståelse varierer undervisningens varighet og intensitet veldig, fra noen få økter til flere måneders trening. Rosenshine og Meisters (1994) sammenfatning viste at effekter av resiprok undervisning ikke var avhengig av antall undervisningsøkter. Antall økter inferenstrening i denne studien ble satt til 8, ideelt gjennomført i en periode på fire til fem uker. Omfanget til øktene ble begrenset til ca. 20 minutter. Denne tidsbegrensningen av inferenstreningen var først og fremst praktisk begrunnet. En betingelse for at inferenstreningen skulle kunne administreres av lærerne, var at treningen måtte være gjennomførbar innenfor rammene for ordinær undervisning i norsk.

Felles for mange typer undervisningsopplegg som har til hensikt å fremme barns leseforståelse, er aktivisering gjennom samtale og diskusjon, enten i hel klasse eller i mindre grupper. van Keer (2004) har undersøkt effekten av undervisning i leseforståelsesstrategier i både små grupper på kun noen få elever, aldersblandede grupper, samt i hel klasse. Undervisningen besto av det van Keer karakteriserer som eksplisitt strategiinstruksjon kombinert med



samtale og diskusjon rundt tekst. I undervisningen ble det lagt vekt på følgende aktiviteter: Aktivisering av kunnskap om tekstens tema, gjøre forutsigelser om tekstens innhold, finne hovedideene i teksten, monitorering og regulering av forståelsen og klassifisering av tekstsjanger og tilpasse lesingen deretter. van Keer fant ingen signifikante forskjeller mellom de tre gruppene med hensyn til barnas utbytte av undervisningen. Rosenshine og Meister (1994) fant tilsvarende resultat for resiprok undervisning, nemlig at gruppestørrelsen ikke spilte noen rolle for deltakernes utbytte av undervisningen. Med andre ord er det ikke forskningsbaserte holdepunkter for at gruppearbeid i seg selv vil bidra til å bedre elevenes leseforståelse. Derfor ble ikke gruppearbeid vektlagt i inferenstreningen i denne studien. De norske femteklassingene som deltok i ELU i Andreassens (2008) studie, arbeidet i grupper på 4 til 5 elever. Andreassen stilte seg imidlertid tvilende til kvaliteten på dette gruppearbeidet spesielt og gruppearbeid som metode for å arbeide med tekster på generelt. På en annen side framhever sentrale forskere som for eksempel Guthrie og Wigfield (Wigfield et al., 2008) med kolleger betydningen av motivasjon for leseforståelse. Guthries forskning på leseforståelse bygger blant annet på å fremme motivasjon gjennom arbeid i små grupper. Kompromisset ble derfor at elevene i inferenstreningen diskuterte inferensspørsmålene til tekstene i par eller i små grupper i flertallet av de åtte leksjonene med tanke på at dette ville kunne motivere elevene og gjøre leksjonene lystbetonte.

I utviklingen av materiell til studien ble det forsøkt å lage tekster som innholdsmessig var interessante for elevene i likhet med et av prinsippene i CORI. Hvorvidt en tekst er interessant eller ei, er naturligvis svært subjektivt, og det er vanskelig å anslå i hvilken grad elevene finner tekstene spennende og motiverende. I inferenstreningen ble det imidlertid brukt tekster med forskjellige tema som ble ansett som passende for aldersgruppen. På en annen side ble det også forsøkt å finne temaer til tekstene som var forholdsvis nøytrale med hensyn til elevenes erfaringer og interesser for å unngå at enkelte grupper elever hadde svært mye eller svært lite kunnskap om det gitte temaet.

Materiellet besto av korte tekster med oppgaver til (materiellet vil bli presentert under). Da nye oppgavetyper ble introdusert for elevene, ble oppgavene gjennomgått i fellesskap i klassen. De fleste av tekstene som ble brukt i inferenstreningen, leste elevene stille hver for seg. Da alle elevene hadde lest teksten, hjalp læreren elevene med å oppklare vanskelige ord eller andre uklarheter i tekstene. Elevene jobbet i par eller små grupper og diskuterte spørsmålene sammen i de fleste øktene, mens i de siste øktene

jobbet elevene alene med oppgavene. Alle økter ble avsluttet med en felles gjennomgang av svarene på oppgavene.

### **7.2.2 Materiell**

Materiellet som ble brukt i inferenstreningen, besto av oppgavene ”Hva forteller setningene” og tekster med og uten grafiske modeller. ”Hva forteller setningene” ble gjennomført som en introduksjon til inferenstreningen i leksjon 1 og 2, mens elevene arbeidet med tekstene gjennom resten av leksjonene, leksjon 2 til 8.

#### **7.2.2.1 ”Hva forteller setningene”**

”Hva forteller setningene” hadde til hensikt å gjøre elevene klar over at lesing innebærer mer enn avkoding. I forhold til de fire komponentene i inferenstrening som ble identifisert i kapittel 5, kan oppgavene ”Hva forteller setningene” betraktes som en bevissthetsfremmende komponent. Oppgavene besto av setninger med spørsmål til. Svarene til spørsmålene sto ikke direkte i setningene. Poenget var at elevene skulle bli oppmerksomme på at de må trekke slutninger ut over det som faktisk står i setningene for å kunne svare på spørsmålet, se eksempel under.

*Ole Jensen fikk ikke låne penger i banken til å kjøpe seg båt. Han begynte å se seg om etter en ekstrajobb*

*Hvorfor ville Ole Jensen ha en ekstrajobb?*

For å finne svaret på oppgaven, at Ole Jensen trengte å tjene penger til å kjøpe båt, må leseren stille opp en situasjonsmodell og gjøre inferens på bakgrunn av informasjon fra begge setningene.

”Hva forteller setningene” hadde også til hensikt å gjøre elevene bevisste på hvilke ord som fikk dem til å gjøre inferenser, derfor ble elevene bedt om å markere hvilke ord som hjalp dem med å finne svar på spørsmålet, som i eksempelet under.

*Apparatet suger inn luft fra baksiden, varmer den opp, og blåser varm luft ut gjennom tuten foran. Apparatet finnes på de fleste norske bad.*

*Hva slags apparat beskrives over?*

*Sett en strek under ord i setningene over som hjalp deg med å finne svaret.*

Hvilke ord som leder fram til løsningen kan være noe individuelt, men ordene ”bad” og ”varm luft” er viktige ledetråder for å finne fram til rett svar,

(hår)føner, og for å skille føner fra andre lignende gjenstander som for eksempel aircondition-anlegg.

#### **7.2.2.2 Inferenstekster**

Etter at elevene var blitt introdusert for inferenstreningen, gikk elevene over fra å jobbe med setninger til å svare på inferensspørsmål knyttet til korte tekster. I alt ble 15 tekster med spørsmål, hvis svar krever at leseren gjorde utfyllende inferenser, brukt i inferenstreningen. I hver økt arbeidet elevene med to til tre inferenstekster. Hver tekst hadde mellom et og fire inferensspørsmål. Tekstene var fagtekster eller nonfiksjonstekster som handlet om temaer innenfor dyreriket, geografi, teknologi, historie og samfunnsfag. For at sen avkodning ikke skulle bli et hinder for elevene, var tekstene forholdsvis korte, mellom 100 og 200 ord. Fordi det var vanskelig å finne korte, meningsfulle tekster som det lot seg stille gode inferensspørsmål til, ble alle tekstene konstruert til inferenstreningen. Alle tekstene som ble brukt i treningen, ble laget med utgangspunkt i autentiske tekster fra for eksempel leksikon og faktabøker, men tekstene ble kortet ned og omskrevet.

Elevene fikk hjelp fra grafiske modeller til å svare på inferensspørsmålene. Hensikten med de grafiske modellene var å illustrere sammenhengen mellom informasjon i teksten og å gjøre elevene klar over at teksten inneholder ”meningsgap” som leseren selv må fylle. De grafiske modellene inneholdt tre ruter som enten var delvis utfylt, eller som skulle fylles ut helt av elevene. To av rutene skulle inneholde opplysninger som sto direkte i teksten og en rute skulle inneholde svaret på inferensspørsmålet. Svaret på inferensspørsmålet var basert på opplysningene fra teksten. Modellene hadde således som mål å konkretisere forholdet mellom informasjon i teksten. I forhold til komponentene i kapittel 5 hører de grafiske modellene inn i komponenten ”skape forbindelser”.

I leksjon 2 ble elevene introdusert for inferenstekster med samme type oppgaver som i ”Hva forteller setningene”, sammen med grafiske modeller. Deretter jobbet elevene med tekster med grafiske modeller. I de første tekstene fikk elevene oppgaver hvor de skulle svare på inferensspørsmål med hjelp fra en grafisk modell hvor to opplysninger fra teksten var fylt inn. Svaret på spørsmålet lå i mellom de to informasjonene fra teksten, og det var nettopp dette modellen hadde til hensikt å illustrere. Under er et eksempel på en tekst med inferensspørsmål fra leksjon 3.

### **Sjiraffen**

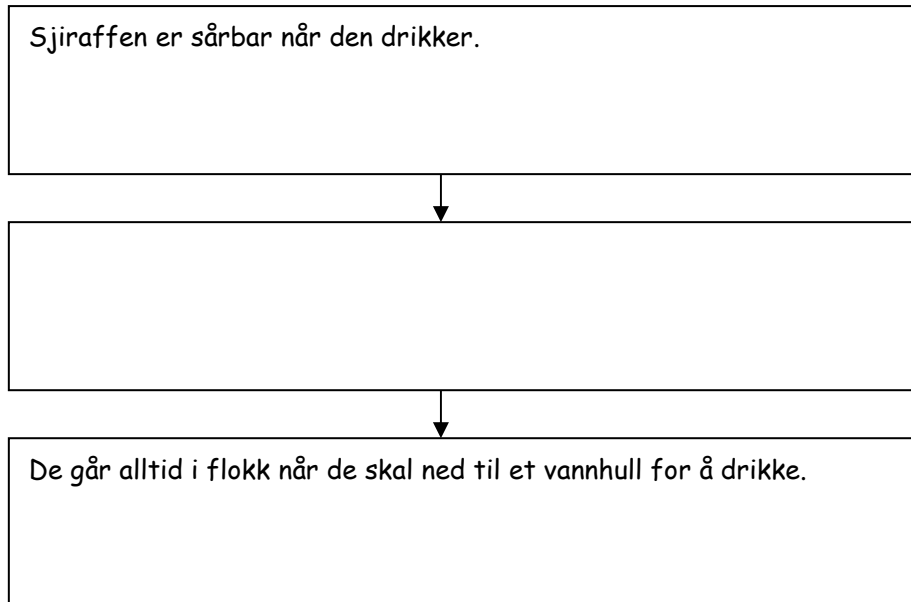
*Sjiraffen kan bli over fem meter høy og er med sin langstrakte kropp verdens høyeste dyr. Den spiser blader fra trær og forsyner seg gjerne av bladene høyt oppe i trekronene. De lange beina og den lange halsen gjør det vanskelig for sjiraffen å drikke. Sjiraffen er mest sårbar når den drikker, da kan den lett bli angrepet av rovdyr.*

*Sjiraffer finnes bare i Afrika, hvor de er spredt over et stort område. De er et vanlig syn på de åpne slettene sør for Sahara.*

*Sjiraffer holder sammen i flokker som i noen tilfeller kan bestå av opptil hundre dyr. De går alltid i flokk når de skal ned til et vannhull for å drikke.*

*Hvorfor går sjiraffene i flokk når de skal drikke?  
Skriv svaret på spørsmålet inn i den tomme ruta i midten.*

Den grafiske modellen til teksten "Sjiraffen" var delvis utfylt med to opplysninger som sto eksplisitt i teksten (se under). Før elevene gikk i gang med å svare på inferensspørsmålene, ble de bedt om å finne igjen informasjonen i øverste og nederste rute inne i teksten. Dette for å understreke overfor elevene at informasjonen i rutene sto eksplisitt i teksten. Deretter skulle elevene finne svar på inferensspørsmålet om hvorfor sjiraffene går i flokk når de skal drikke. Ved å bruke informasjonen om at sjiraffen er sårbar når den drikker, skulle elevene komme fram til at sjiraffene holder vakt for hverandre når de drikker, at de passer på hverandre, vokter hverandre for rovdyr eller lignende. Teksten "Sjiraffen" er gjengitt i appendiks i det format som elevene fikk i treningen.



Elevene fikk også fylle ut grafiske modeller der svaret på inferensspørsmålet var fylt inn, og hvor oppgaven var å finne hvilke ord eller setninger i teksten som ledet til riktig svar på inferensspørsmålet. Dette er for å understreke at man bruker informasjonen fra teksten til å svare på spørsmålene, selv om svaret ikke eksplisitt står i teksten. Ved å identifisere ord i teksten som fører til riktig svar, ble elevene nødt til å overvåke sin egen forståelse. Denne aktiviteten kan karakteriseres som en monitoreringskomponent i inferenstreningen, jevnfør kapittel 5. Under er eksempel på en tekst og en delvis utfylt grafisk modell hvor svaret på inferensspørsmålet var fylt inn i modellen.

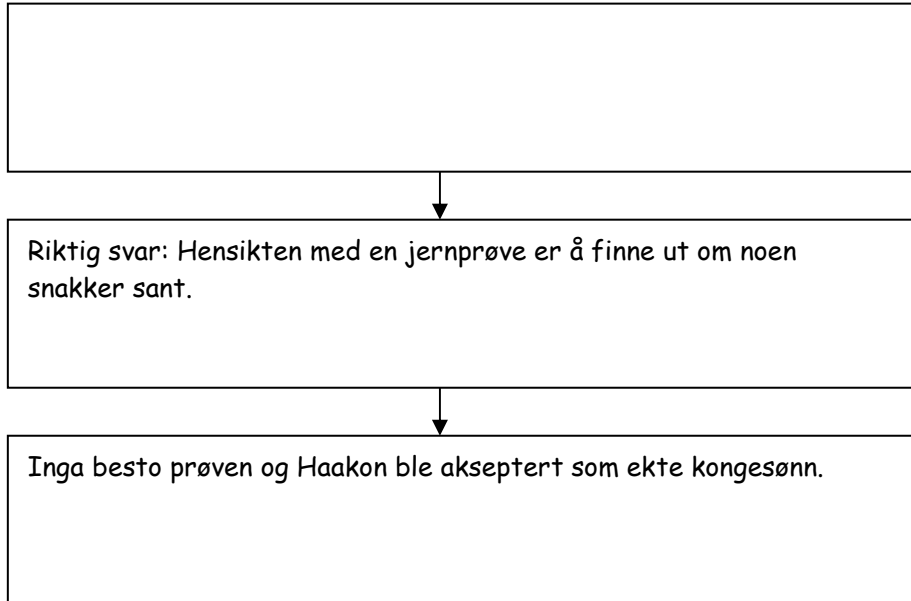
***Inga fra Varteig***

*Da kong Inge døde i 1217, ville birkebeinerne at Håkon Håkonsson skulle bli Norges nye konge. Håkon Håkonsson var sønn av kong Håkon Sverresson og Inga fra Varteig.*

*Det var mange som var uenige i at Håkon Håkonsson skulle bli ny konge i Norge. Grunnen til dette var at mange tvilte på at han virkelig var sønn av kong Håkon Sverresson. Håkon Håkonssons mor, Inga, påsto at sønnen var en ekte kongesønn. Ingas påstander ble av mange ansett som løgn. Uenigheten ble løst ved at Inga måtte gjennomføre en jernprøve. Prøven besto i å bære glovarmt jern uten å ta skade av det. Inga besto jernprøven og Haakon ble akseptert som ekte kongesønn.*

En jernprøve besto av å bære glovarmt jern uten å ta skade av det. Hva var hensikten med en slik prøve?

I den tomme ruta skal du skrive inn ord eller setninger fra teksten som leder til rett svar.



Her var oppgaven å finne ledetråder i teksten som ledet til svaret på inferensspørsmålet, for eksempel "Ingas påstander ble av mange ansett som løgn" eller "mange tvilte på at han virkelig var sønn av kong Håkon Sverresson".

Autentiske tekster inneholder ikke grafiske modeller. Målet med leksjonene var at elevene skulle gjøre inferenser også uten grafiske modeller. Modellene var kun et verktøy for å sette elevene på sporet av hvordan man gjør inferenser. Derfor skulle elevene etter hvert som de ble kjent med de grafiske modellene selv tegne opp modellene. Under er eksempel på en tekst hvor elevene skulle lage hele modellen selv.

### **Gresshopper**

*De aller fleste har en eller annen gang hørt gresshopper synge om sommeren. Gresshoppesangen er kraftig og høyfrekvent. Derfor må du ha god hørsel for å høre den. Egentlig er det ikke riktig å si at gresshoppene synger, for de bruker ikke stemmen når de synger. Gresshoppene har noen små tagger på innsiden av beina. Sangen, eller lyden, lager gresshoppene ved gni beina mot en skarp kant på vingene.*

*Gresshoppene lever av planter og har en kropp som ofte går i ett med omgivelsene. Derfor kan det være vanskelig å få øye på dem. Hvis du vil finne en gresshoppe, er det best å følge sangen. Men lytt til gresshoppens sang mens du er ung, for når du blir gammel, kan det hende at du ikke lenger kan høre den.*

*Hvorfor kan det hende at du ikke lenger kan høre gresshoppene når du blir gammel? Tegn opp en modell og skriv inn svaret og hvilke ord i teksten som hjalp deg med å finne svaret.*

Modellen skulle inneholde eksplisitt informasjon fra teksten. I dette tilfellet opplysningene ”derfor må du ha god hørsel for å høre den” og ”når du blir gammel, kan det hende at du ikke lenger kan høre den”. I tillegg skulle modellen inneholde svaret på inferensspørsmålet, at gamle mennesker hører dårligere enn unge.

I de to siste leksjonene skulle elevene svare på inferensspørsmål uten å fylle ut eller tegne opp grafiske modeller. I tillegg til å svare på inferensspørsmålet skulle elevene streke under ord i teksten som ledet til riktig svar på inferensspørsmålet.

### **7.2.3 Instruksjon til lærerne i eksperimentgruppen**

Materiellet ble introdusert for lærerne i eksperimentgruppen på møter som ble holdt skolevis. Her ble det gjort rede for hva inferenser er og hvordan studien hadde som mål å trene elevene i å gjøre inferenser. Eksempler på materiell fra setningsoppgavene, tekster med og uten grafiske modeller, ble gjennomgått. I tillegg fikk hver lærer i eksperimentgruppen en perm med en generell beskrivelse av inferenstreningen, mål og arbeidsmetode. I tillegg inneholdt permen også en kort beskrivelse av oppgaver og arbeidsform for hver leksjon, for eksempel hvilke oppgaver som skulle gjennomgås i fellesskap og hvilke oppgaver som elevene skulle løse i små grupper eller individuelt. Flere av tekstene inneholdt vanskelige ord, da ble lærerne bedt spesielt om å forklare disse ordene. Eksempel på instruksjon til leksjonen 3 er gjengitt i appendiks. I tillegg til beskrivelsen av hver leksjon inneholdt permen fasit til alle oppgaver med forslag til svar på oppgavene. Innholdet i de åtte leksjonene er oppsummert i tabellen under.

Tabell 16: Oversikt over innhold i leksjonene

Leksjon	1	2	3	4	5	6	7	8
Setningsoppgaver	X	X						
Tekst og grafiske modeller med informasjon fra teksten		X	X		X			
Tekst og grafiske modeller med svar på inferensspørsmålet				X				
Tekst, elevene tegnet og fylte ut grafiske modeller						X		
Tekst og spørsmål uten grafisk modell, i tillegg streket elevene under ord i teksten							X	X

## 7.3 Metode

### 7.3.1 Design

Designet i denne studien kan karakteriseres som kvasieksperiment med pre-, post- og utsatt posttest og kontrollgruppe (Shadish et al., 2002). Eksperimentgruppen fikk inferenstrening som beskrevet over, mens kontrollgruppen gjennomførte sin ordinære undervisning. Hele utvalget ble testet to ganger, ved pretest og posttest. I tillegg ble to kontrollklasser og to eksperimentklasser testet med en utsatt posttest fem uker etter at inferenstreningen var avsluttet. Designet er oppsummert i tabellen under.

Tabell 17: Studiens design

	Pretest	Inferenstrening	Posttest	Utsatt posttest
Eksperimentgruppe	X	X	X	X
Kontrollgruppe	X		X	X

Effekter av inferenstrening ble målt på tre nivåer (jevnfør avsnitt 4.1). Primær overført effekt ble målt med en eksperimentdesignet inferensprøve som inneholdt tilsvarende tekster og spørsmål som ble brukt i inferenstreningen, men uten grafiske modeller. Generalisert effekt av treningen ble målt med en ikke-eksperimentdesignet prøve i leseforståelse. I tillegg ble utsatt effekt, både primær overført og generalisert, målt hos to eksperimentklasser og to kontrollklasser fem uker etter avsluttet inferenstrening.

### 7.3.2 Deltakere

Siden denne studien la opp til innhenting av data på tre tidspunkter, samt oppfølging av lærere og klasser over lengre tid, var det nødvendig å rekruttere deltakerne fra et avgrenset geografisk område. Deltakere er derfor



skjønnsmessig valgt. Materiellet til inferenstreningen og prøver ble utarbeidet på bokmål, derfor ble kun bokmålsklasser inkludert i utvalget.

Åtte rektorer fra fire kommuner på Vestlandet ble kontaktet, hvorav seks skoler fra tre kommuner valgte å delta i studien. Kommune I er en bykommune, som ifølge Statistisk sentralbyrå sin indeks for levekårsproblemer er tilnærmet gjennomsnittlig for en norsk kommune (<http://www.ssb.no>). Gjennomsnittlig inntekt for husholdninger generelt og for husholdninger med mor/far med yngste barn i alderen 0 til 17 år i kommune I ligger noe over landsgjennomsnittet. Kommune II og III består av tettsteder og landlige strøk. Begge disse kommunene har ifølge Statistisk sentralbyrå noe mindre levekårsproblemer enn gjennomsnittet for landet for øvrig. Både i kommune II og III ligger gjennomsnittlig inntekt per husholdning noe over landsgjennomsnittet, mens inntekt for husholdninger mor/far med yngste barn i alderen 0 til 17 år ligger noe under landsgjennomsnittet i kommune II, og noe over landsgjennomsnittet i kommune III.

Skole B hadde elever fra første til tiende trinn, mens de resterende fem skolene kun hadde elever på barnetrinnet. Det totale elevtallet på skolene varierte fra 233 til 680 høsten 2008. Deltakelse i eksperimentgruppe og kontrollgruppe ble avgjort ved loddtrekning på skolenivå. Siden data fra eksperimentgruppen var planlagt brukt i flere analyser enn kontrollgruppen, var det mest hensiktsmessig å ha en større eksperimentgruppe enn kontrollgruppe. Etter trekningen besto eksperimentgruppen av ti klasser fordelt på fire skoler, mens kontrollgruppen besto av seks klasser fordelt på to skoler. Se tabellene under for en oversikt over skolene. Gjennomsnittlig elevtall på eksperimentskolene var 421, og 343 på kontrollskolene.

*Tabell 18: Skoler i utvalget*

Skole	Kommune	Gruppe
A	I	Eksperiment
B	II	Eksperiment
C	III	Eksperiment
D	I	Eksperiment
E	III	Kontroll
F	I	Kontroll

De 16 klassene<sup>5</sup> som inngikk i studien besto av 266 elever våren 2008, da informasjonsbrev om undersøkelsens mål og omfang ble sendt ut til foresatte. Av disse var det 236 elever hvis foresatte ga aktivt samtykke til å delta i undersøkelsen. En elev flyttet i løpet av høsten 2008, og deltok derfor kun på pretester, to elever var bortreist ved skolestart, og deltok derfor kun på posttestene. Enkelte elever var ikke til stede da prøver ble gjennomført, i noen tilfeller lot det seg ikke gjøre å teste disse elevene på et senere tidspunkt. Manglende dataverdier ble registrert for elever som ikke har deltatt på enkeltprøver. Antallet inkluderte deltakere oppgis for alle analyser.

Tabellen under viser oversikt over hvilke klasser som hørte til hvilke skoler, totalt antall elever i klassen, samt hvor mange elever i hver klasse som deltok i studien. Antall elever i klassen varierte fra 13 til 23 elever.

Tabell 19: Klasser i utvalget

Klasse	Skole	Antall elever i klassen totalt	Antall elever som deltok i studien
1*	A	15	14
2*	A	15	14
3*	A	14	13
4	B	22	18
5	B	19	9
6	B	13	11
7	C	15	14
8	C	17	15
9	D	23	23
10	D	21	20
11**	E	17	14
12**	E	17	12
13**	E	18	17
14**	E	15	14
15	F	15	14
16	F	14	14

\*Disse tre klassene hadde todeling i enkelte fag.

\*\* Disse fire klassene hadde tredeling i enkelte fag.

Utvalget totalt besto av 118 jenter og like mange gutter. Kjønnfordelingen var tilnærmet den samme i eksperimentgruppe og kontrollgruppe. Gjennomsnittlig alder beregnet ut fra pretest for inferensprøven og prøven i leseforståelse var 11,21 år for hele utvalget. Elevenes alder varierte fra 10 år og 5 måneder til 11 år og 10 måneder. Gjennomsnittlig alder for

<sup>5</sup> Klasse defineres her om en gruppe elever som har samme kontaktlærer og som er den inndelingen som undervisningen hovedsakelig foregår i.

eksperimentgruppe og kontrollgruppe var henholdsvis 11,20 og 11,23 år. 9,7 % av elevene var tospråklige, her er tospråklighet definert som at de snakker et annet morsmål enn norsk eller skandinavisk hjemme. Av de 23 tospråklige elevene var 15 gutter og 8 jenter.

Tabell 20: Kjønnfordeling, tospråklighet og alder i eksperimentgruppe og kontrollgruppe

	Eksperimentgruppe	Kontrollgruppe	Totalt
N	151	85	236
Jenter	75(49,7 %)	43 (50,6 %)	118 (50 %)
Gutter	76 (50,3 %)	42 (49,4 %)	118 (50 %)
Tospråklige	15 (9,4 %)	8 (10,3 %)	23 (9,7 %)
Alder	11,20	11,23	11,21

Eksperimentgruppen hadde 6 kvinnelige og 4 mannlige lærere, mens kontrollgruppen hadde 3 kvinnelige og 3 mannlige lærere. Omtrent halvveis i undersøkelsen gikk en av de mannlige lærerne i eksperimentgruppen ut i permisjon og ble erstattet av en kvinnelig vikar som gjennomførte deler av inferenstreningen. Per skolestart høsten 2008 varierte lærernes undervisningserfaring fra å ikke ha jobbet som lærer før, til å 23,5 års erfaring i yrket. Gjennomsnittlig undervisningserfaring for lærere i eksperimentgruppen var 10 år, og 8 år i kontrollgruppen. Ved skolestart høsten 2008 hadde en lærer i eksperimentgruppen vært kontaktlærer for sin klasse i 3 år, mens en lærer i kontrollgruppen hadde vært kontaktlærer i 2 år for sin klasse. Øvrige lærere i utvalget hadde vært kontaktlærere for sine klasser i ett år eller mindre.

### 7.3.3 Instrumenter

#### 7.3.3.1 Mål på effekt

##### *Inferensprøve*

Prøven i inferens var en gruppeprøve som hadde til hensikt å måle elevenes evne til å gjøre inferenser på bakgrunn av tekst, og skulle fange opp primær overført effekt av inferenstreningen. I instruksjonen til elevene ble det understreket at de skulle svare på spørsmålene ut fra det de hadde lest i teksten (se appendiks for fullstendig instruksjon). Tekstene og spørsmålene i prøven var derfor bygd opp på samme måte som tekster og spørsmål brukt i treningen. Tekstene inneholdt implisitt informasjon som krevde at leseren gjorde utfyllende inferenser som vist i utdraget under.

*I lyskryss reguleres trafikken av lys. Rødt lys betyr at du må stoppe og vente, gult betyr at du må gjøre deg klar. Når det blir grønt lys, kan du kjøre.*

*I rundkjøringer kjører man rundt en sirkel i veibanen. Man må selv se etter så man ikke kjører på noen, og man må tilpasse farten i forhold til andre biler. I rundkjøringene er det plass til flere biler enn i lyskryss, og de er derfor mer effektive. Ulykker blir som regel mindre alvorlige i rundkjøringer enn i lyskryss.*

Prøven hadde i alt fem tekster og 16 inferensspørsmål. Inferensspørsmålene ble skåret 0, 1, 2. 1 poeng var svar som gjenga relevant informasjon som sto eksplisitt i teksten. 2 poeng ble gitt når svaret uttrykte forståelse for leddet mellom opplysninger gitt i teksten. Under er et eksempel på spørsmål til teksten som er gjengitt over.

*Hvorfor blir ulykker som regel mindre alvorlige i rundkjøringer enn i lyskryss?*

Her ble det gitt 2 poeng for svar som fortalte at farten reduseres i rundkjøringer eller at farten er høyere i lyskryss. 1 poeng ble gitt dersom eleven gjenga relevant informasjon fra teksten om hvordan man kjører i rundkjøringer, med andre ord gjengir informasjon fra en eller begge setningene som forteller leseren om farten:

*I rundkjøringer kjører man rundt en sirkel i veibanen. Man må selv se etter så man ikke kjører på noen og tilpasse farten til de andre bilene.*

I tillegg ble elevene for hvert inferensspørsmål bedt om å markere to ord i teksten som hjalp dem å finne rett svar til spørsmålet. Det vil si, streke under to ord i en eller begge setningene over. Det ble gitt 1 poeng for rett understrekning for hvert inferensspørsmål. Høyeste oppnåelige poengsum sammenlagt for både åpne inferensspørsmål og understrekningsoppgaver var 48.

Samme tekster og spørsmål ble administrert ved alle tre måletidspunkter. Når elevene får de samme spørsmålene flere ganger, er det fare for at elevene husker tekstene og spørsmålene fra gang til gang og derfor får høyere skåre andre og tredje gang enn første gang prøven administreres. En slik læringseffekt vil imidlertid være lik i både kontroll- og eksperimentgruppe.

Elevene fikk 40 minutter å gjennomføre prøven på. Om lag halvparten av elevene ble ferdig med alle oppgavene innenfor tidsrammen ved pretesting. På posttesten ble drøyt 80 % av alle elevene ferdige med alle oppgavene, mens alle ble ferdige før tidsfristen ved utsatt posttest.

### ***Prøve i leseforståelse***

For å måle leseforståelse ble delprøven i leseforståelse i Diagnostic Reading Analysis (Crumpler & Colin, 2004) oversatt fra engelsk til norsk med tillatelse fra forlaget. Prøven har to parallelle sett, sett A og sett B. Diagnostic Reading Analysis er opprinnelig en individuell prøve med muntlige spørsmål og svar, men den norske oversettelsen av prøven ble administrert som en gruppeprøve med skriftlige svar. Hver av de to parallelle prøvene besto av tekster med stigende vanskelighetsgrad. For hvert vanskegradsnivå var det to tekster, en fiksjonstekst og en nonfiksjonstekst. Den norske oversettelsen inkluderte to prøver med to ganger seks tekster med spørsmål. Oppgavesett A ble brukt ved pretest og utsatt posttest, oppgavesett B ble brukt på ved posttesting. Hver tekst besto av mellom 50 og 70 ord.

Prøven hadde flere typer spørsmål, som kan deles inn i to kategorier, faktaspørsmål og tolkningsspørsmål. Faktaspørsmålene krever at eleven trekker ut informasjon fra teksten. Tolkningsspørsmålene krever at leseren tolker, trekker slutninger og reflekterer. Denne kategorien består av tre typer spørsmål. 1) Inferensspørsmål som krever at leseren knytter informasjon sammen fra flere steder i teksten. 2) Prediktive/forutsiende spørsmål som krever at eleven kommer med en rimelig forklaring eller ide i lys av den informasjonen han/hun har lest, typisk "hva tror du skjer videre". 3) Spørsmål om vokabular som krever at eleven trekker slutninger om betydningen av ukjente enkeltord ut fra sammenhengen i teksten. Totalt hadde hver av de to prøvene 34 spørsmål. De fire typene spørsmål gjenspeiler Snow og Sweet (2003) sin definisjon av leseforståelse som bestående av to prosesser, det å trekke ut informasjon (faktaspørsmål) og det å konstruere mening (inferensspørsmål, forutsiende spørsmål og vokabularspørsmål). Oppgavesett A besto av 17 faktaspørsmål og like mange tolkningsspørsmål, mens oppgavesett B besto av 18 faktaspørsmål og 16 tolkningsspørsmål.

Første versjon av oversettelsen ble prøvd ut blant 10 elever på femte trinn våren 2008. Den første versjon viste seg å ha takeffekt, derfor ble de to enkleste tekstene i begge oppgavesettene tatt ut, og to tekster med høyere vanskegrad ble oversatt og tatt inn i prøven.

Elevene hadde 40 minutter til å gjennomføre prøven på, hvilket var tilstrekkelig med tid for flertallet av elevene. Om lag 90 % av elevene ble ferdige med prøven innenfor tidsrammen på pretesten, mens alle ble ferdige innenfor tidsfristen ved posttest og utsatt posttest.

### **7.3.3.2 Kontrollvariabler**

Kontrollvariabler ble inkludert i studien av tre årsaker. For det første for å kunne kontrollere hvorvidt elevene i eksperimentgruppen og kontrollgruppen var sammenlignbare. For det andre for å kunne undersøke hvilke elever som eventuelt ikke profitterte på inferenstreningen, eller hvilke elever som eventuelt profitterte særlig på treningen. For det tredje ble to kontrollvariabler, motivasjon og matematikk, inkludert for å fastslå eventuelle Hawthorne-effekter.

I studien undersøkes effekten av inferenstrening for elevers evne til å gjøre inferenser og for deres prestasjoner i leseforståelse. Som det går fram i kapittel 2, er det en rekke ferdigheter og betingelser som har betydning for leseforståelsen. Særlig er det å ha kjennskap til ords betydning helt avgjørende for å oppnå forståelse av teksten. Derfor ble ordforråd inkludert som kontrollvariabel i studien. Videre er en god avkodning en forutsetning for å kunne forstå teksten, og vansker med leseforståelse kan skyldes problemer med avkodningen. Derfor ble elevenes avkodningsferdigheter kartlagt i studien. Det å gjøre inferenser krever evne til å resonnerer. Variablen nonverbale ferdigheter ble inkludert for å undersøke om elevenes evne til å resonnerer nonverbalt påvirket deres evne til å gjøre inferenser. Motivasjon og matematikk ble også inkludert som kontrollvariabler. Det var to grunner for dette. For det første ble variablene brukt som en kontroll på at eksperimentgruppen og kontrollgruppen var like og dermed sammenlignbare ved pretesten. For det andre ble disse to variablene også målt ved posttest for å kunne avgjøre hvorvidt en eventuell effekt er spesifikt knyttet til lesing, og ikke er en generell effekt på skoleferdigheter, såkalt Hawthorneeffekt.

#### ***Spørreskjema om motivasjon***

Motivasjon ble målt ved hjelp av et spørreskjema som var en oversettelse av deler av spørreskjemaet "Motivation for Reading Questionnaire". Spørreskjemaet er utviklet av Guthrie og Wigfield med kolleger (Baker & Wigfield, 1999; Guthrie, Wigfield, & VonSecker, 2000; Wigfield & Guthrie, 1997), og er et selvrapporteringsskjema for indre og ytre motivasjon. Den norske oversettelsen ble gjort på bakgrunn av skjemaet som er publisert i Guthrie, Wigfield og VonSecker (2000). I denne studien var kun de to konstruktene "curiosity" og "involvement", som måler indre motivasjon, tatt med.

Konstruktet "curiosity" besto av følgende fire item:

- Hvis læreren forteller om et spennende tema, leser jeg kanskje mer om det
- Jeg leser om mine hobbyer og fritidsinteresser for å lære mer om dem
- Jeg leser for å lære noe nytt om temaer som interesserer meg
- Jeg liker å lese bøker om mennesker som bor i andre land

Konstruktet involvement besto av følgende fem item:

- Jeg lager bilder inne i hodet når jeg leser
- Personer jeg leser om i bøker blir som venner for meg
- Jeg liker å lese lange og engasjerende fortellinger eller bøker
- Jeg liker bøker med spenning og mysterier
- Jeg leser ofte eventyrfortellinger

Itemene ble blandet i tilfeldig rekkefølge i spørreskjemaet. Itemene er formulert som påstander. Elevene skulle ta stilling til i hvilken grad de var enig eller uenig i påstandene ved å krysse av på en likert-skala fra 1 til 4, hvor 1 var "veldig uenig" og 4 "veldig enig". Spørreskjemaet ble administrert av lærerne, og gjennomført klassevis.

### ***Ordforrådsprøve***

Ordforrådsprøven ble laget for denne undersøkelsen og er derfor ikke standardisert. Denne undersøkelsen var snarere en utprøving av prøven. En dansk semantisk ordforrådsprøve ble brukt som modell for prøven (<http://laes.hum.ku.dk>).

Prøven er en gruppetest hvor deltakerne får lest opp et ord eller en forklaring, deretter skal de krysse av for hvilket av tre alternative ord som betyr det samme som det første ordet/forklaringen. For eksempel "Flink, er det det samme som arg, dyktig eller fornøyd?". Se appendiks for fullstendig instruksjon.

Ordene i prøven ble valgt ut fra Kunnskapsforlagets ordbok på nett ([www.ordnett.no](http://www.ordnett.no)), i tillegg ble noen ord valgt ut fra lærebøker for sjette trinn. Det ble gjort en kvalitativ vurdering av vanskelighetsgraden på aktuelle ord. Noen av itemene var de samme som i den tilsvarende ordforrådsprøven for niende trinn som ble brukt i forundersøkelsen, se kapittel 6. Et første utkast av ordforrådsprøven ble prøvet ut blant 20 femteklassinger våren 2008. Etter tilbakemelding fra elevene som deltok i utprøvingen, ble prøven vurdert til å ha for mange vanskelige item, noen av de vanskeligste itemene ble derfor tatt ut og noen av svaralternativene ble byttet ut. Den endelige prøven besto av 24

item (se appendiks). Fordi den endelige utgaven av prøven ikke tidligere var prøvd ut, ble det gjennomført en skalaanalyse. Til sammen fire item ble tatt ut av prøven før analysene ble gjennomført. Grunnene for at itemene ble tatt ut var følgende: Et item ("foredrag") korrelerte negativt med øvrige item, et item ("drøfte") korrelerte svært lavt med øvrige item. På to av itemene ("fruktgrøt" og "enerett") hadde mindre enn en tredjedel av elevene svart riktig, med andre ord hadde de mindre riktig enn om de hadde gjettest på riktig svar. Cronbachs alfa ble beregnet til 0.65 for de 20 gjenstående itemene.

### ***Ordkjedeprøve***

Avkodningsferdigheter ble målt ved Høien og Tønnesens (1997) ordkjedeprøve. Prøven kartlegger ordavkodning. Prøven består av ordkjeder, det vil si ord skrevet sammenhengende uten mellomrom, for eksempel "gårhjemishatt". Elevene skal sette en strek mellom ordene der hvor mellomrommene skal være. I løpet av 4 minutter skal elevene sette så mange streker som mulig. Prøven inneholder i alt 90 ordkjeder.

I 1997 ble det laget normer for ordkjedeprøven for tredje til tiende klasse (etter læreplanen av 1997). Normene ble laget på bakgrunn av en utprøving blant elever ved 8 skoler i Rogaland fylke. Siden denne utprøvingen foregikk i månedsskiftet mai – juni, vil det være naturlig å sammenligne skårene til elevene i dette utvalget med femteklasseskårene i normeringen. Gjennomsnittsskåren for de 159 femteklassingene som deltok i normeringen i 1997 var 27,9 poeng, standardavviket var 8,8.

### ***Raven***

Nonverbale ferdigheter ble kartlagt ved hjelp av Raven's Standard Progressive Matrices (Raven, 1958).

### ***Matematikkprøve – M5***

Matematikkferdigheter ble målt ved M5 prøven, del 1 (Tornes, 1996) som er en standardisert prøve beregnet brukt i mai/juni i femte klasse eller i august/september i sjette klasse. Prøven består av oppgaver innen talloppfatning, de fire regningsarter, relasjoner, utsagn/ligninger og taltallsystemet. Matematikkprøven er normert ved "standard nine-skala", som innebærer at resultatene uttrykkes i prøveklasser fra 1 til 9, hvor prøveklasse 1 tilsvarer svakest resultat og 9 best resultat. Resultater på del 1 mellom 50 og 56 poeng tilsvarer prøveklasse 5, som er den midterste prøveklassen.



### **7.3.3.3 Mål på implementering**

#### ***Observasjonsskjema***

Det ble utarbeidet et observasjonsskjema for å kunne gjøre det mulig å sammenligne rammene for undervisning i de 16 klassene. Eksperimentklassene ble observert i timer hvor de hadde inferensleksjoner, kontrollklassene i timer hvor de hadde arbeid av lignende karakter, for eksempel arbeid med arbeidsplan, felles gjennomgang og individuelle oppgaver og gruppeoppgaver.

I skjemaet ble de 20 første minuttene av en skoletime eller de 20 første minuttene etter skifte av en aktivitet registrert. Skjemaet besto av fire intervaller: I de første fem minuttene ble det registrert hvor lang tid det tok før undervisningen kunne begynne etter friminutt eller etter overgang fra en aktivitet til en annen. Det ble registrert hvor mange elever og lærere som var til stede og hvilken type oppgaver eller leksjon som klassen jobbet med. De neste fem minuttene ble det registrert hvor mange elever som avbrøt lærer mens han/hun snakket, hvor mange elever som avbrøt sine medelever der dette ikke angikk oppgavene, og hvor mange elever som snakket samtidig som lærer ga en felles beskjed til klassen. De neste fem minuttene ble graden (liten – middels – stor) av støy i klasserommet som ikke angikk oppgavene registrert. I tillegg ble det registrert generelt inntrykk av hvorvidt elevene fant oppgavene eller arbeidet lystbetont. For eksperimentklassene ble det også registrert hvor mye tid som gikk med til leksjonen. Observasjonsskjemaet er gjengitt i appendiks.

#### ***Loggskjema***

Lærerne i eksperimentgruppen fylte ut et loggskjema etter hver leksjon. I loggskjemaet ble det notert dato for når den enkelte leksjonen ble gjennomført, samt hvor mange av elevene (alle, over halvparten, under halvparten, ingen) som gjennomførte oppgavene. I tillegg ble det registrert hvor mye tid som ble brukt på leksjonene, hvor mange elever som var til stede og eventuelle kommentarer. Eksempel på loggskjema er gjengitt i appendiks.

### **7.3.4 Gjennomføring**

Datainnsamlingen ble gjennomført på tre tidspunkter i løpet av perioden august til desember 2008. Inferenstreningen ble gjennomført mellom pretest og posttest. Den utsatte posttesten ble gjennomført fem uker etter siste leksjon i inferenstreningen. Prøvene ble administrert av lærerne, en forskerassistent med master i spesialpedagogikk og undertegnede. Tabellen under viser hvilke

prøver som ble gjennomført på hvilke tidspunkter, samt hvem som administrerte prøvene og minimum- og maksimumskåre for de ulike prøvene.

Tabell 21: Oversikt over variabler og prøver

Variabler	Prøve	Min. skåre	Maks. skåre	Administrering av tester		
				Pretest	Posttest	Utsatt posttest
Inferens	Inferensprøve	0	48	U	U	U
Leseforståelse	Prøve i leseforståelse	0	34	U	U	U
Motivasjon	Spørreskjema om motivasjon	0	36	L	L	
Ordforråd	Ordforrådsprøve	0	20	A/U		
Avkoding	Ordkjedeprøve	0	90	L		
Nonverbale ferdigheter	Raven	0	60	A		
Matematikkferdigheter	M5, del 1	0	74	L	L	

U= Undertegnede, L= Lærer, A= assistent

Dataene fra undersøkelsen ble analysert med versjon 15 av programvaren Statistical Package for the Social Sciences (SPSS).

#### 7.3.4.1 Pretest

Første datainnsamling ble gjennomført i månedsskiftet august – september. Innsamlingen ble gjennomført i flere økter. En økt på cirka en skoletime per klasse hvor Raven og ordforrådsprøven ble gjennomført. Disse prøvene ble administrert av en forskerassistent. Ordforrådsprøven ble gjennomført to ganger fordi første gangs testing ikke ble administrert slik som intensjonen var. I stedet for å lese ordene høyt for elevene leste elevene ordene på egen hånd. Prøven ble derfor gjennomført en gang til, denne gangen i månedsskiftet september – oktober av undertegnede. Inferensprøven og prøven i leseforståelse ble administrert av undertegnede i en økt på to skoletimer. I tillegg administrerte lærerne ordkjedeprøven, M5 og spørreskjema om motivasjon i ukene 37, 38 og begynnelsen av uke 39. I en klasse ble M5 administrert i uke 40. Elever som ikke var til stede på prøvetidspunktene gjennomførte prøvene på et senere tidspunkt i de tilfeller det var mulig. Enkelte elever var ikke til stede i lengre perioder, slik at det ikke var mulig å gjennomføre pretestene. Spørreskjemaet om motivasjon ble

ikke gjennomført i fire av klassene i utvalget ved pretestene, hvorav tre kontrollklasser og en eksperimentklasse.

#### **7.3.4.2 Inferenstrening**

Inferenstreningen ble gjennomført i perioden uke 39 til uke 48 (skolene hadde ferie i uke 41). Det var vanskelig for alle klassene å starte opp med inferenstreningen samtidig. Fem klasser gikk i gang med inferenstreningen i uke 39, de andre fem begynte i uke 42.

Observasjon av undervisning, både i eksperiment og kontrollgruppe, ble gjennomført i perioden uke 40 til uke 46.

#### **7.3.4.3 Posttest**

Posttestene i leseforståelse og inferens ble gjennomført ca. en uke etter at inferenstreningen var avsluttet. Siden klassene startet opp med inferenstreningen på litt ulikt tidspunkt, ble posttestene i eksperimentklassene gjennomført på ulikt tidspunkt i perioden uke 44 til uke 49. Kontrollklassene ble testet i uke 45. I to av eksperimentklassene var det uro uforenelig med arbeidsro da inferensprøven og prøven i leseforståelse ble gjennomført. Til sammen 4 elever i disse to klassene boikottet inferensprøven.

Matematikkprøven (M5) og spørreskjema om motivasjon ble administrert av lærer uken etter at inferensprøven og prøven i leseforståelse ble gjennomført. I en av eksperimentklassene ble ikke M5-prøven og spørreskjemaet om motivasjon gjennomført grunnet sykdom.

#### **7.3.4.4 Utsatt posttest**

To eksperimentklasser (klasse 07 og 08) og to kontrollklasser (15 og 16) gjennomførte inferensprøven og prøve i leseforståelse fire uker etter pretesten (fem uker etter endt inferenstrening). I perioden mellom posttest og utsatt posttest fikk klassene ordinær undervisning. Klasse 07 hadde imidlertid vikar i denne perioden.

### **7.4 Resultater**

#### **7.4.1 Vurdering av gjennomføring av inferenstreningen**

Rammene for undervisning ble kartlagt ved hjelp av observasjonsskjema av to årsaker. For det første for å undersøke om eksperimentklassene og kontrollklassene hadde samme rammer for undervisning, med andre ord om de var sammenlignbare. For det andre ble implementeringen av inferens-

treningen vurdert for å kunne forklare eventuelle klasseforskjeller i effektmålene.

Observasjonene av eksperimentklassene ga også et inntrykk av implementeringen av inferenstreningen. I tillegg ble implementeringen av inferenstreningen undersøkt ved hjelp av loggskjema som lærerne i eksperimentgruppen fylte ut. Under blir det gjort rede for resultater både fra observasjonene og fra loggskjemaene.

#### **7.4.1.1 Klasseromsobservasjoner**

Alle klassene i utvalget ble observert i 20 minutter, hvor observasjonsskjema ble fylt ut. Antall elever i klassen på observasjonstidspunktet varierte mellom 12 og 21. I to eksperimentklasser var det to lærere til stede under observasjonen, mens i de øvrige klassene var det kun en lærer til stede. 13 klasser, 7 eksperimentklasser og 6 kontrollklasser brukte 4 minutter eller mindre på å komme til ro etter friminutt eller etter ny aktivitet før undervisningen kunne begynne. En eksperimentklasse brukte 5 minutter, og to eksperimentklasser brukte 7 og 8 minutter. I de to sistnevnte klassene ble ”støy som ikke angår arbeidet” vurdert å være ”middels grad”. I øvrige klasser ble grad av støy karakterisert som ”liten grad”.

Inntrykket av hvorvidt arbeidet i inferenstreningen i eksperimentgruppen og arbeidet med andre oppgaver i kontrollgruppen var lystbetont, var forholdsvis likt i begge grupper. I to eksperimentklasser og i en kontrollklasse var inntrykket at klassene opplevde arbeidet i inferenstreningen som lite lystbetont, i øvrige klasser var inntrykket at arbeidet var middels lystbetont.

Observasjonene fra eksperimentklassene ga også et generelt inntrykk av implementeringen av inferenstreningen ut over de data som ble registrert i observasjonsskjemaet. Selv om det ikke var samme leksjon som ble observert i alle klasser, virket det som om inferenstreningen ble gjennomført forholdsvis likt i alle eksperimentklasser. Det var imidlertid noe forskjell mellom klassene på tre punkter. For det første var det noe forskjell mellom lærerne i eksperimentgruppen i forhold til hvor tett opp til opptil instruksjonen til leksjonen inferenstreningen var. Enkelte lærere fulgte instruksjonen svært tett med hensyn til hvilke ord som var foreslått forklart i klassen, og gjennomgangen av riktige svar var tett knyttet til fasiten som lærerne hadde i instruksjonspermen. Andre lærere virket mer ”frie” i forhold til instruksjonen og fasiten. Likevel var det generelle inntrykket helt klart at alle lærerne fulgte instruksjonene til hver leksjon. For det andre var det forskjell i hvor mye tid som ble brukt til å gjennomgå elevenes svar på oppgavene. Mens noen lærere

lot en til to elever fortelle sine svar på oppgavene, gjennomgikk andre lærere så nær som alle elevene sine svar på hvert spørsmål. For det tredje var det forskjell mellom klassene i forhold til hvor mange elever som trengte ekstra hjelp og forklaringer fra lærer.

#### **7.4.1.2 Loggskjema**

To av lærerne i eksperimentgruppa fylte ikke ut loggskjema. I disse tilfellene ga lærerne muntlig informasjon om i hvilken periode inferenstreningen var blitt gjennomført. Det var i tillegg noen loggskjema som manglet dato for når en eller flere leksjoner ble gjennomført.

I leksjon fire og seks skulle elevene lese tre tekster, her var det fire lærere som rapporterte at de ikke hadde rukket å gjennomføre den siste av de tre tekstene i leksjon seks. En lærer rapporterte at under halvparten av elevene hadde rukket å gjennomføre den siste av de tre tekstene i leksjon fire. Ellers rapporterte samtlige lærere at enten alle eller over halvparten av elevene hadde gjennomført samtlige tekster og oppgaver i inferenstreningen.

Flere lærere rapporterte at de brukte mer enn 20 minutter på leksjonene. Fire klasser brukte gjennomsnittlig mellom 20 og 25 minutter per leksjon, mens fire klasser brukte gjennomsnittlig mellom 33 og 44 minutter per leksjon. Gjennomsnittlig tid brukt per leksjon var 31,15 minutter for de 8 klassene som har oppgitt tid per leksjon. I forbindelse med observasjonene ble det notert hvor mye tid eksperimentklassene brukte til inferenstreningen. Da notatene fra observasjonene ble sammenlignet med loggskjemaene, viste det seg at et flertall av lærerne hadde notert lengre tid, ca. fem minutter mer enn det som ble notert i forbindelse med observasjonen.

I sju av klassene ble inferenstreningen gjennomført i løpet av en periode på fire uker (eksklusive en uke skoleferie). En klasse gjennomførte de åtte leksjonene på tre uker, en klasse brukte sju uker, og en klasse brukte åtte uker. I sistnevnte klasse fikk elevene vikar da halvparten av leksjonene var gjennomført, vikaren gjennomførte de siste leksjonene, men klassen fikk et opphold på ca. 4 uker halvveis i leksjonene.

I loggskjemaet hadde lærerne mulighet for å skrive inn kommentarer til hver enkelt leksjon. Noen lærere valgte å ikke skrive inn kommentarer, andre skrev kommentarer til enkelte av leksjonene. Mange av kommentarene handlet om tekster eller oppgaver som læreren og/eller elevene syntes var vanskelige, eller tekster og oppgaver som elevene hadde klart særlig bra. Flere lærere rapporterte om engasjement hos elevene, om elever som hadde fått "a-ha"-

opplevelser, og om elever som hadde fått ”nytt syn på lesing”. Denne typen kommentarer ble for det meste rapportert i forhold til de tre første leksjonene. To lærere rapporterte etter leksjon seks at elevene begynte å bli lei av arbeidsmetoden, mens en annen lærer rapporterte etter samme leksjon at hun merket at elevene brukte erfaringer fra tidligere leksjoner, og at elevene løste oppgavene lettere enn i tidligere leksjoner. To lærere kommenterte at svært ulikt arbeidstempo hos elevene skapte problemer. Den ene læreren mente det ble et problem at andre elever ventet på at de skulle bli ferdige slik at oppgavene kunne gjennomgås. Den andre læreren mente det var et problem at det ble synlig at enkelte elever trengte mye tid på å løse oppgavene, og dette kunne oppleves ubehagelig for elever som trengte mye tid.

#### **7.4.2 Innledende analyser**

I dette avsnittet presenteres først deskriptiv statistikk for kontrollvariabler og effektvariabler, samt korrelasjoner mellom variablene. Dette for å gi en oversikt over datamaterialet og forutsetninger for videre analyser. Videre vil det blitt gjort rede for forskjeller mellom eksperimentgruppen og kontrollgruppen på kontrollvariabler og effektvariabler på de tre måletidspunktene. Dette for å vise i hvilken grad de to gruppene var sammenlignbare ved pretest og for å vise endringer på effektvariablene. Deretter sammenlignes skårer hos gutter og jenter, samt enspråklige og tospråklige sine skårer på pre- og posttest. Implementeringen av inferenstreningen i eksperimentklassene er en forutsetning for å diskutere effekter av treningen. Derfor blir det gjort rede for implementeringen av inferenstreningen mot slutten av dette avsnittet.

##### **7.4.2.1 Deskriptiv statistikk og sammenligning av gruppene**

Gjennomsnitt, standardavvik, skewness og kurtosis for effektvariabler og kontrollvariabler for utvalget totalt, samt for eksperimentgruppe og kontrollgruppe på alle tre måletidspunkt, er oppgitt i tabell 22 til 24. Forskjeller mellom gruppene ble undersøkt med t-tester for uavhengige grupper. Kvadrert eta ble beregnet for å anslå størrelsen på forskjellen mellom gruppene.

*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 22: Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på pretest, samt forskjeller mellom gruppene*

Mål	Eksperimentgruppe					Kontrollgruppe				
	N	M	SD	Skew.	Krt.	N	M	SD	Skew.	Krt.
Ordforråd	144	12.59	3.08	-.31	.00	78	11.92	2.92	-.50	1.42
Ordavkodning	151	35.25	8.87	-.02	-.43	82	31.24	8.08	.35	.61
Nonverbale ferdigheter	148	39.46	8.912	-1.17	1.81	83	38.94	10.22	-1.54	1.96
Motivasjon, totalt	128	25.58	4.08	-.53	1.27	40	25.32	3.74	.11	-.26
Motivasjon, involvement	128	14.15	2.42	-.33	.80	40	14.01	2.52	-.61	.60
Motivasjon, curiosity	128	11.43	2.22	-.59	.60	40	11.31	1.90	.07	-.41
Matematikk	148	51.85	12.12	-.80	.45	84	53.17	11.47	-.70	.34
Inferens, totalt	147	14.13	8.57	.65	.00	84	14.17	7.07	.20	-.55
Inferens, åpne	147	10.02	6.23	.40	-.52	84	9.71	5.28	.21	-.47
Inferens, understrekning	147	4.11	2.88	.89	.71	84	4.46	2.40	.31	-.44
Leseforståelse, totalt	147	22.13	5.55	-.88	.73	84	21.63	5.55	-.56	-.42
Leseforståelse, fakta	147	12.37	2.94	-.99	.69	84	12.13	2.90	-.38	-.48
Leseforståelse, tolkning	147	9.76	2.96	-.49	.26	84	9.50	3.10	-.58	-.27
Leseforståelse, fiksjon	147	12.68	3.21	-1.19	1.33	84	12.40	3.00	-.93	.59
Leseforståelse, nonfiksjon	147	9.45	2.89	-.37	-.24	84	9.23	3.05	-.22	-.73

*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 22, forts.*

Mål	N	M	Totalt				
			SD	Skew.	Krt.	$\eta^2$	P
Ordforråd	222	12.36	3.03	-.34	.41	.01	=.118
Ordavkoding	233	33.84	8.78	.13	-.27	.05	=.001
Nonverbale ferdigheter	231	39.27	9.38	-1.33	1.94	<.01	=.687
Motivasjon, totalt	168	25.52	4.00	-.41	.96	<.01	=.730
Motivasjon, involvement	168	14.12	2.44	-.40	.70	<.01	=.758
Motivasjon, curiosity	168	11.40	2.14	-.48	.46	<.01	=.769
Matematikk	232	52.33	11.88	-.77	.41	<.01	=.419
Inferens, totalt	231	14.15	8.05	.54	-.04	<.01	=.969
Inferens, åpne	231	9.91	5.89	.36	-.44	<.01	=.705
Inferens, understrekning	231	4.24	2.71	.72	.42	<.01	=.348
Leseforståelse, totalt	231	21.95	5.54	-.75	.16	<.01	=.512
Leseforståelse, fakta	231	12.28	2.92	-.77	.22	<.01	=.555
Leseforståelse, tolkning	231	9.67	3.01	-.53	.04	<.01	=.526
Leseforståelse, fiksjon	231	12.58	3.12	-1.09	1.05	<.01	=.519
Leseforståelse, nonfiksjon	231	9.37	2.95	-.32	-.46	<.01	=.581



*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 23: Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på posttest, samt forskjeller mellom gruppene*

Mål	Eksperimentgruppe					Kontrollgruppe				
	N	M	SD	Skew.	Krt.	N	M	SD	Skew.	Krt.
Motivasjon, totalt	113	25.32	4.13	-.57	1.07	82	25.40	3.41	-.26	-.40
Motivasjon, involvement	113	14.11	2.78	-.60	.49	82	14.20	2.42	-.24	-.11
Motivasjon, curiosity	113	11.21	2.17	-.06	-.33	82	11.20	1.65	-.16	.13
Matematikk	135	56.35	11.23	-.84	.24	83	56.14	11.19	-1.09	2.07
Inferens, totalt	139	23.22	8.43	-.25	-.43	84	16.86	7.13	.21	-.10
Inferens, åpne	139	16.55	5.72	-.45	-.19	84	11.82	5.47	-.07	-.52
Inferens, understrekning	139	6.67	3.19	.14	-.46	84	5.04	2.53	.60	1.10
Leseforståelse, totalt	143	24.47	4.84	-1.29	2.79	84	21.06	4.25	-.10	-.53
Leseforståelse, fakta	143	14.88	2.52	-1.87	6.30	84	13.77	2.28	-.42	-.51
Leseforståelse, tolkning	143	9.59	2.91	-.44	-.06	84	7.29	2.63	.10	-.28
Leseforståelse, fiksjon	143	13.13	2.54	-1.19	3.07	84	11.40	2.21	-.30	.49
Leseforståelse, nonfiksjon	143	11.34	2.92	-.82	.27	84	9.65	2.79	.10	-.65

*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 23, forts.*

Mål	N	Totalt					
		M	SD	Skew.	Krt.	$\eta^2$	P
Motivasjon, totalt	195	25.35	3.83	-.49	.80	<.001	=.879
Motivasjon, involvement	195	14.15	2.63	-.49	.36	<.001	=.876
Motivasjon, curiosity	195	11.20	1.96	-.09	-.09	<.001	=.973
Matematikk	218	56.27	11.18	-.92	.87	<.001	=.897
Inferens, totalt	223	20.83	8.52	.02	-.57	.13	<.001
Inferens, åpne	223	14.77	6.06	-.24	-.52	.14	<.001
Inferens, understrekning	223	6.06	3.06	.37	-.18	.07	<.001
Leseforståelse, totalt	227	23.21	4.90	-.71	.16	.11	<.001
Leseforståelse, fakta	227	14.47	2.49	-1.28	3.28	.05	<.001
Leseforståelse, tolkning	227	8.74	3.02	-.16	-.48	.14	<.001
Leseforståelse, fiksjon	227	12.49	2.55	-.70	1.29	.11	<.001
Leseforståelse, nonfiksjon	227	10.71	2.98	-.43	-.51	.07	<.001

*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 24: Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på utsatt posttest*

Mål	Eksperimentgruppe					Kontrollgruppe				
	N	M	SD	Skew.	Krt.	N	M	SD	Skew.	Krt.
Inferens, totalt	29	23.90	7.43	.35	-.15	27	16.78	6.30	-.16	.64
Inferens, åpne	29	17.62	4.81	.16	-.40	27	12.41	4.81	-.21	-.15
Inferens, understrekning	29	6.28	3.138	.56	-.02	27	4.37	2.45	1.45	5.06
Leseforståelse, totalt	29	24.28	4.50	-.81	.03	27	22.26	5.30	-.02	.06
Leseforståelse, fakta	29	13.48	2.39	-.83	.59	27	12.33	2.75	-.34	-.23
Leseforståelse, tolkning	29	10.79	2.93	-.68	-.26	27	9.93	2.93	.10	.45
Leseforståelse, fiksjon	29	13.66	2.94	-1.36	1.66	27	12.41	3.041	-.57	-.12
Leseforståelse, nonfiksjon	29	10.62	2.73	-.54	-.45	27	9.85	3.17	-.09	-.15

*Tabell 24, forts.*

Mål	Totalt						
	N	M	SD	Skew.	Krt.	$\eta^2$	P
Inferens, totalt	56	20.46	7.73	.29	.34	.22	<.001
Inferens, åpne	56	15.11	5.44	-.02	-.04	.23	<.001
Inferens, understrekning	56	5.36	2.96	.91	.93	.11	=.015
Leseforståelse, totalt	56	23.30	5.20	-.40	-.32	.04	=.149
Leseforståelse, fakta	56	12.93	2.61	-.58	-.08	.05	=.100
Leseforståelse, tolkning	56	10.38	2.939	-.29	-.27	.02	=.274
Leseforståelse, fiksjon	56	13.05	3.03	-.90	.25	.04	=.125
Leseforståelse, nonfiksjon	56	10.25	2.95	-.32	-.36	.02	=.334

I henhold til Byrnes (1998) retningslinjer for tolkning av skewness og kurtosis var alle variablene på de tre måletidspunktene tilnærmet normalfordelte (jevnfør avsnitt 6.4.1.1)

På prøven i ordavkoding lå begge gruppene gjennomsnitt noe over prøvens normer for 5. klasse. På matematikk skåret begge gruppene hver for seg og hele utvalget samlet innenfor prøveklasse 5 på pretest. På posttest var de gjennomsnittlige skårene både for eksperimentgruppen, kontrollgruppen og hele utvalget, like over prøveklasse 5.

Spørreskjema om motivasjon ble gjennomført i 12 klasser ved pretest, og i 15 klasser ved posttest. 7 av de 9 itemene hadde noen få missing, det vil si at elevene enten ikke hadde svart, eller hadde krysset av på flere alternativer. Missing varierte fra 0 til 3 % på de 9 itemene. Før analysene ble gjennomført, ble missing erstattet med gjennomsnittskåre for itemet.

En stor andel av elevene hadde latt være å svare på ett eller flere av spørsmålene på inferensprøven. Det var også en del elever som ikke ble ferdig med alle oppgavene på pretest. I skåringen av prøven ble det skilt mellom tre typer missing: Ubesvart, forsøkt besvart men ikke kommet fram til et svar (elevene ble bedt om å sette et kryss på oppgaver som de hadde brukt tid på å besvare, men hvor de ikke hadde kommet fram til et svar), og misforstått oppgaven. Sistnevnte type gjaldt kun for understrekningsoppgavene. Her var det noen som streket under ord i spørsmålet og ikke i teksten. Der hvor elevene hadde misforstått understrekningsoppgaver, ble missing erstattet med gjennomsnittlig skåre på det aktuelle itemet. De to øvrige typene missing ble erstattet med 0 poeng, dette fordi det er rimelig å anta at denne typen missing er uttrykk for at elevene ikke hadde klart å besvare oppgaven, snarere enn at elevene har hoppet over oppgavene fordi de ikke hadde oppdaget oppgavene. Siden de vanskeligste oppgavene var plassert mot slutten av prøven, var det også rimelig å skåre ubesvarte spørsmål mot slutten av prøven som 0 poeng. Ved pretest var Cronbachs alfa lik .82 for hele utvalget etter at missing var erstattet med 0 poeng.<sup>6</sup> Cronbachs alfa var tilnærmet det samme for hele utvalget også ved posttest og utsatt posttest, men Cronbachs alfa var høyere for eksperimentgruppen (posttest:  $\alpha=.81$ , utsatt posttest:  $\alpha=.79$ ) enn kontrollgruppen (posttest  $\alpha=.76$ , utsatt posttest  $\alpha=.70$ ).

Av tabellene over går det fram at både eksperimentgruppen og kontrollgruppen hadde høyere raskårer i inferens på posttest enn på pretest.

---

<sup>6</sup> Kun 16 elever svarte på alle spørsmål på inferensprøven. Cronbach alfa beregnet ut fra disse 16 elevene var .74.

Derfor ble gruppenes gjennomsnittskårer ved pre- og posttest sammenlignet ved hjelp av Wilcoxon test for matchede par. Wilcoxon test viste at forskjellen mellom eksperimentgruppens skårer på posttest var signifikant høyere enn ved pretest,  $p < .001$ . Summen av rangering var 105.00 og 9075 for henholdsvis negativ og positiv rangering. For kontrollgruppen viste Wilcoxon test at forskjellen mellom råskårer på pre- og posttest var signifikant,  $p < .001$ . Summen av rangering var 656.50 og 2424.50 for henholdsvis negativ og positiv rangering.

Også på prøven i leseforståelse var det enkelte elever som ikke hadde besvart alle item, men her var antallet ubesvarte spørsmål betraktelig lavere enn på inferensprøven. Ubesvarte item ble tolket som at eleven ikke har klart å besvare oppgaven, ble derfor skåret som 0 poeng. Cronbachs alfa ved pretest (oppgavesett A) var lik  $.84^7$  for hele utvalget. Ved posttest (oppgavesett B) var alfa lik  $.76^8$  for hele utvalget, men henholdsvis  $.85$  og  $.86$  for kontrollgruppe og eksperimentgruppe. Ved utsatt posttest var Cronbachs alfa lik  $.74$  for kontrollgruppen, og  $.65$  for eksperimentgruppen.

Av tabellene over går det fram at eksperimentgruppens gjennomsnittlige skåre på leseforståelse var høyere på posttest enn ved pretest, mens kontrollgruppen hadde lavere gjennomsnittskåre på posttest enn pretest. For eksperimentgruppen viste Wilcoxon test at forskjellen mellom råskårer ved pretest og råskårer på posttesten signifikant,  $p < .001$ . Summen av rangering var 1875.50 og 6380.50 for henholdsvis negativ og positiv rangering. For kontrollgruppen viste Wilcoxon test at forskjellen mellom råskårer på pre- og posttest ikke var signifikant,  $p = .191$ . Summen av rangering var 1587.50 og 1113.50 for henholdsvis negativ og positiv rangering.

Eksperimentgruppen og kontrollgruppen ble sammenlignet både på effektvariablene inferens og leseforståelse, og kontrollvariablene ordforråd, ordavkoding, nonverbale ferdigheter, motivasjon og matematikk, alle målt ved pretest. Dette for å undersøke hvor like, og dermed hvor sammenlignbare, gruppene var i forkant av inferenstreningen. Det var kun signifikant forskjell mellom gruppene med hensyn til en av kontrollvariablene, ordavkoding. Eksperimentgruppen skåret signifikant høyere enn kontrollgruppen med hensyn til ordavkoding. Forskjellen kan karakteriseres som moderat. Det var

---

<sup>7</sup> Cronbachs alfa var tilnærmet det samme uavhengig om ubesvarte spørsmål er inkludert eller ikke.

<sup>8</sup> Cronbachs alfa var tilnærmet det samme uavhengig om ubesvarte spørsmål er inkludert eller ikke.

ingen signifikante forskjeller mellom gruppene med hensyn de to effektvariablene inferens og leseforståelse.

På posttest var det ingen signifikante forskjeller mellom eksperimentgruppe og kontrollgruppe med hensyn til de to kontrollvariablene motivasjon og matematikk. Derimot var det en signifikant forskjell mellom gruppene med hensyn til inferens og leseforståelse, både for totalskåre og delskåre, i eksperimentgruppens favør. Forskjellene kan karakteriseres som store.

På utsatt posttest var det kun signifikant forskjell mellom gruppene i forhold til inferens, denne forskjellen kan karakteriseres som stor. Det var imidlertid ikke signifikant forskjell mellom eksperimentgruppe og kontrollgruppe med hensyn til leseforståelse, verken på totalskåre eller delskåre.

#### 7.4.2.2 Korrelasjoner mellom effektvariablene og kontrollvariablene ved pretest, samt språk, kjønn og alder

Forholdet mellom variablene målt ved pretesting ble undersøkt med Pearson's r. Korrelasjoner for effektvariabler og kontrollvariabler er oppsummert i tabell 25.

Tabell 25: Oversikt over korrelasjoner mellom kontrollvariabler og effektvariabler

	1	2	3	4	5	6	7
1 Inferens	-						
2 Leseforståelse	.55	-					
3 Motivasjon	.24	.29	-				
4 Ordavkoding	.44	.40	.18	-			
5 Ordforråd	.44	.42	.27	.35	-		
6 Non verbale ferdigheter	.44	.44	.20	.30	.30	-	
7 Matematikk	.47	.49	.16	.45	.27	.44	-

Signifikansnivået var  $p < .05$  for alle variabler

Det var moderate korrelasjoner mellom de fem kontrollvariablene motivasjon, ordavkoding, ordforråd, nonverbale ferdigheter og matematikk. Korrelasjonen mellom de to målene på effekt, inferens og leseforståelse var sterk. Korrelasjonen mellom effektmålene og ordavkoding, ordforråd og nonverbale ferdigheter var alle moderate, mens korrelasjonen mellom effektmålene og motivasjon var noe lavere. Korrelasjonene mellom matematikk og effektmålene kan karakteriseres som middels til sterk.

Det ble også beregnet korrelasjoner mellom variablene språk, kjønn og alder og øvrige variabler ved pretest. Variablen språk var definert som enspråklig og tospråklig. Enspråklig var kodet som 0 og tospråklig var kodet 1. Språk

korrelerte  $-.17$ ,  $p=.009$ , med leseforståelse og  $-.24$ ,  $p<.001$ , med matematikk, det var ingen signifikante korrelasjoner mellom språk og øvrige variabler. Det var også en signifikant korrelasjon mellom kjønn og nonverbale ferdigheter,  $.18$ ,  $p=.006$ , ellers var det ingen signifikante korrelasjoner mellom kjønn og øvrige variabler. Jenter var kodet 1 og gutter 0, hvilket innebærer at jenter assosieres med høyere skåre på nonverbale ferdigheter enn gutter. Alder korrelerte  $.20$ ,  $p=.003$ , med matematikk, det var ingen signifikante korrelasjoner mellom alder og øvrige variabler.

#### **7.4.2.3 Kontroll av Hawthorneeffekt**

Variablene motivasjon og matematikk ble inkludert for å kunne kontrollere om effekter av inferenstreningen kun var spesifikt knyttet til lesing eller om effekten var av generell art, med andre ord Hawthorneeffekt. I dette avsnittet redegjøres det for eksperiment- og kontrollevers resultater på motivasjon og matematikk ved pre- og posttest, nettopp for å avgjøre om effekter av inferenstreningen er reelle, eller kan karakteriseres som Hawthorneeffekter. Dette er en forutsetning for i det hele tatt å diskutere effekter av inferenstreningen på inferens og leseforståelse.

Det er allerede gjort rede for at t-test for uavhengige grupper ikke viste signifikante forskjeller mellom eksperimentgruppen og kontrollgruppen med hensyn til motivasjon på pre- eller posttest. Det var en sterk korrelasjon mellom pretestskårer og posttestskårer både for eksperimentgruppen ( $.70$ ,  $p<.001$ ) og kontrollgruppen ( $.67$ ,  $p<.001$ ). Spørreskjemaet om motivasjon ble ikke gjennomført i alle klassene verken på pretest eller posttest. Det var ikke hensiktsmessig å beregne endringsskåre fra pretest til posttest fordi dette ville ha redusert antall deltakere betraktelig. I stedet ble forholdet mellom endringsskåre for målene på effekt og motivasjon på pretest og posttest undersøkt med Pearson's  $r$ . Det var ingen signifikant korrelasjon mellom endringsskåre i inferens og totalskåre på motivasjon i eksperimentgruppen og kontrollgruppen verken på pretest eller posttest. Analyser av korrelasjon mellom endringsskåre (endring fra pretest til posttest) i inferens og de to konstruktene involvement og curiosity viste signifikant korrelasjon mellom endringsskåre i inferens og curiosity ved pretest i kontrollgruppen,  $.38$ ,  $p=.018$ . Ellers var det ingen signifikante korrelasjoner verken i eksperimentgruppen eller kontrollgruppen mellom endringsskåre i leseforståelse og motivasjon.

I likhet med motivasjon var det heller ikke signifikante forskjeller mellom gruppene verken ved pretest eller posttest med hensyn til matematikk. T-test tar imidlertid ikke høyde for skårer ved pretest når gruppene sammenlignes på

posttest. Enveis kovariansanalyse (ANCOVA) mellom grupper gir imidlertid mulighet for å sammenligne gruppene på posttest etter at det er kontrollert for skårer på pretest. Innledende analyser til ANCOVA viste at det ikke var brudd på forutsetningene med hensyn til normalfordeling, linearitet, homogenitet av varians og homogenitet av regresjonslinjer. Det var imidlertid en svært sterk korrelasjon (.84) mellom avhengig variabel, matematikk målt på posttest, og matematikk ved pretest.

*Tabell 26: Ikke-justerte og justerte gjennomsnitt for matematikk målt ved posttest*

	Ikke-justert			Justert	
	N	M	SD	M	SE
Eksperimentgruppe	135	56.35	11.23	56.70	.53
Kontrollgruppe	83	56.14	11.19	55.54	.68

Etter at det ble justert for pretestskårer i matematikk, ble forskjellen på matematikk målt ved posttest mellom eksperimentgruppe og kontrollgruppe noe større.

*Tabell 27: Enveis ANCOVA for posttestskårer i matematikk kontrollert for matematikk målt ved pretest*

Kilde til varians	df	M Kvadrert	F	P	Partiell $\eta^2$
Matematikk pretest	1	18981.12	496.70	<.001	.70
Gruppe	1	65.79	1.72	=.191	<.01
Error	214				

ANCOVA viste at det ikke var signifikant forskjell mellom eksperimentgruppe og kontrollgruppe i posttestskårer i matematikk etter at det var kontrollert for pretestskårer,  $F(1,214)=4.72$ ,  $p=.191$ .

Oppsummert tyder de to gruppens resultater på kontrollvariablene motivasjon og matematikk på at det ikke var generelle effekter av inferenstreningen. Det er derfor rimelig å konkludere med at effekter av inferenstreningen var spesifikt knyttet til lesing.

#### **7.4.2.4 Sammenligning av gutter og jenter**

For å undersøke om inferenstreningen bidro til å utjevne eller forsterke eventuelle kjønnsforskjeller, eller om et av kjønnene profitterte mer på inferenstreningen enn det andre, jevnfør problemstilling 10c, ble gutter og jenters skårer sammenlignet på de to tidspunktene.



Som det går fram av tabell 28, skåret jentene i utvalget jevnt over noe høyere enn guttene på alle variabler ved pretest, med unntak av for ordforråd hvor guttenes gjennomsnittskåre var noe høyere enn for jentene. Forskjellene mellom gutter og jenter i hele utvalget totalt er imidlertid svært små. Kjønnforskjeller for variabler ved pretest er oppsummert i tabellen under. T-test for uavhengige grupper viste at det kun var signifikante kjønnforskjeller med hensyn til tre variabler. I forhold til nonverbale ferdigheter var det signifikant forskjell mellom gutters og jenters skårer, i jentenes favør. Det var også signifikante forskjeller mellom gutter og jenter med hensyn til fiksjonstekstene i prøven i leseforståelse og på åpne spørsmål i inferensprøven, også her i jentenes favør.

Tabell 28: Gjennomsnitt og standardavvik for gutter og jenter, samt kjønnforskjeller for variablene målt ved pretest

	Gutter			Jenter			$\eta^2$	P
	N	M	SD	N	M	SD		
Ordforråd	110	12.55	3.03	112	12.17	3.05	.00	=.657
Ordavkodning	117	32.94	9.12	116	34.74	8.39	.01	=.118
Nonverbale ferdigheter	118	37.64	10.40	113	40.98	7.87	.03	=.006
Motivasjon, totalt	85	25.28	4.12	83	25.76	3.87	.00	=.445
Leseforståelse, totalt	117	21.48	5.64	114	22.43	5.42	.01	=.193
Leseforståelse, fakta	117	12.10	2.97	114	12.46	2.87	.00	=.347
Leseforståelse, tolkning	117	9.38	3.09	114	9.96	2.91	.00	=.137
Leseforståelse, fiksjon	117	12.12	3.26	114	13.05	2.90	.00	=.023
Leseforståelse, nonfiksjon	117	9.36	2.95	114	9.38	2.96	.00	=.963
Inferens, totalt	117	13.17	7.86	114	15.15	8.15	.00	=.062
Inferens, åpne spørsmål	117	9.09	5.95	114	10.75	5.73	.02	=.033
Inferens, understrekningsoppgaver	117	4.07	2.53	114	4.40	2.89	.00	=.369

Da eksperimentgruppen og kontrollgruppen ble analysert hver for seg, viste det seg at det var signifikant forskjell mellom gutter (N=75, M=12.03, SD=3.37) og jenter (N=72, M=13.36, SD=2.89) i eksperimentgruppen med hensyn til fiksjonstekster på prøven i leseforståelse målt ved pretest,  $t(145)=2.57$ ,  $p=.011$ ,  $\eta^2=.04$ . Det var kun kjønnforskjeller i forhold til åpne spørsmål på inferensprøven på pretesten i eksperimentgruppen (gutter: N=75, M=8.56, SD=6.00, jenter: N=72, M=11.54, SD=6.13),  $t(145)=2.98$ ,  $p=.003$ ,  $\eta^2=.06$ . Da eksperimentgruppen ble analysert isolert, var det også en signifikant forskjell mellom gutter (N=76, M=33.39, SD=9.25) og jenter (N=75, M=37.12, SD=8.10) med hensyn til ordavkodning  $t(149)=-2.632$ ,  $p=.009$ ,  $\eta^2=.04$ , i jentenes favør. I kontrollgruppen var det ingen signifikante kjønnforskjeller på variablene målt ved pretest.

Ved posttest var det ingen signifikante kjønnsforskjeller da utvalget ble analysert samlet. Da gruppene ble analysert hver for seg, var det ingen signifikante kjønnsforskjeller i kontrollgruppen, mens i eksperimentgruppen skåret jentene (N=70, M=10.17, SD=2.69) signifikant høyere på tolkningsspørsmål på prøven i leseforståelse enn guttene (N=73, M=9.03, SD=3.02)  $t(141)=-2.39$ ,  $p=.018$ ,  $\eta^2=.04$ . På øvrige variabler var det ingen kjønnsforskjeller i eksperimentgruppen.

#### 7.4.2.5 Sammenligning av enspråklige og tospråklige elever

I underkant av 10 prosent av deltakerne i utvalget snakket et annet morsmål enn norsk eller skandinavisk hjemme. Skårene til enspråklige og tospråklige ble sammenlignet ved pre- og posttest. Dette for å kunne undersøke om det var forskjell mellom de to gruppene elever med hensyn til hvilken gruppe som eventuelt profitterte mest og minst på inferenstreningen, og om eventuelle forskjeller mellom enspråklige og tospråklige elever ble utjevnet eller forsterket etter inferenstreningen.

Lavt antall tospråklige elever i utvalget begrenset mulighetene for å bruke parametrisk statistikk, derfor ble Mann Whitney U test for uavhengige grupper benyttet som et ikke-parametrisk alternativ til t-test for å sammenligne enspråklige og tospråklige sine skårer. Styrken på forskjeller mellom tospråklige i eksperimentgruppen og kontrollgruppen ble beskrevet ved hjelp av Cohen's *d*. Gjennomsnitt, standardavvik og signifikansnivå er oppgitt i tabellen under.

Tabell 29: Gjennomsnitt, standardavvik og signifikansnivå hos enspråklige og tospråklige målt ved pretest

	Enspråklige			Tospråklige			P	<i>d</i>
	N	M	SD	N	M	SD		
Ordforråd	204	12.46	2.89	18	11.17	4.32	=.167	0.35
Ordavkoding	212	33.83	8.71	21	33.95	9.79	=.970	-0.01
Nonverbale ferdigheter	210	39.63	8.94	21	35.71	12.78	=.199	0.36
Motivasjon	151	25.33	4.06	17	27.20	3.01	=.057	-0.52
Leseforståelse, totalskåre	211	22.24	5.23	20	18.85	7.67	=.052	0.51
Leseforståelse, fakta	211	12.43	2.76	20	10.70	4.00	=.050	0.50
Leseforståelse, tolkning	211	9.81	2.86	20	8.15	4.06	=.072	0.47
Leseforståelse, fiksjon	211	12.75	2.94	20	10.80	4.26	=.050	0.53
Leseforståelse, nonfiksjon	211	9.49	2.82	20	8.05	3.86	=.070	0.43
Inferens, totalt	211	14.32	7.98	20	12.34	8.68	=.245	0.68
Inferens, åpne spørsmål	211	10.03	5.82	20	8.65	6.64	=.282	0.22
Inferens, understrekningsoppgaver	211	4.29	2.73	20	3.69	2.53	=.335	0.23

Mann Whitney U test viste at det ved pretest kun var signifikante forskjeller mellom enspråklige og tospråklige elever på faktaspørsmål og fiksjonstekster på prøven i leseforståelse. Denne forskjellen var i favør av de enspråklige og kan karakteriseres som moderat.

På posttest var det ingen signifikante forskjeller mellom enspråklige og tospråklige i kontrollgruppen. I eksperimentgruppen var det imidlertid signifikante forskjeller mellom enspråklige og tospråklige med hensyn til totalskåre i leseforståelse. Enspråklige (N=129, gjennomsnittlig rangering=74.30) skåret signifikant høyere enn tospråklige (N=14, gjennomsnittlig rangering=50.82) på leseforståelse totalt,  $U=606.50$ ,  $p=.043$ . Når de to spørsmålstypene på prøven i leseforståelse ble analysert hver for seg, viste det seg at det kun var signifikante forskjeller mellom enspråklige (gjennomsnittlig rangering=74.97) og tospråklige (gjennomsnittlig rangering=44.64) med hensyn til faktaspørsmål,  $U=520.00$ ,  $p=.008$ , mens det ikke var signifikante forskjeller med hensyn til tolkningsspørsmål. Da de to teksttypene ble analysert hver for seg, var det signifikante forskjeller mellom enspråklige (gjennomsnittlig rangering=74.78) og tospråklige (gjennomsnittlig rangering=46.39) for nonfiksjonstekster,  $U=544.50$ ,  $p=.014$ , men ikke for fiksjonstekster.

### **7.4.3 Effekter av inferenstreningen**

I dette avsnittet vil det bli gjort rede for endring fra pretest til posttest både i forhold til primæreffekt og generalisert effekt. Med andre ord om inferenstreningen hadde effekt på inferens og leseforståelse, jevnfør problemstilling 8.

#### **7.4.3.1 Måling av effekt**

For å sammenligne individuell endring i inferens og leseforståelse var det nødvendig å transformere skårene på effektmålene. Dette for å kunne sammenligne de to prøvene i leseforståelse og inferensprøven, som ble gjennomført på to tidspunkter. Log odds transformering ble valgt fordi det ga muligheten for sammenligne individuell endring. Fordelen med å bruke log odds framfor å sammenligne endringsskårer basert på råskårer, er at log odds tar høyde for at den enkeltes mulighet for framgang ikke er den samme, men bestemmes av individets resultat på pretesten. For eksempel vil en elev som har 80 % riktige skårer på pretest ha mulighet for kun å gå fram med 20 % ved posttest. Motsatt vil en elev som har 20 % riktig ved pretesten ha mulighet til å gå fram 80 % på posttest. Log odds transformasjonen vil gi alle elevene lik mulighet til endring (Allerup & Elbro, 1998).

For å beregne log oddsskåre for de to effektvariablene ble skårene på pre- og posttest først transformert til odds. Odds ble beregnet ved å dividere prosent skåre riktig med prosent feil skåre. Videre ble logaritmen av oddsen tatt for å skalere variabelen på intervallnivå. Endringsskåre ble beregnet ved å subtrahere log oddstransformerte posttestskåre fra log oddstransformerte pretestskåre. Endringsskåren strekker seg fra minus uendelig til pluss uendelig. En positiv endringsskåre kan tolkes som framgang fra pretest til posttest, mens en negativ endringsskåre kan tolkes som tilbakegang fra pretest til posttest. En endringsskåre på 0 kan tolkes som ingen endring fra pretest til posttest. I tolkningen av log odds transformerte endringsskåre i leseforståelse er det imidlertid viktig å merke seg at prøvene på pre- og posttest var parallelle, men ikke identiske. Hvorvidt prøvene har eksakt samme vanskelighetsgrad vites ikke. Med andre ord er det usikkert hvorvidt en endringsskåre i leseforståelse på 0 poeng faktisk betyr ingen endring eller om en endringsskåre på 0 i virkeligheten representerer en tilbakegang eller framgang.

Når man studerer endring fra pretest til posttest, utgjør ekstremskåre på effektmålene et problem fordi det ikke er mulig å vite hvor langt under minimumsskåren eller over maksimumsskåren en elevs prestasjon virkelig ligger, og følgelig vet man da ingenting om elevens virkelige endring fra pretest til posttest. I analyser av framgang vil derfor elever med ekstremskåre, 0-skåre og toppskåre, bli utlatt. Ved pretest var det tre elever, to fra eksperimentgruppen og en fra kontrollgruppen, som hadde 0-skåre på inferensprøven. Ingen elever hadde 0-skåre på posttestene, verken på inferens, totalskåre på leseforståelse eller på sumskåre på faktaspørsmål og tolkningsspørsmål. På de åpne spørsmålene på inferensprøven hadde fem eksperimentelever 0-skåre på pretesten og en eksperimentelev hadde 0-skåre på både på pre- og posttest. To kontrollelever hadde 0-skåre på pretest, og en kontrollelev hadde 0-skåre på de åpne spørsmålene ved posttest. Ni eksperimentelever og tre kontrollelever hadde 0-skåre på understreknings-spørsmålene på inferensprøven ved pretest, og to eksperimentelever og to kontrollelever hadde 0-skåre på understrekningsoppgavene på posttesten. Ingen elever hadde toppskåre på totalskåre på leseforståelse verken på pretest eller posttest. En eksperimentelev og fire kontrollelever hadde toppskåre på faktaspørsmål ved pretest. Tolv eksperimentelever og to kontrollelever hadde toppskåre på faktaspørsmål ved posttest, og en eksperimentelev hadde toppskåre på faktaspørsmål både på pre- og posttest. Det var ingen elever som hadde toppskåre på tolkningsspørsmål på pretest, men en eksperimentelev hadde toppskåre på tolkningsspørsmål ved posttest. Elever med 0-skåre og elever med toppskåre ble utlatt fra analysene med log odds transformerte

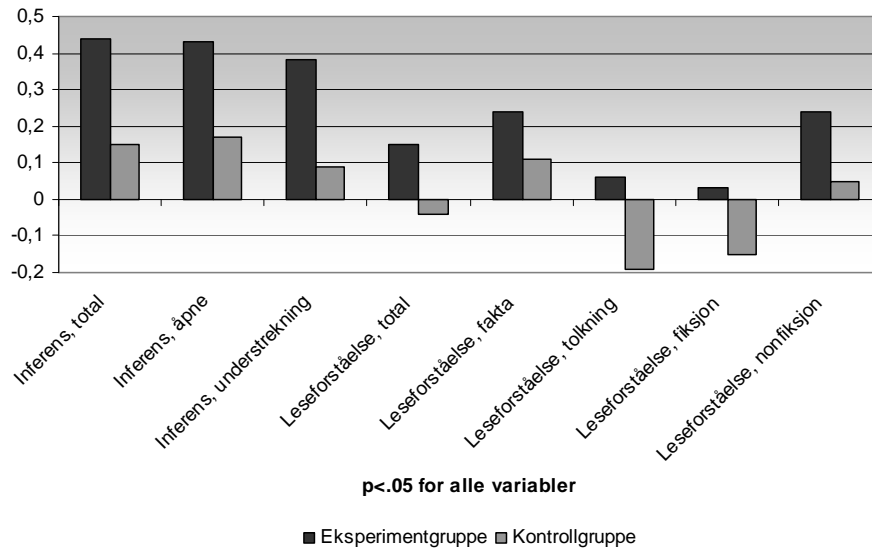
skårer. Grunnen til at elever med 0-skårer på pretesten ble utelatt fra analysene, var at framgangen til disse elevene slik de er målt med effektmålene vil være mindre enn deres virkelige framgang, siden målet på pretesten ikke hadde fanget opp hvor lav elevens egentlige prestasjon var ved pretest. Grunnen til at elever med toppskåre på pretest ble utelatt, var at målet på effekt ikke kunne fange opp disse elevenes framgang siden de ved første måletidspunkt hadde nådd toppskåren. Elever med toppskåre på posttest ble utelatt fordi målet på effekt ikke fanger opp hvor mye over toppskåre prestasjonen virkelig er, og dermed er den målte framgangen nødvendigvis mindre enn den virkelige framgangen.

#### **7.4.3.2 Forskjeller mellom eksperimentgruppe og kontrollgruppe**

I dette avsnittet vil det bli gjort rede for resultater som belyser problemstilling 8, effekten av inferenstreningen for inferens og leseforståelse. Med andre ord om eksperimentgruppen hadde en større framgang fra pretest til posttest på effektmålene enn kontrollgruppen.

Det er flere måter å undersøke om endringen på effektmålene er forskjellig i de to gruppene. For det første ved å undersøke interaksjonseffekter mellom de to måletidspunktene og de to gruppene, med andre ord hvorvidt endringen over tid er forskjellig i eksperimentgruppe og kontrollgruppe. For det andre kan forskjellen mellom de to gruppenes endring uttrykkes ved hjelp av effektstørrelser, beregnet ved hjelp av Cohen's  $d$ . Cohen's  $d$  gir informasjon om styrken på forskjellen mellom eksperimentgruppens og kontrollgruppens endring fra pretest til posttest, i tillegg gir Cohen's  $d$  mulighet for å kunne sammenligne resultatene fra denne effektstudien med tidligere effektstudier av inferenstrening som det er gjort rede for i kapittel 5.

Søylediagrammet i figur 4 viser log odds transformerte endringsskåre hos eksperimentgruppen og kontrollgruppen for inferensprøven og prøven i leseforståelse, samt delsummer for spørsmålstyper og teksttyper. Eksperimentgruppen hadde signifikant høyere endringsskåre enn kontrollgruppen på samtlige samleskåre og delskåre. Kontrollgruppen hadde tilsynelatende hatt en tilbakegang fra pretest til posttest på prøven i leseforståelse. Siden prøven i leseforståelse som ble administrert ved pre- og posttest ikke var den samme, betyr dette at tilbakegangen ikke nødvendigvis var reell, den kan skyldes at vanskelighetsgraden på prøven var noe høyere ved posttest.



Figur 4. Log odds transformerte endringsskårer for eksperimentgruppe og kontrollgruppe

For å undersøke om endringen i effektmålene over de to måletidspunktene var signifikant forskjellig i eksperimentgruppe og kontrollgruppe, ble det benyttet et blandingsdesign hvor between-subjects faktorer og within-subjects faktorer ble kombinert. Ved hjelp av to analyser av varians (ANOVA) med gjentakende målinger ble det undersøkt hvorvidt det var en signifikant interaksjon mellom tid (pre- og posttest) og gruppe (eksperimentgruppe og kontrollgruppe). Det ble gjennomført en analyse for endring i inferens og en for endring i leseforståelse. ANOVA med gjentakende målinger gir informasjon om hovedeffekter av de to variablene og interaksjonen mellom variablene. En signifikant interaksjonseffekt innebærer at endringen i effektmålet er avhengig av gruppe. Kvadrert eta ble beregnet for de tre effektene. For å holde variansen fra between-subjects variabelen og within-subjects variabelen separat, ble kvadrert eta beregnet ved følgende formler (Gamst, Meyers, & Guarino, 2008, s. 363-364):

$$\eta^2_A = \frac{SS_A}{SS_{\text{Between-subjects}}}$$

$$\eta^2_B = \frac{SS_B}{SS_{\text{Within-subjects}}}$$

$$\eta^2_{A \times B} = \frac{SS_{A \times B}}{SS_{\text{Within-subjects}}}$$

ANOVA for inferens (eksperimentgruppe: N=135, kontrollgruppe: N=82) viste at det var en signifikant effekt av between-subjects variabelen (gruppe),  $F(1,215)=7.38$ ,  $p=.007$ ,  $\eta^2=.03$ . Effekten av within-subjects variabelen (tid) var også signifikant,  $F(1,215)=167.82$ ,  $p<.001$ ,  $\eta^2=.39$ . Interaksjonseffekten mellom de to variablene var signifikant,  $F(1,215)=40.88$ ,  $p<.001$ ,  $\eta^2=.10$ , hvilket betyr at endringen i de to gruppene var forskjellig. ANOVA for leseforståelse (eksperimentgruppe: N=141, kontrollgruppe: N=83) viste signifikant effekt av between-subjects variabelen (gruppe),  $F(1,222)=11.06$ ,  $p=.001$ ,  $\eta^2=.05$ . Det var også signifikant effekt av within-subjects variabelen (tid),  $F(1,222)=7.77$ ,  $p=.006$ ,  $\eta^2=.03$ . Det var en signifikant interaksjonseffekt av tid og gruppe,  $F(1,222)=23.42$ ,  $p<.001$ ,  $\eta^2=.09$ .

Cohen's  $d$  ble beregnet ved å dividere differansen mellom gruppenes gjennomsnittlige endringsskåre på samlet standardavvik. Endringsskåre er differansen mellom to måletidspunkter, resultater på posttest subtrahert fra resultater på pretest. De endringsskårene som det henvises til her er log odds transformerte skårer på de to måletidspunktene.

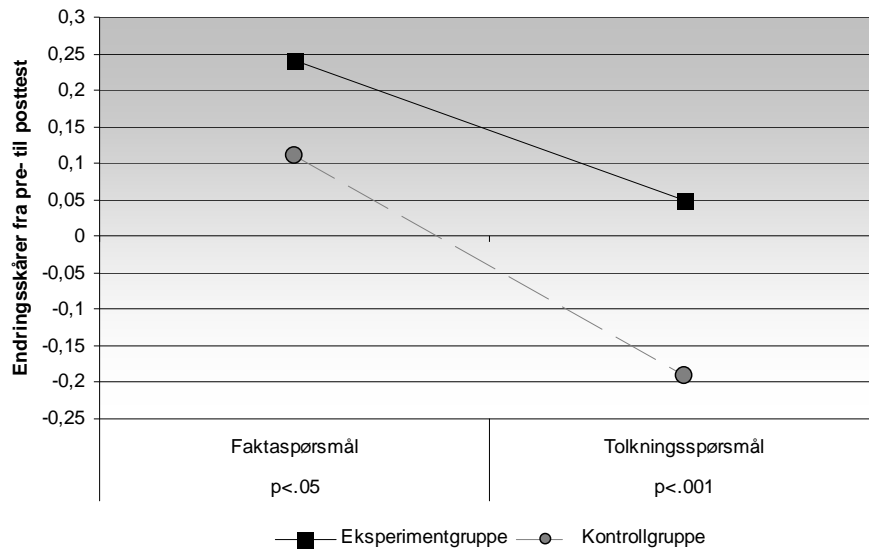
Cohen's  $d$  beregnet på log odds endringsskårer var .92 for totalskåre på prøven i inferens, hvilket tilsvarer sterk effekt. Cohen's  $d$  ble også beregnet for de to typene item som prøven i inferens besto av. Effekten for åpne spørsmål var middels,  $d=.43$ , mens effekten var sterk for understrekningsoppgavene,  $d=.82$ . Cohen's  $d$  for totalskåre på prøven i leseforståelse var .69, hvilket kan karakteriseres som en middels til sterk effekt. Effektstørrelsene for de to spørsmålstypene i prøven i leseforståelse var .45 og .73 for henholdsvis faktaspørsmål og tolkningsspørsmål. Det ble også beregnet effektstørrelser for de to teksttypene i prøven i leseforståelse. Cohen's  $d$  ble beregnet til .46 og .57 for henholdsvis fiksjonstekster og nonfiksjonstekster.

Oppsummert hadde eksperimentgruppen generelt høyere endringsskårer, altså større framgang, enn kontrollgruppen. Resultater fra ANOVA viste signifikante interaksjoner mellom tid og gruppe, hvilket innebærer at eksperimentgruppens framgang var signifikant større enn kontrollgruppens. Beregning av Cohen's  $d$  bekreftet resultatene fra ANOVAen.

### ***Sammenheng mellom endring på faktaspørsmål og tolkningsspørsmål***

Dette avsnittet har til hensikt å belyse forholdet mellom de to spørsmålstypene i prøven i leseforståelse, faktaspørsmål og tolkningsspørsmål. Som det går fram av figur 5, var gruppenes endringsskårer fra pretest til posttest signifikant forskjellige i forhold til begge spørsmålstyper. Figuren viser gruppenes endringsskårer beregnet på råskårer for de to spørsmålstypene hvor

elever med ekstremskårer er inkludert. Endringsskårene for tolkningsspørsmålene er negative. Dette betyr ikke nødvendigvis en tilbakegang, det kan like gjerne være et resultat av at prøven i leseforståelse som ble administrert ved pre- og posttest ikke var den samme.



Figur 5. Eksperimентgruppens og kontrollgruppens endringsskårer (pretest til posttest) basert på råskårer på faktaspørsmål og tolkingsspørsmål

Forholdet mellom framgangen for de to typer item ble undersøkt med Pearson's r. Korrelasjonsanalysene ble gjort på bakgrunn av log odds transformerte skårer av endringsskåre fra pretest til posttest. 16 eksperimentelever og 7 kontrollelever ble utelatt fra analysen grunnet toppskårer på pre- og/eller posttest, ingen hadde skårer på 0. Eksperimентgruppe og kontrollgruppe ble analysert hver for seg. For eksperimентgruppen var det en middels korrelasjon mellom framgang på faktaspørsmål og framgang på tolkingsspørsmål (N=126,  $r=.40$ ,  $p<.001$ ), hvor høy skåre på faktaspørsmål assosieres med høy skåre på tolkingsspørsmål. For kontrollgruppen var det derimot ingen signifikant korrelasjon mellom de to typer item (N=76,  $r=.11$ ,  $p=.330$ ). Analysen ble gjentatt med endringsskåre basert på råskårer, resultatet sammenfallende som ved log odds transformerte endringsskåre.

Korrelasjonsanalysene kan ikke gi svar på hvorvidt forskjell i endring på de to spørsmålstypene var avhengig av gruppe, med andre ord om mønsteret i



endring i faktaspørsmål og tolkningsspørsmål var forskjellig i de to gruppene. Som det går fram av figur 5, er linjene som representerer de to gruppene forholdsvis like, men kontrollgruppen har en noe bråere stigning. En analyse av interaksjonen mellom spørsmålstype og gruppe vil avsløre om graden av stigning på de to linjene er signifikant forskjellige, eller om de to linjene kan betraktes som parallelle. For å undersøke dette ble det benyttet et blandingsdesign med kombinasjon av between-subjects og within-subjects faktorer. Ved hjelp av ANOVA med gjentakelse ble det undersøkt hvorvidt det var en signifikant interaksjon mellom gruppe (eksperiment og kontroll) og spørsmålstype (tolkning og fakta).

Gruppe var between-subjects variabel, og endringsskåre i spørsmålstype var within-subjects variabel (eksperimentgruppe: N=126, kontrollgruppe: N=76). Fordi resultatene ikke var sammenfallende når analysene ble gjort med endringsskårer basert på råskårer og med endringsskårer basert på log odds endringsskårer, vil resultatene fra analyser med begge typer endringsskårer bli rapportert i dette avsnittet.

Det var en signifikant effekt av between-subjects variabelen (gruppe) både når analysen ble kjørt med endringsskårer basert på råskårer,  $F(1,200)=10.14$ ,  $p<.001$ ,  $\eta^2=.08$ , og når analysen ble kjørt med log odds transformerte endringsskårer,  $F(1,200)=19.72$ ,  $p<.001$ ,  $\eta^2=.09$ . Effekten av within-subjects variabelen (spørsmålstype) var også signifikant både i analysen av endringsskårer basert på råskårer  $F(1,200)=189.26$ ,  $p<.001$ ,  $\eta^2=.48$ , og i analysen av log odds transformerte endringsskårer,  $F(1,200)=60.77$ ,  $p<.001$ ,  $\eta^2=.23$ . Effekten av interaksjonen mellom de to variablene var signifikant da analysen ble kjørt med endringsskårer basert på råskårer,  $F(1,200)=6.46$ ,  $p=.012$ , men effekten av interaksjonen var liten,  $\eta^2=.02$ . Dette betyr at forholdet mellom endring i faktaspørsmål og endring i tolkningsspørsmål ikke var det samme i de to gruppene. En signifikant interaksjon mellom gruppe og spørsmålstype innebærer altså at eksperimentgruppen ikke fulgte samme utvikling som kontrollgruppen. Endring på den ene typen spørsmål bidrar med endring på den andre typen spørsmål og omvendt. Da interaksjonen mellom gruppe og spørsmålstype ble undersøkt med log odds transformerte skårer, var imidlertid resultatet ikke signifikant,  $F(1,200)=2.80$ ,  $p=.096$ ,  $\eta^2=.01$ . Dette resultatet betyr at endringen i de to gruppene var parallell, det vil si at endringen ikke var forskjellig for de to spørsmålstypene. En parallell endring i gruppene innebærer at forskjellen mellom eksperimentgruppen og kontrollgruppen var en generell nivåforskjell. Denne ikke-signifikante interaksjonen mellom gruppe og spørsmålstype innebærer at eksperiment-

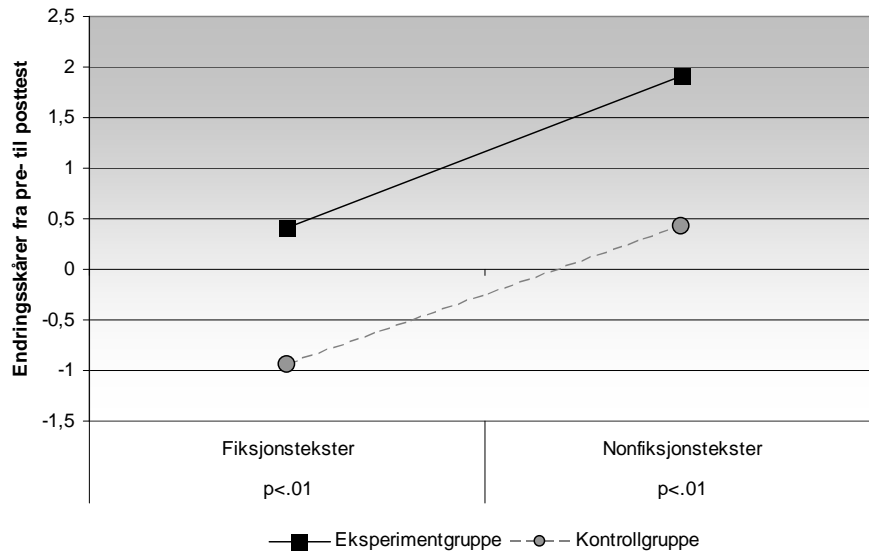
gruppen fulgte den naturlige utviklingen, her representert ved kontrollgruppen, mens inferenstreningen har påskyndet utviklingen.

Grunnen til at resultatet av ANOVA med transformerte og ikke-transformerte skårer blir forskjellig, kan ha med gruppenes fordeling på de to spørsmålstypene på de to tidspunktene å gjøre. Det var mange elever med høye skårer på faktaspørsmålene, særlig på posttesten i eksperimentgruppen. På tolkningsspørsmålene var det imidlertid ikke like mye takeffekt. Hos elevene med toppskåre vet man ikke hva den virkelige skåren er, kanskje blir den virkelige ferdigheten hos gruppen med høye skårer også undervurdert. Dersom prøven ikke hadde hatt et "tak", ville de to linjene i figuren over trolig vært enda mer parallelle fordi det er sannsynlig at eksperimentelevenes endringsskåre på faktaspørsmålene ville ha vært høyere, og stigningen på eksperimentgruppens linje i figuren over ville ha vært bråere. Log oddstransformeringen bidrar nettopp med å kompensere for takeffekter ved å "fjerne" taket. Gruppenes fordeling på de to spørsmålene på pre- og posttest er nærmere beskrevet i avsnittet 7.4.2. De elevene som skåret høyt på faktaspørsmål på pretest, hadde mindre mulighet for å forbedre seg på faktaspørsmål enn på tolkningsspørsmål ved posttest. Odds har nettopp til hensikt å utjevne forskjeller mellom den enkeltes mulighet for framgang. Dermed blir forskjellene mellom log oddsskårene hos elever med høye skårer både på pre- og posttest større enn hos elever med lavere skårer. Dette lar seg best illustrere med et eksempel: Elev A har en total poengsum på 15 på faktaspørsmål på pretest og 17 poeng på posttest, og får log oddsskårer på henholdsvis 0.88 og 1.23 på pre- og posttest. Endringsskåre basert på råskårer blir for denne eleven 2 poeng. Endringsskåre basert på log oddsskårer blir 0.35 poeng. Elev B har en total poengsum på 14 på faktaspørsmål på pretest og 16 poeng på posttest, og får log oddsskårer på pre- og posttest på henholdsvis 0.67 og 0.90. Endringsskåre basert på råskåre for elev B blir, i likhet med elev A, 2 poeng. Endringsskåre basert på log oddsskårer blir 0.23 for elev B. Elev C har 13 og 15 råskårepoeng på faktaspørsmålene på pre- og posttest og får log oddsskårer på 0.51 og 0.70. Endringsskåren blir fremdeles 2 poeng når den beregnes på råskårer, men kun 0.19 når den beregnes på log oddsskårer. Selv om differansen mellom råskårene på pre- og posttest er den samme, blir log odds endringsskåren mindre jo lengre ned på skalaen man kommer, og høyere desto lengre opp på skalaen man kommer. Siden det var flere elever øverst på skalaen på faktaspørsmål enn på tolkningsspørsmål, blir log oddsskårene justert ned for flere elever på faktaspørsmålene enn for tolkningsspørsmålene. Dermed blir ikke de to spørsmålstypene helt sammenlignbare.

Fordi det var flere elever som hadde endring fra pretest til posttest på øvre del av skalaen på faktaspørsmålene enn på tolkningsspørsmålene, ble utvalget delt inn i to grupper etter deres skåre på faktaspørsmålene på pretesten. Dette var for å undersøke om interaksjonen mellom spørsmålstype og gruppe var den samme blant elever med lav skåre på faktaspørsmålene på pretest og elever med høy skåre. De to gruppene "høy" og "lav" ble definert ut fra utvalgets gjennomsnitt på faktaspørsmålene på pretesten. Gruppen med lavt gjennomsnitt besto av elever med skårer under utvalgets gjennomsnitt, og gruppen med høyt gjennomsnitt besto av elever med skårer over utvalgets gjennomsnitt. I eksperimentgruppen var det 64 elever i kategorien lav skåre på faktaspørsmålene ved pretest, og 68 elever i kategorien høy skåre. I kontrollgruppen var det 41 elever med lav skåre og 36 med høy skåre. ANOVA med gjentakende målinger ble gjennomført for de to gruppene høy (eksperimentgruppe: N=67, kontrollgruppe: N=35) og lav (eksperimentgruppe: N=59, kontrollgruppe: N=41) skåre på faktaspørsmål på pretest, med både endringsskårer basert på råskårer og endringsskårer basert på log odds endringsskårer. Resultatene fra ANOVA med endringsskårer basert på råskårer viste at det var en signifikant interaksjon mellom gruppe og spørsmålstype blant elever med høy skåre på faktaspørsmålene ved pretest,  $F(1,100)=7.199$ ,  $p<.01$ ,  $\eta^2=.03$ . Det var imidlertid ingen signifikant interaksjon mellom spørsmålstype og gruppe blant elever med lav skåre på faktaspørsmålene ved pretest. Da analysene ble gjentatt med log odds transformerte endringsskårer var interaksjonen ikke signifikant, verken for elever med høy eller lav skåre på faktaspørsmålene ved pretest. Disse resultatene støtter antakelsen om at log oddsskårene bidrar til å "fjerne" takeffekt slik at forholdet mellom de to spørsmålstypene i eksperimentgruppen og kontrollgruppen blir parallell, og dermed blir interaksjonen ikke signifikant.

### ***Sammenheng mellom endring på fiksjonstekster og nonfiksjonstekster***

I dette avsnittet belyses forholdet mellom framgang på de to ulike teksttyper i prøven i leseforståelse. Figur 6 viser gruppenes endringsskårer basert på råskårer fra pre- til posttest på fiksjonstekster og nonfiksjonstekster.



Figur 6. Eksperimentgruppens og kontrollgruppens endringsskårer (pretest til posttest) basert på råskårer på fiksjonstekster og nonfiksjonstekster

Pearson's  $r$  ble beregnet på log odds transformerte skårer. 17 eksperimentelever og 2 kontrollelever ble utelatt fra analysen på grunn av toppskårer på pre- og/eller posttesten, ingen elever hadde skårer på 0. I eksperimentgruppen var det en middels, signifikant korrelasjon ( $N=124$ ,  $r=.25$ ,  $p<.01$ ) mellom de to teksttypene, hvor høy skåre på den ene typen tekster assosieres med høy skåre på den andre typen tekst. I kontrollgruppen var korrelasjonen mellom fiksjonstekster og nonfiksjonstekster ikke signifikant ( $N=81$ ,  $r=.11$ ,  $p=.334$ ). Analysen ble gjentatt med endringsskårer basert på råskårer, resultatet sammenfallende som ved log odds transformerte endringsskårer. Som det allerede er gjort rede for, ble Cohen's  $d$ , beregnet til å være middels for begge typer tekster.

Interaksjonen mellom teksttype og gruppe ble undersøkt ved hjelp av ANOVA etter samme prosedyre som for spørsmålstyper, jevnfør avsnittet over (eksperimentgruppe:  $N=124$ , kontrollgruppe:  $N=81$ ). Det var en signifikant effekt av gruppe både når analysen ble gjort med råskårer  $F(1,203)=21.46$ ,  $p<.01$ ,  $\eta^2=.10$ , og log odds transformerte skårer,  $F(1,203)=20.45$ ,  $p<.01$ ,  $\eta^2=.09$ . Det var også signifikant effekt av teksttype både ved råskårer,  $F(1,203)=31.07$ ,  $p<.01$ ,  $\eta^2=.13$ , og log odds transformerte skårer,  $F(1,203)=35.39$ ,  $p<.01$ ,  $\eta^2=.15$ . I motsetning til analysene av spørsmålstyper, jevnfør avsnittet over, var det sammenfallende resultat med

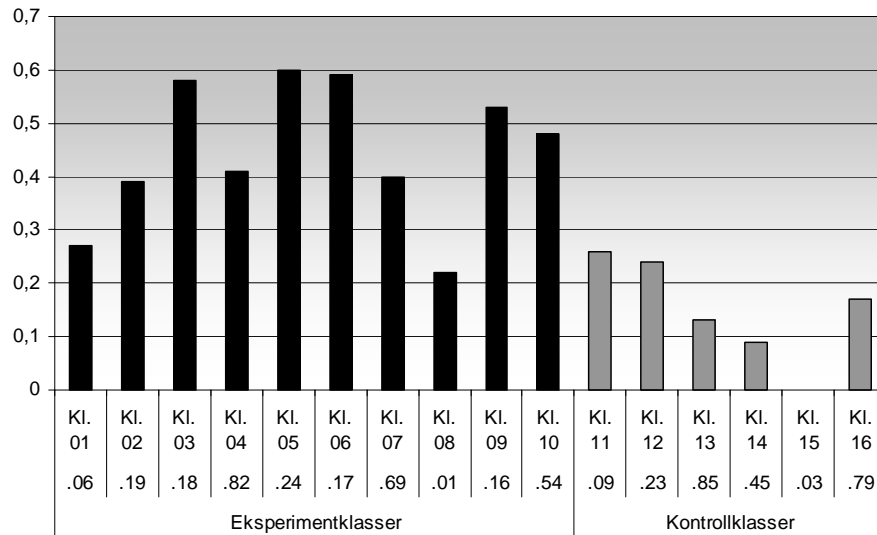
hensyn til interaksjon beregnet på råskårer,  $F(1,203)=0.18$ ,  $p<.670$ ,  $\eta^2>.01$ , og log odds transformerte skårer,  $F(1,203)=0.001$ ,  $p<.973$ ,  $\eta^2>.01$ . Det var ingen signifikant interaksjon mellom teksttype og gruppe, verken når analysen ble gjort med råskårer eller log odds transformerte skårer.

#### **7.4.3.3 Endring i eksperimentklasser og kontrollklasser**

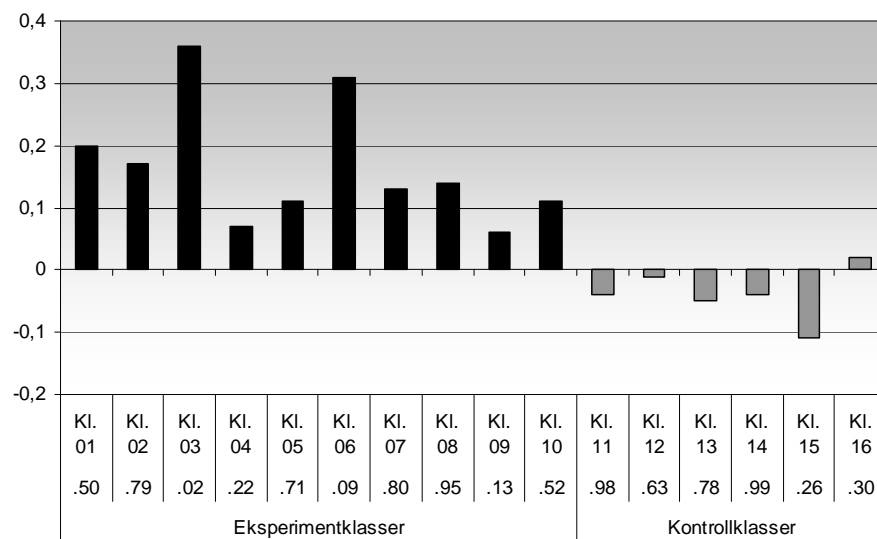
For å undersøke om enkelte klasser hadde gått særlig mye fram eller tilbake fra pretest til posttest sammenlignet med øvrige klasser i samme gruppe, ble klassenes gjennomsnittlige skårer beregnet og sammenlignet ved hjelp av t-test og Mann-Whitney U test. Disse analysene ble gjennomført for å undersøke om eksperimentgruppens og kontrollgruppens tendenser skyldtes særlig framgang eller tilbakegang i enkeltklasser.

Figur 7 viser klassenes log odds transformerte endringsskårer på inferensprøven fra pretest til posttest. T-tester for uavhengige grupper ble gjennomført for å sammenligne de enkelte klassenes endringsskåre med øvrige klasser i samme gruppe. Eksperimentklasse 08 hadde signifikant lavere endringsskåre ( $N=15$ ,  $M=.22$ ,  $SD=.22$ ) enn øvrige eksperimentklasser ( $N=120$ ,  $M=.46$ ,  $SD=.34$ ),  $t(133)=2.73$ ,  $p=.007$ ,  $\eta^2=.05$ . Resultatet ble bekreftet av Mann Whitney U test. T-test viste at forskjellen mellom eksperimentklasse 01 ( $N=13$ ,  $M=.27$ ,  $SD=.20$ ) og øvrige eksperimentklasser ( $N=122$ ,  $M=.45$ ,  $SD=.34$ ) ikke var signifikant. Derimot viste Mann Whitney U test at forskjellene mellom klasse 01 og øvrige eksperimentklasser var signifikant,  $U=506.50$ ,  $p=.033$ . T-test viste at kontrollklasse 15 ( $N=14$ ,  $M=.00$ ,  $SD=.32$ ) hadde en signifikant lavere endringsskåre enn øvrige klasser i kontrollgruppen ( $N=68$ ,  $M=.18$ ,  $SD=.28$ ),  $t(80)=2.19$ ,  $p=.031$ ,  $\eta^2=.06$ . Mann Whitney U test viste imidlertid at forskjellen mellom klasse 15 og øvrige kontrollklasser ikke var signifikant.

Figur 8 viser klassenes endringsskårer basert på log odds på prøven i leseforståelse fra pretest til posttest. T-tester for uavhengige grupper ble gjennomført for å sammenligne de enkelte klassenes endringsskåre med øvrige klasser i samme gruppe. Klasse 03 skilte seg ut med signifikant høyere endringsskåre ( $N=11$ ,  $M=0.36$ ,  $SD=0.33$ ) enn øvrige eksperimentklasser ( $N=130$ ,  $M=0.13$ ,  $SD=0.30$ ),  $t(139)=-2.409$ ,  $p=.017$ ,  $\eta^2=.04$ . Resultatet ble bekreftet av Mann Whitney U test. Det var ingen signifikante forskjeller mellom kontrollklassenes endringsskårer.



Figur 7. Sammenligning av klassenes log odds transformerte endringsskårer fra pretest til posttest på inferensprøven, hvor signifikansnivå er basert på t-test



Figur 8. Sammenligning av klassenes log odds transformerte endringsskårer fra pretest til posttest på prøven i leseforståelse, hvor signifikansnivå er basert på t-test

#### **7.4.3.4 Sammenhenger mellom klasseopplysninger og eksperimentklassenes endring**

Sammenhenger mellom klassenes endringsskårer og opplysninger om lærerne i eksperimentgruppen, opplysninger om eksperimentklassene, inferenstreningens varighet og intensitet og klassenes skårer på kontrollvariablene ved pretest ble undersøkt med korrelasjonsanalyser. Hensikten med analysene var å undersøke sammenhengen mellom rammene for undervisning og framgangen til de ulike eksperimentklassene.

Fordi klassene ble brukt som enheter i korrelasjonsanalysene, ble korrelasjonene beregnet ved hjelp av Spearman's rho, som er en ikke-parametrisk korrelasjonskoeffisient, og dermed en mer konservativ analyse enn for eksempel Pearson's r. Klassenes gjennomsnittlige endringsskårer på effektmålene ble beregnet for log odds transformerte endringsskårer.

Det forelå tre opplysninger om lærerne i eksperimentgruppen, antall år som lærer, hadde vært klassens kontaktlærer og lærers totale undervisningserfaring. Med unntak av én lærer hadde samtlige lærere kun vært kontaktlærer for klassen i ett år eller mindre da undersøkelsen startet opp. Variasjonen på denne variabelen var derfor for liten til at det var hensiktsmessig å undersøke korrelasjoner med andre variabler. Det var ingen korrelasjon mellom klassens endringsskårer og læreres totale undervisningserfaring. Det var en variabel som angikk selve klassen, nemlig antall elever i klassen. Det var en signifikant negativ korrelasjon mellom antall elever i klassen og elevenes log odds endringsskårer i leseforståelse,  $-0.95$ ,  $p < .001$ . Dette betyr at høy skåre på endring i leseforståelse assosieres med lavt antall elever i klassen. Her må det imidlertid tas i betraktning at analysen er basert på kun ti klasser og at den høye korrelasjonen i hovedsak bygger på de to klassene med høyest antall elever i klassen (se tabell 30). Det var ikke signifikant korrelasjon mellom antall elever i klassen og endring i inferens.

*Effekter av inferenstrening blant norske sjetteklassinger*

*Tabell 30: Eksperimentklassenes endringsskårer og opplysninger om klasser og lærere i eksperimentgruppen*

Klasse	Log odds endringsskårer		Inferenstreningens varighet		Opplysninger om klassenes kontaktlærere		
	Inferens	Leseforståelse	Uker	Tid per leksjon	Ant. elever i klassen	Ant. år som kontaktlærer	Total undervisnings-erfaring
01	.27	.20	4 uker	39 min.	15	0	0
02	.40	.17	3 uker	44 min.	15	1	5.5
03	.58	.36	4 uker	44 min.	14	1	13.5
04	.41	.07	4 uker	33 min.	22	1	2.5
05	.60	.11	7 uker	-	19	0.5	3
06	.59	.31	8 uker	-	13	1	12
07	.40	.13	4 uker	25 min.	15	0	23.5
08	.22	.14	4 uker	25 min.	17	0	15
09	.53	.06	4 uker	20 min.	23	3	10
10	.48	.11	4 uker	20 min.	21	0	18



Når det gjelder inferenstreningens intensitet og varighet, var det ikke hensiktsmessig å korrelere denne variabelen med endringsskårene fordi åtte av de ti klassene i eksperimentgruppen gjennomførte inferenstreningen i løpet av tre til fire uker. Derimot var det større variasjon i klassenes gjennomsnittlige tidsbruk per leksjon. To lærere hadde ikke rapportert tidsbruk per leksjon, derfor ble kun åtte klasser inkludert i denne analysen. Resultatene fra korrelasjonsanalysene viste en signifikant positiv korrelasjon mellom tidsbruk per leksjon og endringsskårer i leseforståelse,  $.76, p=.029$ , hvilket innebærer at høy endringsskåre assosieres med høyt antall minutter brukt på inferenstreningen per leksjon. Det var ingen signifikant korrelasjon mellom tidsbruk per leksjon og endringsskårer i inferens.

Spearman's rho ble også brukt til å undersøke sammenhenger mellom klassenes gjennomsnittlige endringsskårer og klassenes gjennomsnittlige skårer på kontrollvariablene ordforråd, ordavkoding, nonverbale ferdigheter og motivasjon ved pretest. Analysene viste en signifikant positiv korrelasjon mellom log odds endringsskårer i leseforståelse og ordavkoding,  $.66, p=.037$ , hvilket innebærer at høy skåre på ordavkoding assosieres med høy endringsskåre i leseforståelse. Det var ingen signifikante korrelasjoner mellom endringsskårer i inferens eller leseforståelse og øvrige kontrollvariabler.

#### **7.4.4 I hvilken grad forklarer kontrollvariablene endring?**

##### **7.4.4.1 Effekt av inferenstrening kontrollert for ordforråd, ordavkoding og nonverbale ferdigheter**

Analysene av effekt som er beskrevet tidligere, viste at eksperimentgruppen hadde større framgang på effektmålene enn kontrollgruppen. Det var nødvendig å undersøke om eksperimentgruppens framgang kunne forklares av forskjeller i kontrollvariablene ordavkoding, ordforråd og nonverbale ferdigheter. Grunnene til dette er for det første at eksperimentgruppen og kontrollgruppen ikke var matchede med hensyn til ordavkoding, målt med ordkjedetesten. For det andre ble elevene ikke testet med verken ordavkoding, ordforråd eller nonverbale ferdigheter ved posttest.

Enveis ANCOVA mellom grupper ble brukt til å undersøke hvorvidt det var effekt av inferenstreningen på inferens og leseforståelse, kontrollert for ordavkoding, ordforråd og nonverbale ferdigheter. I analysene ble gruppe (eksperimentgruppe og kontrollgruppe) brukt som uavhengig variabel. Avhengig variabel var log odds transformerte skåre på inferens/leseforståelse ved posttest. Deltakernes log odds transformerte skårer på inferens/

leseforståelse på pretest, ordavkoding, ordforråd og nonverbale ferdigheter ble brukt som kovariater. Korrelasjonene mellom inferensskårer ved pretest og ordavkoding, ordforråd og nonverbale ferdigheter var henholdsvis .40, .41, og .41, mens korrelasjonene mellom pretestskårer i leseforståelse og ordavkoding, ordforråd og nonverbale ferdigheter var henholdsvis .39, .41 og .42.

Forutsetninger for bruk av ANCOVA ble sjekket. Det var ingen brudd på forutsetningene med hensyn til normalfordeling, linearitet, homogenitet av varians og homogenitet av regresjonslinjer.

Innledende analyser til ANCOVA for inferens viste at korrelasjonene mellom avhengig variabel (inferensskårer ved posttest) og kovariatene ordavkoding, ordforråd, nonverbale ferdigheter og inferens ved pretest var henholdsvis .34, .50, .42 og .60.

Tabell 31: Ikke-justerte og justerte gjennomsnitt for inferens ved posttest

	Ikke-justert			Justert	
	N	M	SD	M	SE
Eksperimentgruppe	130	0.00	0.32	0.00	0.02
Kontrollgruppe	75	-0.28	0.30	-0.27	0.03

Tabell 34 viser at forskjellen mellom gruppenes gjennomsnittlige log odds inferensskårer ved posttest ble noe mindre etter at det var justert for deltakernes skårer på ordavkoding, ordforråd og inferens på pretest.

Tabell 32: Enveis ANCOVA for inferens målt ved posttest kontrollert for inferens ved pretest, ordforråd og ordavkoding

Kilde til varians	df	M Kvadrert	F	P	Partiell $\eta^2$
Inferens (pretest)	1	3.43	69.35	<.001	.26
Ordavkoding	1	0.04	0.77	=.771	.00
Ordforråd	1	0.91	18.47	<.001	.09
Nonverbale ferdigheter	1	0.41	8.34	=.004	.04
Gruppe	1	3.43	69.35	<.001	.26
Error	199				

Tabell 32 viser at det var en signifikant forskjell mellom gruppene med hensyn til inferens på posttest også etter at det er kontrollert for ordavkoding, ordforråd, nonverbale ferdigheter og inferens på pretest. Effekten av gruppe kan karakteriseres som sterk, hvilket betyr at hvorvidt deltakeren hører til eksperiment eller kontrollgruppe hadde stor innflytelse på resultatet på inferens ved posttest. Pretestskårer på inferens hadde en sterk innflytelse på

skårer på inferens ved posttest kontrollert for øvrige kovariater. Ordforråd hadde en middels sterk innflytelse på skårer på posttestskårer i inferens, mens nonverbale ferdigheter hadde liten til middels effekt på posttestskårer i inferens. Det var ingen signifikant effekt av ordavkodning.

Innledende analyser til ANCOVA for leseforståelse viste at korrelasjonene mellom avhengig variabel (leseforståelse ved posttest) og kovariatene ordavkodning, ordforråd, nonverbale ferdigheter og pretestskårer i leseforståelse var henholdsvis .48, .46, .42 og .61.

Tabell 33: Ikke-justerte og justerte gjennomsnitt for leseforståelse ved posttest

	Ikke-justert			Justert	
	N	M	SD	M	SE
Eksperimentgruppe	134	0.46	0.33	0.44	0.02
Kontrollgruppe	76	0.23	.0.25	0.27	0.03

Av tabell 33 går det fram at forskjellene mellom gruppene med hensyn til leseforståelse ble noe mindre etter at det ble justert for deltakernes skårer på ordavkodning, ordforråd, nonverbale ferdigheter og leseforståelse målt ved pretest.

Tabell 34: Enveis ANCOVA for leseforståelse ved posttest, kontrollert for pretestskårer i leseforståelse, ordforråd, ordavkodning og nonverbale ferdigheter

Kilde til varians	df	M Kvadrert	F	P	Partiell $\eta^2$
Leseforståelse (pretest)	1	2.46	50.07	<.001	.20
Ordavkodning	1	0.48	9.77	=.002	.05
Ordforråd	1	0.36	7.32	=.007	.04
Nonverbale ferdigheter	1	0.31	6.32	=.013	.03
Gruppe	1	1.40	28.42	<.001	.12
Error	204				

Av tabell 34 går det fram at det også var en signifikant forskjell mellom gruppene med hensyn til leseforståelse på posttest. Eksperimentgruppen hadde en signifikant framgang sammenlignet med kontrollgruppen etter at det er kontrollert for skåre på leseforståelse ved pretest, samt ordavkodning, nonverbale ferdigheter og ordforråd. Effekten for gruppe kan karakteriseres som sterk, med andre ord var eksperimentgruppens framgang betydelig sammenlignet med kontrollgruppen. Pretestskårer på leseforståelse hadde en sterk innflytelse på skårer på posttestskårer i leseforståelse. Ordforråd og ordavkodning hadde en middels sterk innflytelse på leseforståelseskårer på posttest. Det var også en signifikant effekt av nonverbale ferdigheter, men denne effekten var svak.

#### 7.4.4.2 Sammenhenger mellom endring og ordforråd, ordavkoding og nonverbale ferdigheter

For å undersøke sammenhengen mellom endring på effektmålene og kontrollvariablene ordforråd, ordavkoding og nonverbale ferdigheter, ble korrelasjonsanalyser gjennomført. Dette for å finne svar på om kontrollvariablene kunne forklare variasjonen i individuell framgang på effektmålene, jevnfør problemstilling 11. Siden det var eksperimentgruppen som hadde størst framgang, ble kun eksperimentelever inkludert i analysene.

Forholdet mellom log odds endringsskårer fra pretest til posttest i inferens og leseforståelse og kontrollvariablene ved pretest ble undersøkt ved å beregne Pearson's r. Resultatene er oppsummert i tabell 35.

Tabell 35: Korrelasjoner mellom log odds endringsskårer i inferens og leseforståelse og kontrollvariabler

	Ordforråd	Ordavkoding	Nonverbale ferdigheter	Motivasjon
Inferens	.02	.12*	.07	.02
Leseforståelse	.02	.08	-.03	.08

\* Korrelasjonen var signifikant på  $p < .05$

Det var en signifikant korrelasjon mellom ordavkoding og endringsskåre i inferens, men denne korrelasjonen var svært lav. Det var ingen øvrige signifikante korrelasjoner mellom kontrollvariablene og endringsskårer i inferens og leseforståelse.

#### 7.4.5 Endring hos ulike elevgrupper

I dette avsnittet gjøres det rede for resultater som svarer til problemstilling 10. For det første om det var forskjell i endring i inferens og leseforståelse hos elever med sterk, gjennomsnittlig og svak leseforståelse. For det andre om det var forskjell i endring i inferens og leseforståelse hos elever med gode, gjennomsnittlige og svake ferdigheter i inferens. Videre undersøkes det om det var forskjell i endring hos gutter og jenter, og mellom enspråklige og tospråklige elever.

##### 7.4.5.1 Forskjeller mellom elever på ulike ferdighetsnivå

For å finne ut om elevenes forutsetninger med hensyn til inferens og leseforståelse hadde sammenheng på hvordan de profitterte på inferenstreningen, ble elevene i utvalget delt inn i tre nivåer: Elever med svake (nivå 1), middels (nivå 2) og gode (nivå 3) ferdigheter i inferens og leseforståelse. Nivåene ble definert ut fra gjennomsnittsskårer på pretest. Nivå

1 besto av elever som skåret ett standardavvik under gjennomsnittet, mens nivå 3 besto av elever som skåret ett standardavvik over gjennomsnittet. Nivå 2 besto av elever som hadde skårer mellom minus og pluss ett standardavvik fra gjennomsnittet. Elevene ble gruppert separat for inferens og leseforståelse. Dette innebærer at nivåene for inferens og nivåene for leseforståelse ikke besto av de samme individene. En del elever ble imidlertid kategorisert i samme nivå både på inferens og leseforståelse. Eksperimentgruppen hadde henholdsvis 28, 81, og 26 elever på nivå 1, 2 og 3 i inferens, og henholdsvis 21, 96 og 9 elever på nivå 1, 2 og 3 i leseforståelse. Kontrollgruppen hadde 13, 57 og 12 elever på henholdsvis nivå 1, 2 og 3 i inferens og 19, 55 og 9 på henholdsvis nivå 1, 2 og 3 i leseforståelse.

Det ble gjennomført to toveis ANOVAer, en for inferens og en for leseforståelse. Hensikten med disse analysene var å undersøke betydningen av gruppe og nivå i inferens og leseforståelse for endring i inferens og leseforståelse. med andre ord om det var interaksjon mellom gruppe (eksperiment og kontroll) og nivå.

Toveis ANOVA for endring i inferens viste at det ikke var signifikant interaksjonseffekt av variablene nivå i inferens og gruppe,  $F(2,211)=0.905$ ,  $p=.406$ . Det var heller ikke signifikant interaksjonseffekt av variablene nivå i leseforståelse og gruppe i ANOVA for endring i leseforståelse,  $F(2,218)=0.092$ ,  $p=.912$ . Dette resultatet innebærer at det ikke var forskjell i hvordan elever med svake, middels og gode ferdigheter i inferens og leseforståelse, slik det er definert her, profitterte på inferenstreningen.

#### **7.4.5.2 Forskjeller mellom gutter og jenter**

For å undersøke om det var forskjell i hvordan gutter og jenter profitterte på inferenstreningen, ble gutter i eksperimentgruppen sammenlignet med gutter i kontrollgruppen, og jenter i eksperimentgruppen med jenter i kontrollgruppen. For å undersøke om kjønnsforskjellene ble mindre etter inferenstreningen, ble endringsskårene til gutter og jenter i eksperimentgruppen sammenlignet.

Forskjeller mellom gutter og jenter i de to gruppene ble sammenlignet ved hjelp av t-test for uavhengige grupper. I tabell 36 og 37 sammenlignes gutter og jenter i eksperimentgruppe og kontrollgruppe.

Tabell 36: Sammenligning av log odds endringsskårer fra pretest til posttest hos gutter i eksperimentgruppe og kontrollgruppe

Mål	Eksperiment- gruppe			Kontrollgruppe			df	t	$\eta^2$	P
	N	M	SD	N	M	SD				
Inferens, totalt	68	0.48	0.38	41	0.12	0.29	107	3.15	.20	<.001
Inferens, åpne	65	0.48	0.35	39	0.12	0.35	102	5.09	.20	<.001
Inferens, understrekning	63	0.38	0.34	39	0.07	0.32	100	4.59	.17	<.001
Leseforståelse, totalt	73	0.15	0.30	41	-0.05	0.23	112	3.71	.11	<.001
Leseforståelse, Fakta	65	0.26	0.38	36	0.07	0.28	91.5 2	2.88	.08	=.010
Leseforståelse, tolkning	73	0.04	0.34	41	-0.18	0.37	112	3.15	.08	=.002
Leseforståelse, fiksjon	66	0.10	0.44	40	-0.12	0.31	104	2.63	.06	=.010
Leseforståelse, nonfiksjon	73	0.19	0.30	40	0.00	0.31	111	3.22	.09	=.002

Tabell 37: Sammenligning av log odds endringsskårer fra pretest til posttest hos jenter i eksperimentgruppe og kontrollgruppe

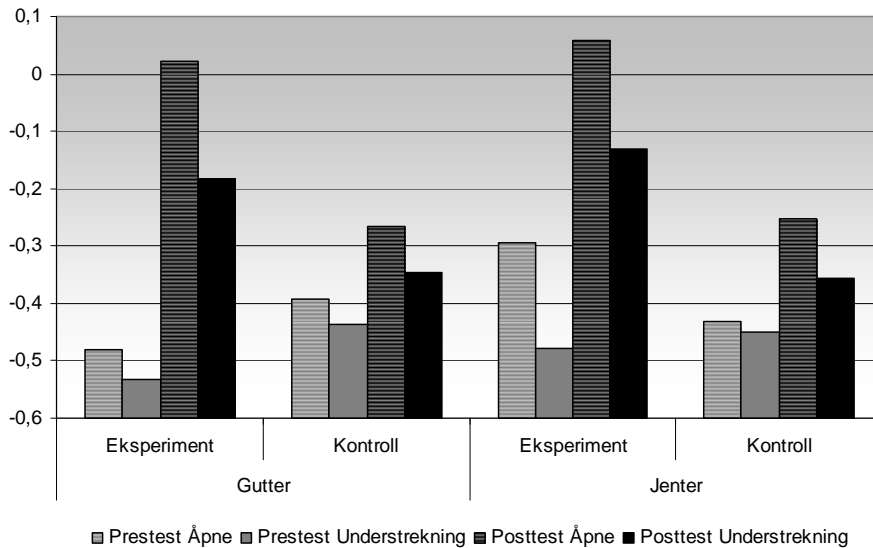
Mål	Eksperiment- gruppe			Kontrollgruppe			df	t	$\eta^2$	P
	N	M	SD	N	M	SD				
Inferens	67	0.39	0.28	41	0.17	0.30	106	4.64	.12	<.001
Inferens, åpne	66	0.38	0.32	41	0.21	0.31	105	2.75	.07	=.007
Inferens, understrekning	64	0.39	0.35	39	0.11	0.43	101	3.60	.11	<.001
Leseforståelse	68	0.15	0.32	42	-0.03	0.24	108	3.16	.08	=.002
Leseforståelse, Fakta	62	0.22	0.40	40	0.15	0.41	100	0.88	.01	=.381
Leseforståelse, tolkning	67	0.06	0.36	42	.20	0.23	106. 94	4.64	.17	<.001
Leseforståelse, fiksjon	58	-0.04	0.46	41	-0.19	0.35	97	1.78	.03	=.091
Leseforståelse, nonfiksjon	68	0.29	.35	42	0.10	0.35	108	2.76	.07	=.007

Av tabell 36 og 37 går det fram at var det var signifikant forskjell mellom gutter i eksperimentgruppen og kontrollgruppen i forhold til endringsskåre i inferens, totalskåre i leseforståelse, for begge typer spørsmål i leseforståelse og begge teksttyper. Forskjellen mellom gutter i de to gruppene var størst med hensyn til endringsskåre i inferens. Det var også signifikant forskjell mellom

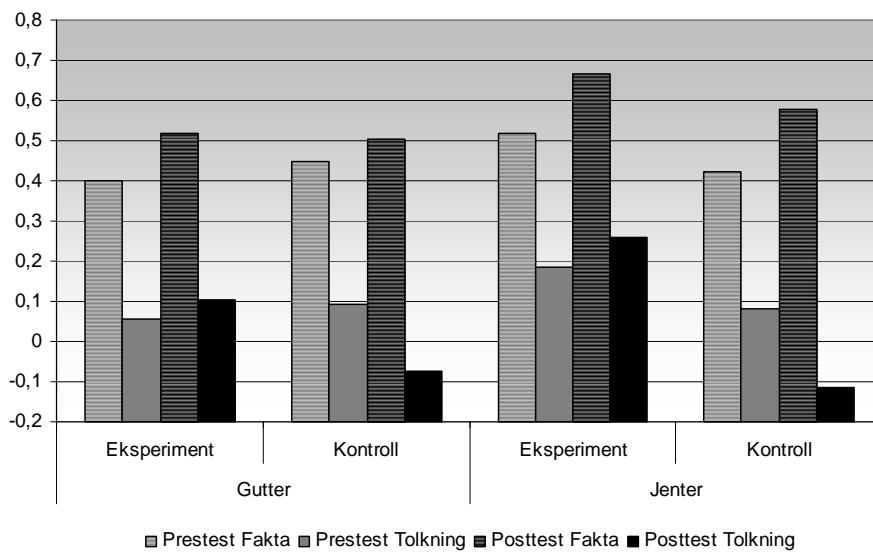
jenter i eksperimentgruppen og kontrollgruppen med hensyn til både inferens og totalsum på leseforståelse. Det var imidlertid ikke signifikant forskjell mellom jenter i de to gruppene med hensyn til faktaspørsmål eller fiksjonstekster. Forskjellen mellom eksperimentjenter og kontrolljenter var signifikant i forhold til tolkningsspørsmål og nonfiksjonstekster. Forskjellen mellom jenter i de to gruppene var størst med hensyn til tolkningsspørsmål.

Umiddelbart kan resultatene i tabellene over kan gi inntrykk av at inferenstreningen hadde større effekt blant gutter enn blant jenter. For å undersøke dette ble ANOVAer med gjentakende målinger gjennomført for både inferens og leseforståelse, hver med de tre uavhengige variablene: kjønn x gruppe x tid. I analysene ble både kjønn og gruppe lagt inn som between-subjects variabler, mens tid ble lagt inn som within-subjects variabel. ANOVAen for inferens viste at treveisinteraksjonen ikke var signifikant,  $F(1,213)=2.748$ ,  $p=.099$ ,  $\eta^2<.001$ . ANOVAen for leseforståelse viste heller ikke noen signifikant treveisinteraksjon,  $F(1,220)=0.054$ ,  $p=.816$ ,  $\eta^2<.001$ . Dette betyr at det ikke var forskjell i hvordan gutter og jenter profitterte på inferenstreningen, verken vurdert med inferensprøven eller prøven i leseforståelse.

Det er sannsynlig at treningen kunne hatt forskjellige effekter blant gutter og jenter. For eksempel kunne det være at guttene fikk forholdsvis mer ut av treningen ved faktaspørsmål, sammenlignet med tolkningsspørsmål og lesing av fiksjonstekster (sammenlignet med nonfiksjonstekster), enn jentene. For å undersøke dette ble kjønnsforskjeller undersøkt ytterligere med ANOVAer med gjentakende målinger for både inferens og leseforståelse, hver med følgende fire uavhengige variabler: Kjønn x gruppe x tid x spørsmålstype/teksttype. Kjønn og gruppe ble lagt inn som between-subjects variabler i analysen, mens tid og spørsmålstype/teksttype ble lagt inn som within-subjects variabler. ANOVAen for inferens viste at interaksjonen kjønn x gruppe x tid x spørsmålstype (åpne spørsmål og understreknings spørsmål) ikke var signifikant,  $F(1,195)=3.779$ ,  $p=.053$ ,  $\eta^2=.002$ . ANOVAen for leseforståelse viste at interaksjonen kjønn x gruppe x tid x spørsmålstype (faktaspørsmål og tolkningsspørsmål) ikke var signifikant,  $F(1,199)=2.530$ ,  $p=.113$ ,  $\eta^2=.001$ . Interaksjonen kjønn x gruppe x tid x teksttype (fiksjonstekster og nonfiksjonstekster) var heller ikke signifikant,  $F(1,201)=.397$ ,  $p=.529$ ,  $\eta^2<.001$ .

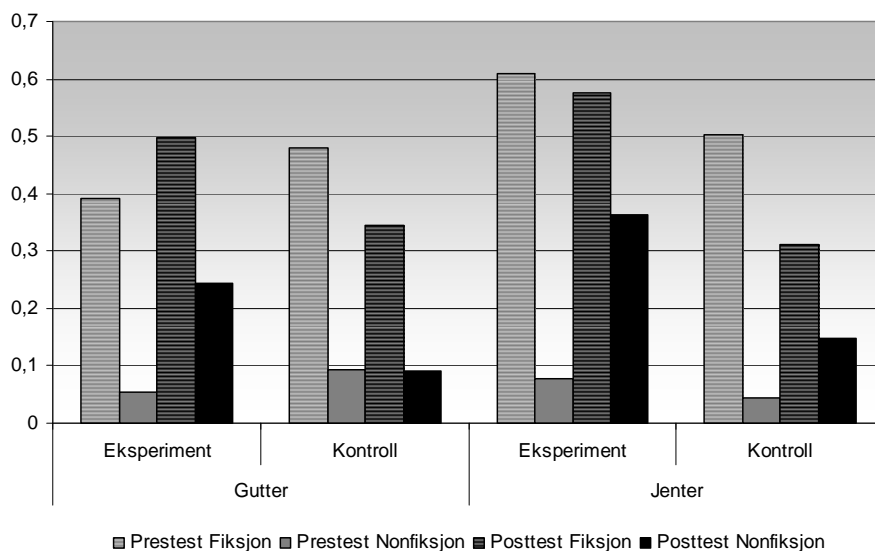


Figur 9. Log oddsskårer på åpne spørsmål og understrekingsspørsmål på inferensprøven hos gutter og jenter i eksperimentgruppe og kontrollgruppe



Figur 10. Log oddsskårer på faktaspørsmål og tolkningsspørsmål på prøven i leseforståelse hos gutter og jenter i eksperimentgruppe og kontrollgruppe





Figur 11. Log oddsskårer på fiksjonstekster og nonfiksjonstekster på prøven i leseforståelse hos gutter og jenter i eksperimentgruppe og kontrollgruppe

Oppsummert hadde ikke inferenstreningen forskjellig effekt på gutter og jenter, slik effekten ble målt med åpne spørsmål og understreknings spørsmål på inferensprøven og faktaspørsmål, tolkningsspørsmål, fiksjonstekster og nonfiksjonstekster på prøven i leseforståelse. Figur 9 til 11 viser log oddsskårer på spørsmålstype og testtype ved pre- og posttest i inferens og leseforståelse hos jenter og gutter i eksperimentgruppe og kontrollgruppe.

#### 7.4.5.3 Forskjeller mellom enspråklige og tospråklige elever

For å undersøke om det var forskjeller i hvordan enspråklige og tospråklige profitterte på inferenstreningen, ble enspråklige og tospråklige i kontroll- og eksperimentgruppe sammenlignet. Videre ble også enspråklige og tospråklige i eksperimentgruppen sammenlignet for å undersøke om inferenstreningen bidro til større eller mindre forskjeller mellom enspråklige og tospråklige.

Forskjeller i log odds endringsskårer fra pretest til posttest i inferens og leseforståelse mellom tospråklige og enspråklige i eksperimentgruppen, og sammenligning av tospråklige i eksperimentgruppe og kontrollgruppe ble analysert ved Mann Whitney U test. Cohen's  $d$  ble beregnet for sammenligning størrelsen på forskjeller mellom tospråklige i eksperimentgruppen og kontrollgruppen. T-test for uavhengige grupper ble benyttet for å undersøke

forskjeller mellom enspråklige i eksperimentgruppen og kontrollgruppen. Kvadrert eta ble beregnet på bakgrunn av t-verdier.

Tabell 38: Sammenligning av log odds endringsskårer fra pretest til posttest hos enspråklige i eksperimentgruppen og kontrollgruppen

Mål	Eksperiment- gruppe			Kontrollgruppe			df	t	$\eta^2$	P
	N	M	SD	N	M	SD				
Inferens, totalt	123	0.43	0.34	75	0.16	0.30	196	5.69	.14	<.001
Inferens, åpne	119	0.42	0.32	73	0.18	0.33	190	5.04	.12	<.001
Inferens, understreking	116	0.37	0.34	71	0.09	0.38	185	5.16	.13	<.001
Leseforståelse, totalt	128	0.15	0.30	76	-0.04	0.22	202	4.24	.08	<.001
Leseforståelse, Fakta	116	0.24	0.40	69	0.12	0.36	183	1.98	.02	=.049
Leseforståelse, tolkning	127	0.05	0.35	76	-0.20	0.25	194. 49	5.84	.15	<.001
Leseforståelse, fiksjon	112	0.01	0.46	74	-0.16	0.35	180. 30	2.84	.04	=.008
Leseforståelse, nonfiksjon	128	0.25	0.33	75	0.05	0.31	201	4.22	.08	<.001

Tabell 39: Sammenligning av log odds endringsskårer fra pretest til posttest hos tospråklige i eksperimentgruppen og kontrollgruppen

Mål	Eksperimentgruppe			Kontrollgruppe			U	d	P
	N	M	SD	N	M	SD			
Inferens, totalt	12	0.43	0.34	7	0.05	0.24	10.00	1.63	=.007
Inferens, åpne	12	0.56	0.32	7	0.08	0.33	13.00	1.48	=.014
Inferens, understreking	11	0.48	0.33	7	0.03	0.30	11.50	1.43	=.014
Leseforståelse, totalt	13	0.15	0.34	7	-0.01	0.41	37.50	0.43	=.526
Leseforståelse, fakta	11	0.32	0.33	7	0.05	0.32	26.50	0.75	=.276
Leseforståelse, tolkning	13	0.08	0.38	7	-0.08	0.69	36.00	0.29	=.611
Leseforståelse, fiksjon	12	0.24	0.37	7	-0.13	0.55	22.00	0.79	=.090
Leseforståelse, nonfiksjon	13	0.16	0.52	7	0.07	0.20	38.00	0.21	=.550

Det var signifikant forskjell mellom tospråklige i eksperimentgruppe og kontrollgruppe med hensyn til endringsskårer i inferens, denne forskjellen var svært sterk. Det var imidlertid ikke signifikant forskjell mellom tospråklige i eksperimentgruppe og kontrollgruppe i forhold til leseforståelse.

T-tester viste at det var signifikante forskjeller mellom enspråklige i eksperimentgruppe og kontrollgruppe på samtlige mål på effekt, både totalskårer og delskårer. Effekten var sterkest for prøven i inferens og tolkningsspørsmål på prøven i leseforståelse. Effekten for faktaspørsmål var liten og effekten for fiksjonstekster var liten til middels sterk.

Mann Whitney U test viste at det ikke var signifikant forskjell i log odds endringsskåre i inferens mellom tospråklige (N=12, gjennomsnittlig rangering=81.83) og enspråklige (N=123, gjennomsnittlig rangering=66.65) i eksperimentgruppen,  $U=572.00$ ,  $p=.199$ . Det var heller ingen signifikant forskjell da åpne spørsmål og understrekningsoppgaver ble analysert hver for seg. Heller ikke i forhold til log odds endringsskåre i leseforståelse var det signifikant forskjell mellom tospråklige (N=13, gjennomsnittlig rangering=75.27) og enspråklige (N=128, gjennomsnittlig rangering=70.57) i eksperimentgruppen,  $U=776.50$ ,  $p=.692$ , det var heller ingen signifikant forskjell mellom enspråklige og tospråklige da spørsmålstyper og teksttyper ble analysert hver for seg.

På grunn av lavt antall tospråklige elever ble ikke interaksjonseffekter for enspråklige og tospråklige i eksperimentgruppe og kontrollgruppe beregnet. Med andre ord kan det ikke slås fast om enspråklige og tospråklige profitterte likt på inferenstreningen.

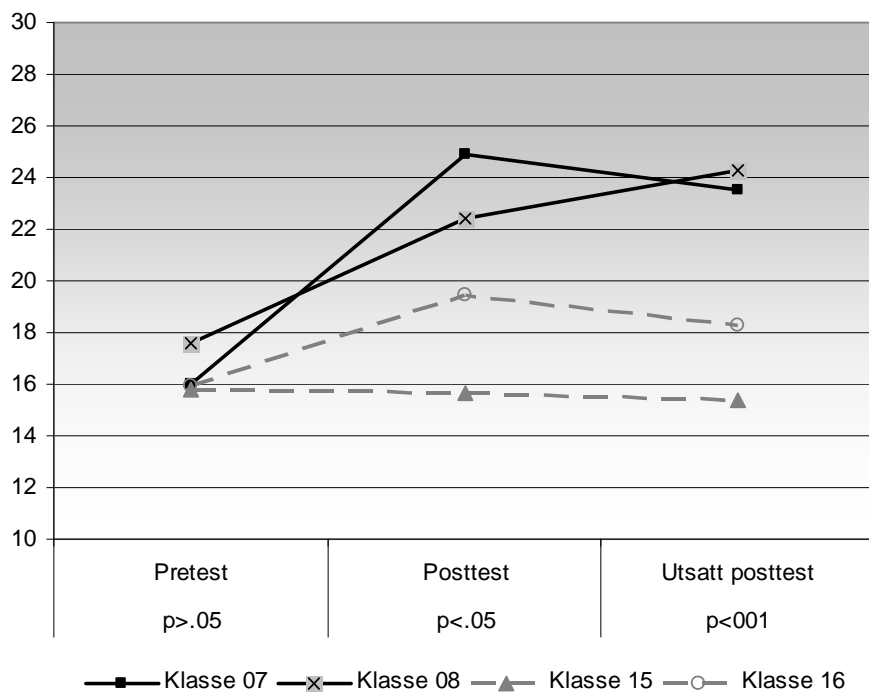
#### **7.4.6 Langtidseffekter av inferenstreningen?**

For å undersøke effektene av inferenstreningen på sikt, om effektene vedvarte og om effektene ble sterkere eller svakere etter avsluttet inferenstrening, ble to eksperimentklasser med til sammen 29 elever og to kontrollklasser med til sammen 27 elever testet med utsatt posttest i inferens og leseforståelse fem uker etter at eksperimentklassene hadde avsluttet inferenstreningen. Mellom posttest og utsatt posttest fikk alle elevene ordinær undervisning. I dette avsnittet fokuseres det på disse fire klassenes utvikling på effektvariablene.

T-test for uavhengige grupper viste at det ikke var signifikant forskjell mellom den reduserte eksperimentgruppen og kontrollgruppen med hensyn til kontrollvariablene ordforråd, ordavkodning, nonverbale ferdigheter, matematikk og motivasjon målt ved pretest. Det var heller ikke signifikante

forskjeller mellom de reduserte gruppene med hensyn til inferens eller leseforståelse ved pretest.

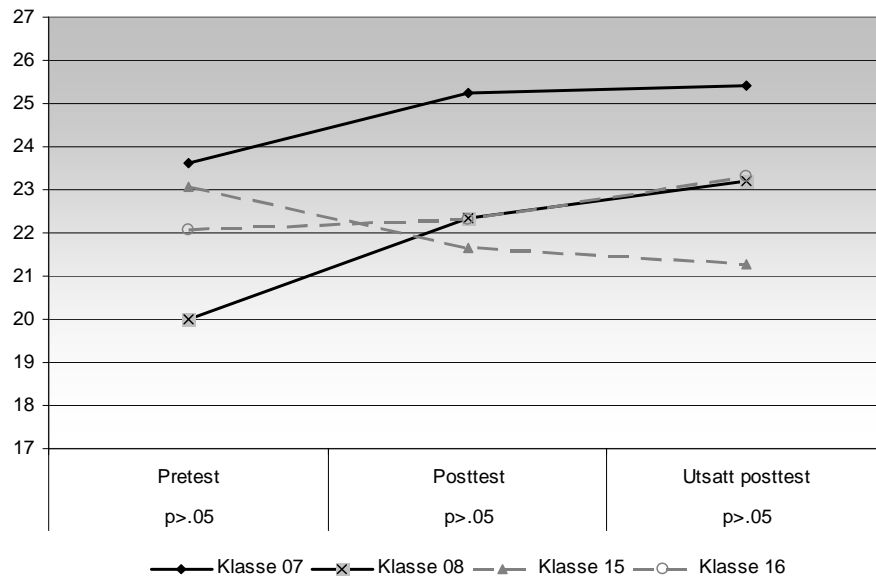
Figur 12 viser de fire klassenes utvikling i råskårer på de tre måletidspunktene for inferens. Klasse 07 og 08 var eksperimentklasser, mens klasse 15 og 16 var kontrollklasser. Tre av klassene gikk fram fra pretest til posttest, mens kontrollklasse 15 gikk noe tilbake fra pretest til posttest. Begge kontrollklasser hadde tilbakegang fra posttest til utsatt posttest. Eksperimentklasse 08 fortsatte framgangen fra posttest til utsatt posttest, mens eksperimentklasse 07 hadde tilbakegang.



Figur 12. Utviklingen i råskårer i inferens til de fire klassene som deltok på alle tre måletidspunkter

Figur 13 viser de fire klassenes utvikling på de tre måletidspunktene for leseforståelse. Alle klassene, med unntak av kontrollklasse 15, gikk fram fra pretest til posttest og fra posttest til utsatt posttest. Eksperimentklasse 08 og kontrollklasse 16 gikk om lag like mye fram fra posttest til utsatt posttest. Merk her at to prøvesett ble brukt til å måle leseforståelse på de tre tidspunktene. Prøvesett A ble brukt på pretest og utsatt posttest, mens prøvesett B ble brukt på posttesten. Som det tidligere er gjort rede for var det

en signifikant forskjell i skårer på prøven i leseforståelse på posttest mellom hele eksperimentgruppen og hele kontrollgruppen. Forskjellen skåre på prøven i leseforståelse ved posttest mellom de reduserte gruppene var imidlertid ikke signifikant.



Figur 13. Utviklingen i råskårer i leseforståelse til de fire klassene som deltok på alle tre måletidspunkter

Forskjell mellom eksperimentgruppens og kontrollgruppens endring i målene på effekt over de tre måletidspunktene ble undersøkt ved å kombinere between-subjects faktorer (gruppe: Eksperimentgruppe og kontrollgruppe) og within-subject faktorer (tid: Inferens/leseforståelse målt ved pre-, post- og utsatt posttest) i en enveis ANOVA for gjentakende målinger.

ANOVA for inferens på de tre måletidspunktene viste en signifikant interaksjon mellom gruppe og tid,  $F(2,102)=4.36$ ,  $p=.015$ ,  $\eta^2=.05$ . Analyse av kontraster i den avhengige variabelen viste at det kun var signifikant interaksjon mellom gruppe og inferens ved pretest og utsatt posttest,  $F(1,51)=0.02$ ,  $p=.005$ ,  $\eta^2=.11$ , mens interaksjonen mellom gruppe og inferens ved posttest og utsatt posttest ikke var signifikant,  $F(1,51)=7.70$ ,  $p=.966$ ,  $\eta^2<.01$ . Den ikke-signifikante interaksjonen mellom gruppe og inferens på posttest og utsatt posttest betyr at eksperimentelevene ikke hadde ytterligere framgang etter avsluttet inferenstrening. Interaksjon mellom gruppe og

inferens på pretest og utsatt posttest innebærer at effekten av inferenstreningen vedvarte.

ANOVA for leseforståelse viste at det ikke var signifikant interaksjon mellom gruppe og tid,  $F(2,102)=2.73$ ,  $p=.070$ ,  $\eta^2=.03$ . Siden tidligere analyser viste at det var forholdsvis stor effekt for tolkningsspørsmål mens effekten for faktaspørsmål var beskjeden, ble langtidseffekter av de to spørsmålstypene undersøkt. ANOVA for tolkningsspørsmål viste at det var en ikke-signifikant gruppe x tid interaksjon,  $F(2,100)=3.00$ ,  $p=.055$ ,  $\eta^2=.02$ . Dette innebærer at det ikke var forskjell i gruppenes endring over tid. Siden interaksjonen var like over det konvensjonelle signifikansnivået, ble analyse av kontraster i den avhengige variabelen likevel gjennomført. Resultatene viste at det var signifikant interaksjon mellom gruppe og leseforståelse ved pretest og utsatt posttest,  $F(1,50)=4.21$ ,  $p=.045$ ,  $\eta^2=.07$ , mens interaksjonen mellom gruppe og leseforståelse ved posttest og utsatt posttest ikke var signifikant,  $F(1,50)=0.37$ ,  $p=.545$ ,  $\eta^2<.01$ . Altså er det samme tendens for tolkningsspørsmål som for inferens, effekten av inferensprøven vedvarte, men det var ingen ytterligere framgang for eksperimentgruppen. Det ble ikke gjennomført tilsvarende analyse for faktaspørsmålene fordi variablene ikke tilfredsstilte forutsetningen om homogenitet i kovariansmatrisen. Box's test av likhet i kovariansmatriser bekreftet at den observerte kovariansen for den avhengige variabelen (tid) var lik på tvers av gruppene. Med andre ord var det lite variasjon mellom eksperimentgruppe og kontrollgruppe på den uavhengige variabelen.

## **7.5 Diskusjon**

I hvilken grad resultatene besvarer forskningsspørsmålene som ble presentert i innledningen i dette kapittelet, blir diskutert i dette avsnittet. Videre vil muligheter for generalisering også bli diskutert, før det konkluderes i forhold til målene for studien.

### **7.5.1 Resultater og problemstillinger**

#### **7.5.1.1 Effekter av inferenstreningen på inferens og leseforståelse**

Resultatene fra analyser av interaksjon mellom tid og gruppe viste at eksperimentgruppen hadde en klart større framgang i inferens enn kontrollgruppen. Inferenstreningen hadde med andre ord effekt på elevenes evne til å gjøre inferenser. Siden inferensprøvens innhold og struktur lå tett opp mot det som elevene hadde trent på i de åtte leksjonene, er en sterk effekt på inferensprøven tegn på at inferenstreningen var implementert. En slik primær overføringseffekt er følgelig en viktig forutsetning for å tolke generaliserte effekter.

Inferenstreningen hadde også en generalisert effekt på leseforståelse. Teoretisk er dette en bekreftelse på årsaksforholdet mellom inferens og leseforståelse, at leseforståelse er en følge av ferdigheter i inferens. De praktisk-pedagogiske konsekvensene av effektene av inferenstreningen innebærer at man kan øve elever i inferens for å bedre deres leseforståelse. Studien har også skissert hvordan inferens kan trenes via grafiske modeller. Sterk effekt av inferenstrening på elevenes evne til å gjøre inferenser, er i seg selv verdifullt, men det som særlig er positivt, er at inferenstreningen hadde generalisert effekt på et generelt mål på leseforståelse.

Både eksperimentgruppen og kontrollgruppen hadde signifikant høyere gjennomsnittskåre på inferens på posttesten enn på pretesten. Dette kan skyldes at samme prøve ble administrert på begge måletidspunkter. På posttesten kjente allerede elevene i begge grupper både tekstene og spørsmålene, hvilket kan ha ført til at elevene brukte mindre tid på oppgavene og derfor fikk løst flere oppgaver, og i neste omgang har fått flere poeng på prøven. Oppgavene i seg selv kan ha bidratt til en viss læringseffekt, i og med at tekstene og spørsmålene var konstruert på samme måte, og at elevene dermed ble flinkere til å løse inferensspørsmålene fordi de gjennom prøven fikk trening i å svare på inferensspørsmål. En slik læringseffekt vil imidlertid ha vært den samme for både eksperimentelever og kontrollelever.

Prøven i leseforståelse besto, i motsetning til inferensprøven, ikke av de samme tekstene og oppgavene ved pretest og posttest. Da de to gruppens gjennomsnittsskåre på pretest og posttest ble sammenlignet, viste det seg at eksperimentgruppens skåre på leseforståelse ved posttest var signifikant høyere enn ved pretest. Kontrollgruppen hadde et noe lavere gjennomsnitt, dog ikke signifikant lavere, ved posttest enn ved pretest. Dette kan skyldes at prøven i leseforståelse ikke var den samme på de to tidspunktene. Derfor er det også sannsynlig at prøvesettet som ble benyttet til å måle leseforståelse ved posttesten var noe vanskeligere enn prøvesettet som ble brukt ved pretesten.

Effekten på leseforståelse var noe mindre enn på inferens. Det er naturlig at inferenstreningen hadde større effekt på den prøven som er nærmest knyttet til innholdet i treningen enn på prøven som krever at deltakerne i større grad overfører kunnskapen fra treningen. Tekstene som ble brukt i inferenstreningen kan karakteriseres som nonfiksjonstekster. Siden resultatene viste at det var en middels sterk effekt for både fiksjonstekster og nonfiksjonstekster, kan dette tyde at det var overføringseffekt fra nonfiksjonstekster til

fiksjonstekster. Med andre ord kan man ikke spore begrensinger i generalisering av effekten av leseforståelse i sjanger med det målet som er brukt.

Analyser av effekt for leseforståelse fra pretest til posttest viste at det var forskjell i effekt for type spørsmål (faktaspørsmål og tolkningsspørsmål). Mens effekten for faktaspørsmålene kun var liten, var effekten for tolkningsspørsmålene middels til sterk. Faktaspørsmålene krevde kun at elevene fant svaret på spørsmålene direkte i teksten, med andre ord å trekke ut informasjon. Tolkningsspørsmålene derimot, krevde i større grad refleksjon og rekonstruksjon av mening på bakgrunn av teksten. Sistnevnte spørsmål representerer en dypere leseforståelse enn for faktaspørsmålene. Inferenstreningen gjorde ikke bare elevene dyktigere på å rekonstruere mening, men også på å hente ut informasjon direkte fra teksten. Elevene fikk en spesifikk trening i inferens som kan knyttes til det å konstruere mening, hvilket tolkningsspørsmålene hadde til hensikt å måle. Så hvorfor var da ikke effekten av treningen spesifikk? Med andre ord, hvorfor var det ikke kun effekt på tolkningsspørsmålene? En forklaring kan være at det blir lettere for elevene å trekke ut informasjon fra teksten fordi de har konstruert en dypere forståelse. Eller sagt på en annen måte blir det enklere å forstå delene når man har en gjennomgripende forståelse av helheten.

Analyser av korrelasjon i eksperimentgruppen viste at høy endringsskåre på den ene typen spørsmål var assosiert med høy endringsskåre på den andre typen spørsmål og omvendt. Resultatene fra analyser av interaksjonen mellom spørsmålstype og gruppe med endringsskårer beregnet på råskårer og log oddsskårer var motstridende. Selv om det er uklart hvorvidt det var en interaksjon, er det sikkert at dersom det var en interaksjon var den av svært liten betydning. Kvadrert eta for endringsskåre basert på råskåre ble beregnet til ubetydelige .02.

Effekten av inferenstreningen ble målt både med hensyn til avstand i tid og i innholdsmessig avstand til treningen. Ikke overraskende var effektene av inferenstreningen sterkest for inferensprøven som var det effektmålet som innholdsmessig lå nærmest inferenstreningen. Inferensprøven var designet spesielt til denne studien og dermed eksperimentdesignet, mens prøven i leseforståelse ikke var designet spesielt for denne studien. Som det tidligere er gjort rede for, fant Rosenshine og Meister (1994) i en oppsummering av effektstudier av resiprok undervisning betraktelig høyere effektstørrelser for mål på effekt som var eksperimentdesignet enn for ikke-eksperimentdesignede, standardiserte prøver, jevnfør avsnitt 4.1. Denne



studien viser samme tendens, selv om forskjellene mellom effektstørrelser målt med eksperimentdesignede tester og ikke-eksperimentdesignede tester var betraktelig mindre enn det Rosenshine og Meister fant.

#### **7.5.1.2 Langtidseffekter**

To eksperimentklasser og to kontrollklasser ble testet med inferensprøven og prøve i leseforståelse med utsatt posttest fem uker etter at inferenstreningen ble avsluttet. Inferensprøven besto av de samme tekstene og oppgavene som ved pre- og posttest, mens prøven i leseforståelse besto av de samme tekstene og oppgavene som ved pretest.

På utsatt test posttest var det fortsatt signifikant forskjell mellom eksperimentgruppen og kontrollgruppen med hensyn til inferens. Interaksjonen mellom gruppe og tid var signifikant. Det var imidlertid ingen signifikant forskjell i endringen fra posttest til utsatt posttest mellom de to gruppene. Derimot var det en signifikant forskjell mellom gruppene i endringen fra pretest til utsatt posttest i eksperimentgruppens favør. Det vil si at effekten ble opprettholdt etter avsluttet inferenstrening, men det var ingen ytterligere framgang. Det at effekten ble opprettholdt etter at treningen var avsluttet kan skyldes flere ting. For det første at elevene fortsatte å anvende den metoden de hadde lært gjennom de grafiske modellene etter at treningen var avsluttet. For det andre kan inferenstreningen ha påvirket lærerne, ved at de fokuserte mer på inferens og leseforståelse i sin leseundervisning.

Gjennomsnittskårene til eksperimentklassene viste at den ene klassen, klasse 08, fortsatte framgang på inferensprøven på utsatt posttest, mens den andre klassen, klasse 07, hadde en tilbakegang fra posttest til utsatt posttest. Læreren som gjennomførte inferenstreningen i klasse 07 underviste ikke klassen i perioden mellom posttest og utsatt posttest. Selv om utvalget ved utsatt posttest var lite, og en derfor må være forsiktig å trekke forhastede konklusjoner særlig på klassenivå, er det interessant å se at nettopp klasse 07 hadde tilbakegang på inferensprøven.

I motsetning til inferensprøven, var interaksjonen mellom gruppe og de tre prøvetidspunktene ikke signifikant i forhold til prøven i leseforståelse. Videre analyser av endringen på de to spørsmålstypene på de tre prøvetidspunktene viste imidlertid at det var en signifikant interaksjon mellom gruppe og tolkningsspørsmål fra pretest til utsatt posttest, men ikke fra posttest til utsatt posttest. Dette betyr at effekten av inferenstreningen ble opprettholdt i forhold til tolkningsspørsmål, men i likhet med inferensprøven var det ingen

ytterligere framgang for eksperimentgruppen etter at inferenstreningen var avsluttet.

I tolkningen av langtidseffekter, både for inferens og leseforståelse, må det tas i betraktning at langtidseffekter kun ble målt for en liten gruppe av utvalget. Det er ikke gitt at resultatene ville vært de samme dersom hele utvalget var blitt testet med utsatt posttest. Dette fordi den statistiske styrken blir mindre når antall deltakere reduseres.

### **7.5.1.3 Betingelser for effekt**

#### ***Kontrollvariabler***

Ordforråd, ordavkoding og nonverbale ferdigheter ble inkludert som kontrollvariabler i effektstudien. I dette avsnittet diskuteres det i hvilken grad kontrollvariablene kan forklare variasjon i individuell framgang i inferens og leseforståelse, jevnfør problemstilling 11. Som det går fram av tabell 35 var det ubetydelige sammenhenger mellom endringskårer i inferens og leseforståelse og kontrollvariablene. ANCOVA viste at kontrollvariablene hadde en liten til middels effekt på variasjonen i posttestskårene i inferens og leseforståelse, jevnfør avsnitt 7.4.4.1.

Slik ordforråd, ordavkoding og nonverbale ferdigheter ble målt, var effekten av inferenstreningen i liten grad betinget av kontrollvariablene. Dette innebærer at elevene i eksperimentgruppen hadde framgang uavhengig av deres ferdigheter i ordforråd, ordavkoding og nonverbale ferdigheter. Forskjellen i framgang mellom eksperimentgruppen og kontrollgruppen kunne ikke forklares ved disse variablene.

#### ***Ulike grupper elever***

I dette avsnittet vil problemstillinger som angår forskjell i framgang for grupper av elever bli drøftet, jevnfør problemstilling 10. Først diskuteres framgangen på de to effektmålene for elever med ulike ferdigheter på pretesten med hensyn til inferens og leseforståelse. Deretter sammenlignes kjønnsforskjeller, før framgang hos enspråklige versus tospråklige blir drøftet.

Det var ingen signifikant interaksjon mellom elevenes nivå i inferens og endring i inferens. Interaksjonen mellom nivå i leseforståelse og endring i leseforståelse var heller ikke signifikant. Disse resultatene indikerer at det ikke var forskjell i hvordan elever med ulike ferdigheter i inferens og leseforståelse profitterte på inferenstreningen. Materiellet som ble brukt i inferenstreningen var laget for bruk i ordinære klasser. Et betimelig spørsmål

er derfor om man i det hele tatt kunne forvente at elever med spesielt god eller svak leseforståelse skulle profittere spesielt mye på en trening som har til hensikt å favne bredt. Kendeou et al. (2007) hevder at svake lesere ikke nødvendigvis vil tjene på å trene på det de gode leserne gjør, fordi elevens kognitive kapasitet er begrenset. Elever med svak leseforståelse har ifølge Kendeou et al. behov for å bruke sine kognitive ressurser på mer grunnleggende prosesser, som for eksempel å danne seg et mål for hvorfor de leser. Motsatt kan man tenke seg at elever med svært gode ferdigheter i leseforståelse allerede gjør mer avanserte inferenser enn de det legges opp til i inferenstreningen.

Det var i utgangpunktet få kjønnsforskjeller i utvalget. Interaksjonen mellom de avhengige variablene inferens og leseforståelse og de uavhengige variablene tid x gruppe x kjønn var ikke signifikant. Med de målene på endring som inngikk i studien, ble det ikke påvist forskjell i hvordan gutter og jenter profitterte på inferenstreningen. Interaksjonen kjønn x gruppe x tid x spørsmålstype/teksttype var heller ikke signifikant. Treningen hadde heller ikke forskjellig effekt for gutter eller jenter, slik effekt ble målt. Med andre ord bidro treningen verken til å jevne ut eller øke kjønnsforskjeller.

Det var få forskjeller mellom enspråklige og tospråklige elever ved pretestene, det var kun moderate forskjeller mellom elevgruppene med hensyn til faktaspørsmål og fiksjonstekster i de enspråkliges favør. Enspråklige i eksperimentgruppen hadde signifikant større endring fra pretest til posttest enn enspråklige i kontrollgruppen både med hensyn til inferens og leseforståelse uansett spørsmålstype og teksttype. Også tospråklige elever i eksperimentgruppen hadde signifikant større endring fra pretest til posttest enn tospråklige i kontrollgruppen med hensyn til inferens. Derimot var det ikke signifikant forskjell mellom tospråklige i eksperimentgruppen og tospråklige i kontrollgruppen, verken for totalskåre i leseforståelse, faktaspørsmål, tolkningsspørsmål, fiksjonstekster eller nonfiksjonstekster. På grunn av lavt antall tospråklige deltakere ble ikke interaksjonseffekter beregnet for enspråklige og tospråklige. Det er derfor ikke mulig å fastslå om det var forskjell i hvordan enspråklige og tospråklige elever profitterte på inferenstreningen eller om effektene av treningen var forskjellig for de to gruppene. Tendensen i denne studien viser at inferenstreningen hadde en primær overført effekt for tospråklige, men ingen generalisert effekt på leseforståelse. Denne tendensen ville vært interessant å følge opp i studier med betraktelig høyere antall tospråklige elever.

### ***Forhold i eksperimentklassene***

For å kunne identifisere betingelser for framgang ble rammer for undervisning i klassene kartlagt. Rammene for inferenstreningen varierte på flere områder. For det første var det forskjell i antall elever totalt i klassen, fra 13 til 23 elever. Korrelasjonsanalyser viste at eksperimentklassene med størst antall elever hadde lavest framgang i leseforståelse og omvendt. Med andre ord var inferenstreningen mest vellykket i de minste klassene med hensyn til leseforståelse. Det må imidlertid tas i betraktning at korrelasjonsanalysene kun inkluderte ti klasser, og enkeltklassers resultater fikk stor innflytelse på korrelasjonen. For det andre var det variasjon i hvor mye undervisningserfaring lærerne i eksperimentgruppen hadde. En lærer hadde ikke jobbet som lærer før da studien startet opp like etter skolestart høsten 2008. En annen lærer i eksperimentgruppen hadde derimot over 20 års undervisningserfaring. Korrelasjonsanalyser viste ingen sammenheng mellom lærernes undervisningserfaring og eksperimentelevenes framgang. Dette kan tolkes som at inferenstreningen var robust siden det var effekt selv for klasser hvor lærerne hadde lite undervisningserfaring. For det tredje viste observasjonene at det var noe forskjell i hvor lang tid klassene i eksperimentgruppen brukte på å komme til ro etter friminutt eller etter bytte av aktivitet. Eksperimentklassene som brukte mest tid til å komme til ro skilte seg ikke negativt ut i forhold til framgang i inferens og leseforståelse. For det fjerde rapporterte flere lærere gjennom loggskjemaene at de brukte mer enn 20 minutter på leksjonene. Lærerne ble bedt om å rapportere hvor mye tid som gikk med spesielt til inferenstreningen. Det var en positiv sammenheng mellom gjennomsnittlig rapportert tid brukt per inferensleksjon og framgang i leseforståelse. Positiv sammenheng mellom tid brukt per leksjon og framgang i leseforståelse indikerer at det er behov for å bruke mer tid per tekst enn det som det ble lagt opp til i denne studien dersom inferenstreningen skal anvendes i en ordinær undervisningssituasjon. Det siste punktet hvor det var forskjeller mellom eksperimentklassene var i forhold til inferenstreningens intensitet. Flertallet av klassene gjennomførte treningen i løpet av tre til fire uker. To klasser brukte imidlertid sju og åtte uker på inferenstreningen, hvorav den ene klassen fikk et opphold i inferenstreningen på fire uker. Disse to klassene skilte seg ikke nevneverdig fra øvrige eksperimentklasser med hensyn til endringsskårer i inferens og leseforståelse. I tillegg viste informasjon fra loggskjemaene at det i noen klasser var stor forskjell på arbeidstempoet hos elevene og dette kan ha gjort at det tok lengre tid å gjennomføre leksjonene.

Lærerne som deltok i denne studien, gjennomførte inferenstreningen etter de rammer og begrensninger som var nødvendig med tanke på å måle effekten av treningen. Antall tekster, instruksjon og tidsbruk var definert på forhånd. Selv

om implementeringsdataene viste at det var variasjon fra klasse til klasse i hvordan treningen ble gjennomført, var fleksibiliteten noe begrenset. Lærere som underviser i inferens i en ordinær undervisningssituasjon, vil i større grad ha mulighet for å tilpasse for eksempel vanskelighetsgrad på materiell, instruksjon og tidsbruk til elevenes individuelle behov.

## **7.5.2 Muligheter for generalisering**

I dette avsnittet diskuteres muligheter for generaliserbarhet. Det settes også fokus på både forutsetninger og begrensinger ved studien som må tas i betraktning i tolkningen av resultatene. I tillegg diskuteres forhold ved studien som med fordel kunne ha vært gjort annerledes.

### **7.5.2.1 Forutsetninger knyttet til gruppene**

En forutsetning for å diskutere effekter av inferenstreningen er at eksperimentgruppen og kontrollgruppen var sammenlignbare før oppstart av inferenstreningen. Fordeling av skoler i eksperiment og kontrollgruppe ble foretatt med tilfeldig utvelgning. I og med at fordelingen av eksperimentelever og kontrollelever var tilfeldig, er det rimelig å forutsette at resultatene fra inferenstreningen ville vært de samme dersom kontrollgruppen hadde mottatt inferenstreningen.

Det var ingen forskjeller mellom eksperimentgruppen og kontrollgruppen med hensyn til effektvariablene inferens og leseforståelse ved pretest. I forhold til kontrollvariablene var det kun signifikant forskjell mellom gruppene med hensyn til ordavkodning. Gruppene var svært like med hensyn til kjønnsfordeling og andel av tospråklige elever. Opplysninger om klassene og skolene som deltok i studien viste at eksperimentgruppen og kontrollgruppen hadde forholdsvis like ytre rammer. Det var forskjeller mellom klassene i begge grupper med hensyn til antall elever i klassen og antall elever på skolen totalt, men om lag like stor forskjell i både eksperimentgruppen og kontrollgruppen. Det var generelt stor forskjell i undervisningserfaringen hos de 16 lærerne som deltok i studien, men gjennomsnittlig undervisningserfaring hos lærere i eksperimentgruppen og kontrollgruppen var forholdsvis lik, henholdsvis 10 og 8 år. Erfaringer fra observasjonene var at tre eksperimentklasser skilte seg ut ved at elevene brukte noe mer tid på å komme til ro etter friminutt, ellers viste observasjonene at eksperimentelever og kontrollelever hadde forholdsvis like rammer med hensyn til arbeidsro. Ut fra de opplysninger som foreligger om de ytre rammene for undervisning i studien, er det rimelig å konkludere med at de ytre undervisningsrammene var forholdsvis like for elever i eksperimentgruppen og kontrollgruppen. Det er

derfor ingen holdepunkter for at resultatene ville vært annerledes dersom inferenstreningen hadde blitt gjennomført blant kontrollelevne.

I denne studien var det inkludert flere variabler i tillegg til effektvariablene som hadde til hensikt å kontrollere i hvilken grad elevene i de to gruppene var sammenlignbare. Det var signifikant forskjell mellom gruppene med hensyn til en av disse kontrollvariablene, ordavkodning, til fordel for eksperimentgruppen. Med hensyn til generalisering var det et ankepunkt at gruppene ikke var helt like med hensyn til ordavkodning. ANCOVA viste imidlertid at ordavkodning ikke hadde nevneverdig innflytelse på posttestskårer i inferens, og kun en liten til middels effekt på posttestskårer i leseforståelse. I tillegg viste korrelasjonsanalyser at sammenhengen mellom ordavkodning og endring i inferens og leseforståelse var liten. Oppsummert ser ikke ordavkodning ut til å ha nevneverdig innflytelse på effektvariablene, og forskjellen mellom gruppene må derfor ansees som lite betydningsfull.

En annen forutsetning for å tolke resultatene er om effektene kan knyttes til lesing spesifikt eller om det er snakk om en generell forventningseffekt. Variabelen matematikkferdigheter ble inkludert i studien nettopp for å kunne fastslå om effekter var spesifikt knyttet til lesing, med andre ord for å kunne kontrollere for en eventuell Hawthorneeffekt. Det var ingen signifikante forskjeller i matematikk mellom eksperimentgruppe og kontrollgruppe på posttesten. Derfor er en rimelig konklusjon at effekter av inferenstreningen var spesifikt knyttet til lesing. Som det tidligere er gjort rede for, kan effekten av trening som har til hensikt å fremme leseforståelse, påvirkes av deltakerens motivasjon. Det at elevene er klar over at de får eller ikke får en undervisning som har til hensikt å gjøre dem flinkere i inferens og leseforståelse, kan gjøre dem særlig motiverte eller lite motiverte for å prestere på prøvene. Derfor ble variabelen motivasjon også inkludert i studien som en kontroll for Hawthorneeffekt. Det var ikke signifikante forskjeller mellom eksperimentgruppen og kontrollgruppen med hensyn til motivasjon, verken på pretest eller posttest. Dette betyr at inferenstreningen ikke hadde innflytelse på eksperimentelevenes motivasjon, slik motivasjon var målt. Det var heller ingen signifikant korrelasjon mellom endringsskårene på effektmålene og motivasjon i noen av gruppene. Det er derfor ikke sannsynlig at eksperimentelevenes framgang kan knyttes til deres motivasjon, slik motivasjon ble målt.

For å kunne tolke generaliserte effekter av inferenstreningen, var det nødvendig å kontrollere at inferenstreningen faktisk ble implementert. Loggskjema og observasjoner ble brukt som mål på implementering. Av de

åtte lærerne som fylte ut loggskjema, ble det rapportert at samtlige tekster og oppgaver ble gjennomført, med unntak av en tekst som ikke ble gjennomført i fire klasser. Dette er et tegn på at bruken av materialet var godt implementert. Når det gjelder hvordan lærerne formidlet instruksjonen til tekstene og oppgavene, er det vanskelig å vurdere om elevene har fått akkurat den samme instruksjonen i alle eksperimentklasser. Selv om erfaringene fra observasjonene viste at det var noe forskjell i hvordan lærerne presenterte og forklarte tekster og oppgaver, var det tydelig at alle lærere fulgte instruksjonen som hørte til de ulike leksjonene, i alle fall i de leksjonene hvor det ble observert. I tillegg til loggskjemaer og observasjoner viste eksperimentgruppen stor framgang på inferensprøven, hvis tekster og oppgaver var tilsvarende til tekster og oppgaver i inferenstreningen. Med andre ord har elevene hatt framgang nettopp i forhold til det de trente på i de åtte leksjonene. Dette er også et tegn på at inferenstreningen ble gjennomført i samsvar med hvordan treningen var planlagt gjennomført.

#### **7.5.2.2 Forutsetninger knyttet til måleinstrumenter**

Som det er gjort rede for i resultatkapittelet, var det ekstremverdier på instrumentene som målte effekter av inferenstreningen. Dette gjorde det umulig å fange opp den reelle effekten hos enkelte elever. Noen svært få elever hadde skåret 0 på inferensprøven ved pretest, og noen elever hadde toppskåre på delskårer på leseprøven både på pretest og posttest. Problemer knyttet til at noen få elever hadde takeffekter og gulveffekter, ble løst ved å ekskludere disse elevene fra analyser. I forhold til inferens hadde elevene som ble ekskludert 0-skåre på pretest, mens det ikke var noen elever som hadde 0-skåre på posttest. Med andre ord har ikke ekskluderingen av disse elevene ført til å forsterke effekten av inferenstreningen, snarere tvert imot. I forhold til delskårer i leseforståelse hadde flertallet av de ekskluderte elevene med toppskåre på posttesten, altså var den virkelige endringsskåren til flertallet av de ekskluderte elevene trolig høyere enn den målte. Siden flertallet av de ekskluderte elevene hørte til i eksperimentgruppen, ville heller ikke dette ha bidratt til å forsterke effekter av inferenstreningen. Det er altså ikke grunnlag for å hevde at effekten av inferenstreningen skyldes ekstremskårer ved målene på effekt.

Siden verken inferensprøven eller prøven i leseforståelse var standardisert på norsk, vet man ikke hva skåringsnivåene svarer til i den virkelige verden. Prøvene gir ikke informasjon om lesealder, følgelig gir ikke prøvene kunnskap om hvor mange ukers eller måneders leseferdighet effekten av inferenstreningen svarer til.

Prøven i leseforståelse hadde delskårer for faktaspørsmål og tolkningsspørsmål. Som det er gjort rede for i avsnitt 7.3.3.1, gjenspeiler disse spørsmålstypene ulike sider av leseforståelse. Prøven hadde også spørsmål knyttet til tekster i to sjangre, fiksjonstekster og nonfiksjonstekster. Like fullt var alle spørsmål del av samme prøve og det kan derfor stilles spørsmål ved hvorvidt spørsmålstypene og teksttypene faktisk kan betraktes som forskjellige variabler. På en annen side dekker den samlede skåre i leseforståelse bredt både i forhold til krav til forskjellig type leseforståelse og sjanger og er således et forholdsvis valid mål på leseforståelse.

Autentiske tekster kunne med fordel ha blitt brukt for en ekstern validering av målene på effekt. Dette fordi autentiske tekster kan gi verdifull informasjon om hva elevenes framgang betyr i praksis. For eksempel kunne autentiske tekster med spørsmål blitt inkludert både i pre- og posttestingen. En ville da kunne målt eksperimentelevenes prosentvise riktige svar før og etter inferenstreningen. Utvalget av norske prøver i leseforståelse med autentiske tekster er begrenset. Det finnes kartleggingsprøver i lesing med autentiske tekster, men mange av disse er lite egnet til å fange opp effekter av undervisning som har til hensikt å bedre leseforståelse, siden kartleggingsprøvene har til hensikt å fange opp de svake leserne. Ved å benytte kartleggingsprøver er det derfor sannsynlighet for at takeffekten ville vært stor, særlig på posttesten. Et alternativ til pre- og posttester med autentiske tekster kunne vært å få lærerens vurdering av elevenes evne til å svare på spørsmål knyttet til fagtekster før og etter inferenstreningen.

Leseforståelse er et verktøy i mange skolefag, og det er sannsynlig at inferenstreningen kan ha virket inn på elevenes måte å tolke fagtekster i skolefag som for eksempel samfunnsfag og naturfag. Dette fordi tekstene som ble brukt i treningen stort sett var fagtekster, og fordi effekten av treningen var så godt som like sterk både for fiksjonstekster og non fiksjonstekster. Denne studien gir imidlertid ikke svar på hvilke andre områder inferenstreningen kan ha overføringseffekt til.

Som det er gjort rede for i kapittelet om ferdigheter i leseforståelse, korrelerer ordforråd og leseforståelse høyt i tidligere studier. Resultatene fra denne studien viste samme tendens, men det var imidlertid kun moderate korrelasjoner mellom ordforråd og leseforståelse og inferens på pretest. Forskjeller mellom tidligere studier og denne studien i styrken på sammenhengen mellom ordforråd og mål på leseforståelse, kan skyldes at ordforråd inneholder både et dybdeaspekt og et breddeaspekt. Videre kan måten ordforrådet måles på ha konsekvenser for deltakernes respons til



prøven og dermed deres skåre på ordforråd. Beck og McKeown (1991) hevder at det finnes flere typer og grader av ordkunnskap som handler om dybde i ordforrådet. For det første hevder de at det er flere typer kunnskap involvert i å kunne ord, for eksempel evnen til å kunne definere et ord med andre ord, og det å kjenne til flere betydninger av ordet. Den ovennevnte ordforrådsprøven krever at deltakerne kjenner til flere betydninger av ordet, men kan ikke gi informasjon om for eksempel hvorvidt deltakerne er i stand til å bruke ordet aktivt i diskusjon og i produksjon av egne tekster. Videre graderer Beck og McKeown kunnskap om ord fra å ikke ha kjennskap til ordet til å kunne bruke ordet aktivt. Ordforrådsprøven som er brukt i denne studien, krever at deltakerne er i stand til å gjenkjenne det riktige ordet som passer til den oppgitte forklaringen, men gir ikke informasjon om hvor godt deltakerne kjenner ordet, for eksempel om de kan gjenkjenne feil bruk av ordet. Prøven som er brukt i denne studien gir på den ene siden begrenset informasjon om deltakernes dybde i ordforrådet i og med at de ikke får muligheten til selv å definere ord. På en annen side har distraktorene i prøven en semantisk likhet med det riktige ordet, noe som krever at deltakerne har såpass presis kunnskap om forklaringsordet at de kan skille ut synonymet fra de semantisk like distraktorene.

Et annet relevant spørsmål som har konsekvenser for måling av ordforråd, er hvorvidt ordene som brukes til å måle ordforråd er representative. Ordnes representativitet handler om bredden i ordforråd, noe som ifølge Tannenbaum et al. (2006) er sterkere knyttet til leseforståelse enn dybde (jevnfør avsnitt 2.2.2). Skal ordene være representative for det norske språket, eller skal de være representative for de ordene sjetteklassinger møter i sin hverdag? Ordforrådsprøven som er brukt i denne studien er en blanding, ordene er valgt både fra en norsk ordbok og fra litteratur som sjetteklassinger bruker i skolen.

Et viktig poeng med pedagogisk forskning er ikke bare å finne ut av hva som gir effekter, men også hvilke elever undervisningen gir effekter for. Hensikten med kontrollvariablene var ikke bare å kontrollere for hvorvidt de to gruppene var like, men også for å kunne identifisere elever som eventuelt profitterte særlig lite eller mye på inferenstreningen. Kontrollvariablene kunne i liten grad identifisere disse elevene. Et betimelig spørsmål er derfor om andre kontrollvariabler burde vært inkludert. En aktuell kontrollvariabel som med fordel kunne vært inkludert, er fagspesifikk kunnskap. Som det er gjort rede for i avsnitt 2.3.3, har kunnskap om tekstens tema betydning for leserens forståelse. Det er derfor sannsynlig at elever med lite relevant kunnskap om temaene i tekstene som inngikk i inferenstreningen profitterte mindre på inferenstreningen enn elever med mye relevant kunnskap om tekstenes tema.

En svakhet ved denne studien er at det ikke foreligger data på langsiktige effekter av inferenstreningen ut over mål på effekt fem uker etter at inferenstreningen var avsluttet. Det ville derfor vært en klar fordel om elevene hadde blitt testet med målene på effekt for eksempel seks måneder etter avsluttet inferenstrening.

### **7.5.2.3 Begrensninger i og muligheter for generalisering**

Et ankepunkt som angår generalisering, er hvorvidt resultatene fra studien kan generaliseres til andre elever. Utvalget i denne studien kan ikke karakteriseres som tilfeldig i og med skolene ble valgt ut fra et geografisk begrenset område, hvilket setter visse begrensninger for generaliserbarheten. Det ble imidlertid gjort en skjønnsmessig utvelgning av skoler som fikk forespørsel om deltakelse i studien. Det ble forsøkt å velge skoler som var representative for den norske skole med hensyn til antall elever på skolen, geografi, både skoler i bykommuner og kommuner i landlige områder, og sosioøkonomiske forhold i skolekretsen. Rektor i samråd med aktuelle lærere avgjorde hvorvidt skolen skulle delta i studien. Studien forpliktet lærerne til å samarbeide om innhenting av data og til å gjennomføre åtte undervisningsøkter dersom de ble trukket ut til å delta i eksperimentgruppen. Det er mulig at skoler og lærere som takket ja til å delta, var spesielt motiverte og interesserte i leseopplæring. Dette ville i så fall ha slått likt ut for både eksperimentgruppen og kontrollgruppen, siden deltakelse i gruppe ble avgjort ved loddtrekning etter at skoler og lærere allerede var rekruttert.

På tross av at generaliserbarheten reduseres når utvalget ikke var tilfeldig valgt, er effektstudier uten ekte randomisering av praktiske årsaker mest vanlig i studier gjennomføres i skolen. Selv om ekte randomisering er idealet for eksperimentelle studier, viser Pressley (2006) til flere studier som tyder på konklusjonene fra randomiserte eksperimenter og kvasieksperimenter ofte er sammenfallende. National Reading Panel (2000) har undersøkt effekter som funksjon av kvalitet i undersøkelser i effektstudier av trening i fonologisk bevissthet. National Reading Panel fant at effektene var sterkest i studier som hadde solide design. Med andre ord kan effektene være en indikasjon på studiens kvalitet.

Denne studien omfattet materiell til åtte økter med inferenstrening. Selv om resultatene viste effekt av inferenstreningen, er det ikke mulig å trekke slutninger om effekter av tilsvarende materiell. For eksempel dersom en lærer konstruerer tilsvarende tekster med tilsvarende grafiske modeller, kan hun ikke være sikker på å få den samme effekten. Effektene som det er referert til i

denne studien, gjelder også bare for klasser som gjennomførte samtlige tekster og oppgaver, med unntak av noen få tekster. Det ville selvsagt vært mulig å teste ut effekten av kun deler av materialet. Dette kunne vært gjort ved å inkludere flere eksperimentgrupper som kun arbeidet med deler av materialet i inferenstreningen. Dette ble gjort i noen av de tidligere effektstudiene av inferenstrening (se kapittel 5). Flere eksperimentgrupper ville imidlertid ha krevd flere deltakere, noe som ikke var gjennomførbart innenfor rammene for denne studien.

Et annet spørsmål som angår generalisering, er hvorvidt man kan generalisere til andre som gjennomfører inferenstreningen. Med andre ord om lærere som gjennomfører inferenstreningen i andre klasser kan forvente å få tilsvarende effekter som i denne studien. Alle klasser og skoler er noe forskjellige, derfor vil rammene for inferenstreningen være annerledes dersom den gjennomføres i andre klasser. Muligheten for å generalisere til andre som gjennomfører inferenstreningen, avhenger derfor av hvor robust inferenstreningen er. Sammenhenger mellom opplysninger om rammer for inferenstreningen i de ti eksperimentklassene og klassenes framgang er en indikasjon på hvor robust treningen var. Det var signifikante primær overføringseffekter og generaliserte effekter for samtlige eksperimentklasser. Det var imidlertid forskjell i rammer for undervisningen i de ti eksperimentklassene på flere områder, jevnfør avsnitt 7.5.1.3. Disse forskjellene mellom eksperimentklassene kan tolkes på at inferenstreningen er robust. Samtlige eksperimentklasser hadde framgang både i forhold til prestasjoner i inferens og leseforståelse på tross av variasjon i rammer for inferenstreningen, tidsbruk per leksjon og inferenstreningens intensitet. Det er derfor rimelig å konkludere med at effektene av inferenstreningen kan generaliseres til forskjellige varianter av gjennomføring av undervisningsopplegget.

### **7.5.3 Konklusjon**

Denne effektstudien viste at inferenstrening hadde en effekt på elevenes evne til å gjøre inferenser og på deres leseforståelse. Effektene for inferens og tolkningsspørsmål på prøven i leseforståelse var opprettholdt fem uker etter endt inferenstrening, men det var ikke ytterligere effekt. Det var ikke mulig å identifisere elever som profitterte særlig mye eller lite ved hjelp av kontrollvariablene ordavkodning, ordforråd eller nonverbale ferdigheter, slik disse variablene ble målt. Både gutter og jenter profitterte på inferenstreningen.

Samtlige eksperimentklasser hadde framgang på målene på effekt, men klasser med få elever hadde sterkere effekt enn klasser med mange elever. Det

var størst effekt av inferenstreningen i de klassene som hadde brukt mest tid per leksjon. Det var effekt av inferenstreningen tross forskjeller i rammer for undervisning i eksperimentklassene. Dette betyr at gjennomførelsen av inferenstreningen kan generaliseres.

Lesing er resultatet de to komponentene avkoding og språkforståelse ganget med hverandre. Med en overvekt av forskning på undervisning i avkoding, er denne studien bidrag til hva som skal til for å fremme elevens forståelse for tekst, ikke bare avkodingen av den. Inferenstreningen i denne undersøkelsen satte fokus på at lesing også inkluderer tenkning, og at leseren selv må bidra med egen kunnskap for å rekonstruere den mening som ligger i teksten. Som metaanalysen i kapittel 5 viste, er det få effektstudier av inferenstrening. Derimot er det langt flere effektstudier av undervisning som trener flere ferdigheter i leseforståelse samtidig. Denne typen studier kan gi kunnskap om effekter av leseforståelsesundervisning generelt, men ikke om effekten av enkeltferdigheter for den overordnede leseforståelse, og dermed heller ikke om årsaksforbindelser i leseforståelse. Denne effektstudien hadde nettopp fokus på en enkelt ferdighet av betydning for leseforståelse. Den teoretiske motivasjonen for studien var å undersøke årsaksforholdet mellom inferens og leseforståelse, og resultatene kunne bekrefte inferens som en årsak til leseforståelse. Pedagogisk har denne studien vist at elevers evne til å gjøre inferenser kan fremmes gjennom inferenstrening via grafiske modeller. Særlig verdifullt ved treningen var at den også hadde overføringseffekt til leseforståelse. Inferenstrening vil derfor være et utbytterikt verktøy i undervisningen i leseforståelse hos norske elever.



## **8. Oppsummering og konklusjoner**

Denne avhandlingen har satt fokus på betydningen av inferens for leseforståelse, samt effekter av å trene inferens blant skolebarn. I kapittel 2 ble det redegjort for inferens som en av flere ferdigheter av betydning for den overordnede leseforståelsen. Inferens er et begrep som i faglitteraturen er langt fra entydig. I kapittel 3 ble derfor en taksonomi av inferenstyper med utgangspunkt i teoretiske perspektiver på inferens presentert. Metaanalysen i kapittel 5 bekreftet at inferensbegrepet defineres og kategoriseres på mange måter. Like fullt lot det seg gjøre å kategorisere studiene i metaanalysen i forhold til taksonomien som ble presentert i kapittel 3. Metaanalysen viste større effekter for brobyggende inferenser enn for utfyllende inferenser, men det ble imidlertid påpekt at den pedagogiske verdien av å trene brobyggende inferens ikke nødvendigvis er større enn ved å trene utfyllende inferenser, da begge inferenstyper er nødvendige for å oppnå leseforståelse. Effektstudien i kapittel 7 fokuserte på utfyllende inferens. Effektene av trening i utfyllende inferens i studien i kapittel 7 var noe sterkere enn de vektede gjennomsnittlige effektstørrelsene av trening i samme inferenstype i metaanalysen.

I tillegg til å belyse hva inferens i lesing innebærer, har avhandlingen hatt en pedagogisk målsetting om å undersøke hvordan inferens kan trenes med tanke på å bedre skoleelevers evne til å gjøre inferenser og øke deres leseforståelse. Gjennom metaanalysen ble de fire komponentene å skape bevissthet om inferens, monitorering, aktivisering av relevant kunnskap og å skape forbindelser, identifisert. Metaanalysen ga imidlertid ikke svar på hvilke av de fire komponentene som var mest effektive for å fremme elevenes ferdigheter i inferens og leseforståelse. Dette fordi det var flere studier i utvalget som trente mer enn en komponent. Dette kan tolkes som at det var bred enighet om hvordan inferens kan trenes. Inferenstreningen som ble prøvd ut gjennom effektstudien blant norske elever, inneholdt tre av komponentene som ble utpekt gjennom metaanalysen. I forbindelse med utviklingen av materialet til inferenstreningen i effektstudien ble det gjennomført en forundersøkelse som ble presentert i kapittel 6. Det var en stor forskjell på resultatene i forundersøkelsen og effektstudien. Mens påvirkningen som elevene forundersøkelsen ble eksponert for ikke ga effekt, var det sterke effekter av inferenstreningen i effektstudien. Noe av forskjellen mellom betingelsene som deltakerne i de to studiene ble eksponert for, og som kanskje kan forklare forskjellen i resultatene, ligger i bevisstheten om, eller fokuset på, det leseren selv tilføyer teksten for å gi den mening. I inferenstreningen ble ”meningsgapet” mellom to informasjoner i teksten synliggjort ved en grafisk modell. Elevene fylte bokstavelig talt inn det implisitte i teksten i en modell.

Det at leseren selv tilføyer kunnskap til opplysningene i teksten, hadde ikke på langt nær like stort fokus i forundersøkelsen. Dette er en indikasjon på at det å gjøre elevene bevisste på at de selv må bidra med kunnskap for å få en sammenhengende forståelse av teksten, bør være sentralt i undervisningen i inferens. Inferenstreningen i effektstudien var tilpasset en norsk skolekontekst og prøvd ut innenfor rammene for ordinær undervisning. Dette gjør at resultatene fra effektstudien er anvendbare for den pedagogiske praksis, og et argument for det bør fokuseres på inferens i leseundervisningen i norsk skole.

Det overordnede målet for denne avhandlingen var å undersøke effekter av inferenstrening. I kapittel 4 ble det argumentert for måling av effekter på ulike nivåer og med både eksperimentdesignede og ikke-eksperimentdesignede instrumenter. Validitetsmessige problemstillinger knyttet til eksperimentelle design ble også diskutert i kapittelet, samt etiske overveielser knyttet til pedagogiske effektstudier. Både metaanalysen i kapittel 5 og effektstudien i kapittel 7 viste sterke effekter av inferenstrening både på inferens spesielt, og for leseforståelse generelt. Metaanalysen viste svært sterke effekter for inferens både målt ved prøve i primæreffekt og primær overført effekt. Effektstudien viste også sterk effekt av inferenstrening på primær overført effekt av inferens, dog noe mindre enn den gjennomsnittlige vektete effektstørrelsen i metaanalysen. Det var også en middels til sterk generalisert effekt på leseforståelse av inferenstrening både i metaanalysen og i effektstudien. Teoretisk innebærer dette at inferens kan betraktes som en årsak til leseforståelse. Metaanalysen viste videre at den primære overførte effekten avtok etter at inferenstreningen var avsluttet. Effektstudien viste at effektene både på tolkningsspørsmål på prøven i leseforståelse og inferens ble opprettholdt etter avsluttet trening, men det var ingen ytterligere framgang hos elevene som hadde deltatt i inferenstreningen. Dette kan bety at undervisning i å gjøre inferenser må foregå kontinuerlig for at elevene skal fortsette å utvikle sine ferdigheter i inferens.

Effektstudien forsøkte også å finne svar på hvilke variabler som betinget framgang på effektmålene. I effektstudien ble kontrollvariablene ordforråd, ordavkodning og nonverbale ferdigheter inkludert, disse variablene hadde, slik de ble målt, liten innflytelse på effekten. Det ble heller ikke påvist forskjeller i hvordan gutter og jenter profitterte på inferenstreningen. Verken metaanalysen eller effektstudien kunne svare på hvilke elever som profitterte særlig mye eller lite på inferenstreningen med hensyn til ferdigheter i inferens og leseforståelse. Enkelte av studiene i metaanalysen gjennomførte inferenstrening blant elever som hadde særlig gode eller svake ferdigheter i lesing. Det var imidlertid ikke hensiktsmessig å sammenligne effekter for

disse studiene, blant annet fordi inferenstreningen i disse studiene høyst sannsynlig var tilpasset nettopp elever med enten gode eller svake ferdigheter i lesing.

Internasjonale undersøkelser som PIRLS og PISA har avdekket at norske elever ikke har så gode leseferdigheter som ønsket og forventet, tatt i betraktning de ressurser som brukes på utdanning i Norge. Læreplanen har deretter fått ambisiøse mål for lesing og leseforståelse. Læreplanen er imidlertid ordknapp når det gjelder hvordan disse målene skal nås. Denne avhandlingen har vist at inferens er en ferdighet som bidrar til å bedre elevers leseforståelse. Således har avhandlingen bidratt med innsikt i nettopp hvordan vi kan arbeide med leseforståelse for nå ambisjonene i Kunnskapsløftet og dermed styrke norske elevers leseferdigheter. Det å kunne lese med god forståelse er blitt et fundament for å tilegne seg den kunnskap som er nødvendig for å kunne delta i samfunnet. Lesing blir derfor ofte karakterisert som et verktøy i skole og utdanning. Leseforståelse er imidlertid ikke bare vesentlig i en akademisk sammenheng. Å ha en god leseforståelse er også et viktig grunnlag for glede og utbytte av lesing ellers i livet.





## Referanser

- Allerup, P., & Elbro, C. (1998). Comparing differences in accuracy across conditions or individuals: An argument for the use of log odds. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 51(2), 409-424.
- Anderson, R. C., & Pearson, D. P. (1984). A schema-theoretic view of basic processes in reading comprehension. I R. Barr, M. L. Kamil & P. Mosenthal (red.), *Handbook of reading research* (s. 255 - 291). New York: Longman.
- Anderson, V. (1992). A teacher development project in in transactional strategy instruction for teachers of severely reading-disabled adolescents. *Teaching and Teacher Education*, 8(4), 391-403.
- Andreassen, R. (2008). *Eksplisitt leseforståelsesundervisning i norske femteklasser: et feltekspertiment: doktorgradsavhandling*. Stavanger: Universitetet i Stavanger, Det humanistiske fakultet.
- Apthorp, H. S. (2006). Effects of a supplemental vocabulary program in third-grade reading/language arts. *Journal of Educational Research*, 100(2), 67-79.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading. *Reading Research Quarterly*, 34(4), 452.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Beck, I. L., & McKeown, M. (1991). Conditions of vocabulary acquisitions. I R. Barr, M. L. Kamil, P. B. Mosenthal & D. Pearson, P. (red.), *Handbook of reading research* (vol. 2, s. 789-814). New York: Longman.
- Boulineau, T., Fore, C., Hagan-Burke, S., & Burke, M. D. (2004). Use of story-mapping to increase the story-grammar text comprehension of elementary students with learning disabilities. *Learning Disability Quarterly*, 27(2), 105-121.
- 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.
- Brandao, A. C. P., & Oakhill, J. (2005). "How do You Know This Answer?"-- Children's Use of Text Data and General Knowledge in Story Comprehension. *Reading and Writing: An Interdisciplinary Journal*, 18, 687-713.

- Bratsberg, B., Raaum, O., Hægeland, T., Adult Literacy and Life Skills, & Nasjonalt senter for leseopplæring og leseforskning. (2006). *Lese- og tallforståelse, utdanning og arbeidsmarkedssuksess*. Stavanger: Universitetet i Stavanger, Lesesenteret.
- Brown, R., Pressley, M., van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, 88(1), 18-37.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: basic concepts, applications, and programming*. Mahwah, N.J.: L. Erlbaum Associates.
- Cain, K. (1996). Story knowledge and comprehension skill. I C. Cornoldi & J. Oakhill (red.), *Reading comprehension difficulties. Processes and intervention* (s. 167-192). Mahwah, NJ: Lawrence Erlbaum.
- 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. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, 76(4), 683-696.
- Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, 29(6), 850-859.
- Cain, K., Oakhill, J., & Bryant, P. (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., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30(3), 681-694.
- Carr, E. M., Dewitz, P., & Patberg, J. (1983). The Effect of Inference Training on Children's Comprehension of Expository Text. *Journal of Reading Behavior*, 15(3), 1-18.
- Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. I H. W. Catts & A. G. Kamhi (red.), *The connections between language and reading disabilities* (s. 25-40). New York: Lawrence Erlbaum Associates.
- Cipielewski, J., & Stanovich, K. E. (1992). Predicting Growth in Reading-Ability from Childrens Exposure to Print. *Journal of Experimental Child Psychology*, 54(1), 74-89.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. utg.). Hillsdale, N. J.: Laurence Erlbaum.
- Connors, F. A. (2009). Attentional control and the Simple View of reading. *Reading and Writing*, 22(5), 591-613.
- Crumpler, M., & Colin, M. (2004). *Diagnostic reading analysis*. London: Hodder education.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33(6), 934-945.
- Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora, & Kalleberg, R. (2006). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi/elektronisk ressurs*. Oslo: Forskningsetiske komiteer.
- Dewitz, P., Carr, E. M., & Patberg, J. P. (1987). Effects of Inference Training on Comprehension and Comprehension Monitoring. *Reading Research Quarterly*, 22(1), 99-121.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: young children learning at home and school*. Baltimore, Md.: P.H. Brookes Pub. Co.
- Dickson, S. V., Simmons, D. C., & Kameenui, E. J. (1998). Text organization: Research bases. I D. C. Simmons & E. J. Kameenui (red.), *What reading research tells us about children with diverse learning needs. Bases and basics* (s. 239-277). Mahwah, NJ: Lawrence Erlbaum.
- Elbro, C. (2007). *Læsevanskeligheder*. København: Gyldendal.
- Floyd, R. G., Bergeron, R., & Alfonso, V. C. (2006). Cattell-Horn-Carroll cognitive ability profiles of poor comprehenders. *Reading and Writing*, 19(5), 427-456.
- Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). *Analysis of variance designs. A conceptual and computational approach with SPSS and SAS*. New York: Cambridge university press.
- Gardill, M. C., & Jitendra, A. K. (1999). Advanced story map instruction: Effects on the reading comprehension of students with learning disabilities. *Journal of Special Education*, 33(1), 2-17+28.
- Gaskins, I. (2003). Taking charge of reader, text, activity and context variables. I A. P. Sweet & C. E. Snow (red.), *Rethinking reading comprehension* (s. 141 - 175). New York: The Guilford Press.
- Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. I J. Oakhill & C. Cornoldi (red.), *Reading comprehension difficulties. Processes and intervention* (s. 1-13). Mahwah, NJ Lawrence Erlbaum.

- Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education (RASE)*, 7(1), 6-10.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, 101(3), 371-395.
- Guthrie, J. T., & Humenick, N. M. (2004). Motivating students to read. Evidence for classroom practices that increase reading motivation and achievement. I P. McCardle & V. Chhabra (red.), *The voice of evidence in reading research* (s. 329-354 ). Baltimore: P.H. Brookes.
- Guthrie, J. T., Taboada, A., & Coddington, C. S. (2007). Engagement practices for strategy learning in concept- oriented reading instruction. I D. S. McNamara (red.), *Reading comprehension strategies. Theories, interventions and technologies* (s. 241 - 266). London: Lawrence Erlbaum Associates.
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., m.fl. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96(3), 403 - 423.
- Guthrie, J. T., Wigfield, A., & Perencevich, K. C. (2004). *Motivating reading comprehension: concept-oriented reading instruction*. Mahwah, N.J.: L. Erlbaum Associates.
- Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of Integrated Instruction on Motivation and Strategy Use in Reading. *Journal of Educational Psychology*, 92(2), 331.
- Hansen, J. (1981). The Effects of Inference Training and Practice on Young Children's Reading Comprehension. *Reading Research Quarterly*, 16(3), 391-417.
- Hansen, J., & Pearson, P. D. (1983). An Instructional Study: Improving the Inferential Comprehension of Good and Poor Fourth-Grade Readers. *Journal of Educational Psychology*, 75(6), 821-829.
- Hays, W. L. (1994). *Statistics* (5. utg.). Fort Worth: Harcourt Brace.
- Holmes, B. C. (1985). The Effects of a Strategy and Sequenced Materials on the Inferential Comprehension of Disabled Readers. *Journal of Learning Disabilities*, 18(9), 542-546.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127-160.
- Høyen, T., & Lundberg, I. (2000). *Dysleksi. Fra teori til praksis* (2. utg.). Oslo: Gyldendal akademisk.
- Høyen, T., & Tønnesen, G. (1997). *Ordkjedetesten*. Stavanger: Stiftelsen dysleksiforskning.

- Idol, L. (1987). Group story mapping - a comprehension strategy for both skilled and unskilled readers. *Journal of Learning Disabilities*, 20(4), 196-205.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
- Johnson, D. D., & Johnson, B. v. H. (1986). Highlighting vocabulary in inferential comprehension instruction. *Journal of Reading* 29, 622-625.
- Joshi, R. M. (2005). Vocabulary: A Critical Component of Comprehension. *Reading & Writing Quarterly*, 21(3), 209-219.
- Kamil, M. L. (2004). Vocabulary and comprehension instruction. Summary and implications of the national reading panel findings. I P. McCardle & V. Chhabra (red.), *The voice of evidence in reading research* (s. 213-234). Baltimore: P.H. Brookes.
- Kavli, A.-B. (2008). *Policy impact of PIRLS 2006 in Norway*. Paper presentert på IEA General Assembly.
- Keenan, J. M. (1993). Thoughts about the minimalist hypothesis. Commentary on Garnham on reading-inference. *Psychology*, 4(2).
- 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.
- Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567-1577.
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. (2007). Comprehension in preschool and early elementary children: Skill development and strategy interventions. I D. S. McMamara (red.), *Reading comprehension strategies: Theories, interventions, and technologies* (s. 27-45). Mahwah, NJ: Erlbaum.
- Kintsch, W. (1993). Information accretion and reduction in text-processing: Inferences. *Discourse Processes*, 16(1-2), 193-202.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W. (2005). Comprehension. I S. G. Paris & S. A. Stahl (red.), *Children's reading comprehension and assessment*. (s. 71-92). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Kjærnsli, M., Lie, S., Olsen, R. V., Roe, A., & Turmo, A. (2004). *Rett spor eller ville veier? : norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003*. Oslo: Universitetsforl.

- Konheim-Kalkstein, Y. L., & van den Broek, P. (2008). The effect of incentives on cognitive processing of text. *Discourse Processes*, 45(2), 180-194.
- Kramer, S. H., & Rosenthal, R. (1999). Effect sizes and significance levels in small-sample research. I R. H. Hoyle (red.), *Statistical strategies for small sample research* (s. 59-79). London: SAGE publications.
- Kristoffersen, T. K. O. (2006). Hold Banen. I C. T. Haukeland (red.), *NM norsk i midten 5: Damm*.
- Kunnskapsdepartementet, & Utdanningsdirektoratet. (2006). *Læreplanverket for Kunnskapsløftet* (Midlertidig utg.). Oslo: Utdanningsdirektoratet.
- Levin, J. R. (2005). Randomized classroom trials on trial. I G. D. Phye, D. H. Robinson & J. R. Levin (red.), *Empirical methods for evaluation educational interventions* (s. 3-27). San Diego: Elsevier.
- Lie, S., Kjærnsli, M., Roe, A., & Turmo, A. (2001). *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i en internasjonalt perspektiv*. Oslo: Institutt for lærerutdanning og skoleutvikling.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *American Psychologist*, 48(12), 1181-1209.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif.: Sage.
- Lund, T. (2002). *Innføring i forskningsmetodologi*. Oslo: Unipub.
- Mayer, R. E. (2005). The failure of educational research to impact educational practice: Six obstacles to educational reform. I G. D. Phye, D. H. Robinson & J. R. Levin (red.), *Empirical methods for evaluating educational interventions* (s. 67-81). San Diego: Elsevier.
- McGee, A., & Johnson, H. (2003). The Effect of Inference Training on Skilled and Less Skilled Comprehenders. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 23(1), 49-59.
- McKoon, G., & Ratcliff, R. (1992). Inference During Reading. *Psychological Review*, 99(3), 440-466.
- Mezynski. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, 53(2), 253-279.
- Mucherah, W., & Yoder, A. (2008). Motivation for Reading and Middle School Students' Performance on Standardized Testing in Reading. *Reading Psychology*, 29(3), 214-235.
- Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 International report. IEA's study of reading literacy*

- achievement in primary schools in 35 countries*. Chestnut Hill, Mass.: International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report. IEA's progress in international reading literacy study in primary schools in 40 countries*. Chestnut Hill, Mass.: International Study Center, Lynch School of Education, Boston College.
- Nasjonalt læremiddelsenter. (1997). *Kartlegging av lesedugleik. 9. klasse (tidlegare 8. klasse). Lærarrettleiing*.
- Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, 73(2), 139-158.
- Nation, K., Clarke, P., & Snowling, M. J. (2002). General cognitive ability in children with reading comprehension difficulties. *British Journal of Educational Psychology*, 72(4), 549-560.
- Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, 70(1), B1-B13.
- 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*. Washington, DC: National Institute for Child Health and Human Development.
- Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. I D. McNamara (red.), *Reading comprehension strategies. Theories, interventions and technologies* (s. 47-71). London: Lawrence Erlbaum Associates.
- Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing*, 18(7), 657-686.
- Ogle, L. T., Begnum, A. C., & Solheim, R. G. (2008). *Comparisons of Teachers of Language Minority Fourth-graders in Norway and the United States: Results from PIRLS 2006*. Paper presentert på The 36rd IEA International Research Conference
- Palincsar, A. S. (2003). Collaborative approaches to comprehension instruction. I A. P. Sweet & C. E. Snow (red.), *Rethinking reading comprehension* (s. 99-114). New York: The Guilford Press.



- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension forstering and comprehension monitoring activities. *Cognition and Instruction, 1*, 117-175.
- Pallant, J. (2005). *SPSS survival manual. A step by step guide to data analysis using SPSS for Windows (Version 12)* (2. utg.). Maidenhead: Open University Press.
- Paris, S. G., & Hamilton, E. E. (2009). The development of children's reading comprehension. I S. E. Israel & G. G. Duffy (red.), *Handbook of research on reading comprehension* (s. 32-53). New York: Routledge.
- Pearson, D., P., & Fielding, L. (1991). Comprehension instruction. I R. Barr, M. L. Kamil, P. B. Mosenthal & D. Pearson, P. (red.), *Handbook of reading research* (vol. 2, s. 815-860). New York: Longman.
- Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94*(1), 23-33.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*(4), 461-469.
- Pressley, M. (1998). Comprehension strategies instruction. I J. Osborn & F. Lehr (red.), *Literacy for all. Issues in teaching and learning* (s. 113-133). New York: The Guilford Press.
- Pressley, M., Eldinary, P. B., Graskins, I., Schuber, T., Bergman, J. L., Almasi, J., m.fl. (1992). Beyond direct explanation - transactional instruction of reading- comprehension strategies. *Elementary School Journal, 92*(5), 513-555.
- Pressley, M., Graham, S., & Harris, K. (2006). The state of educational intervention research as viewed through the lens of literacy intervention. *British Journal of Educational Psychology, 76*(1), 1-19.
- Programme for International Student Assessment. (2001). *Knowledge and skills for life: first results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- Programme for International Student Assessment. (2004). *Learning for tomorrow's world: first results from PISA 2003*. Paris: Organisation for Economic Co-operation and Development.
- Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*(4), 289-312.

- Raven, J. C. (1958). *Standard progressive matrices: sets A, B, C, D and E*. London: J.C. Raven : H.K. Lewis.
- Reutzel, D. R., & Hollingsworth, P. M. (1988). Highlighting Key Vocabulary - a Generative-Reciprocal Procedure for Teaching Selected Inference Types. *Reading Research Quarterly*, 23(3), 358-378.
- Reyna, V. F. (2004). Why scientific research? The importance of evidence in changing educational practice. I P. McCardle & V. Chhabra (red.), *The voice of evidence in reading research* (s. 47-58). Baltimore: P.H. Brookes.
- Rosenshine, B., & Meister, C. (1994). Reciprocal Teaching - a Review of the Research. *Review of Educational Research*, 64(4), 479-530.
- Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavioral research: methods and data analysis* (3 utg.). Boston: McGraw-Hill.
- Samuelstuen, M. S., & Braten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology*, 46(2), 107-117.
- Seigneuric, A., & Ehrlich, M.-F. (2005). Contribution of Working Memory Capacity to Children's Reading Comprehension: A Longitudinal Investigation. *Reading and Writing: An Interdisciplinary Journal*, 18(7-9), 617-656.
- Seigneuric, A., Ehrlich, M.-F., Oakhill, J., & Yuill, N. (2000). Working memory resources and children's reading comprehension. *Reading and Writing*, 13(1-2), 81-103.
- Senter for leseforskning, & Progress in International Reading Literacy Study. (2003). *En Norsk kortversjon av den internasjonale rapporten om 10-åringers lesekunnskaper*. Stavanger: Senter for leseforskning.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Sjøberg, S. (2008). Norsk skole: styrt av Pisa i Paris? *Utdanning*, 4.
- Skog, O. J. (2004). *Å forklare sosiale fenomener. En regresjonsbasert tilnærming*. Oslo: Gyldendal akademisk.
- Snow, C. E., & Sweet, A. P. (2003). Reading for comprehension. I A. P. Sweet & C. E. Snow (red.), *Rethinking reading comprehension* (s. 224). New York: The Guilford Press.
- Solheim, R. G., Begnum, A. C., van Daal, V., Gabrielsen, N. N., Universitetet i Stavanger, Progress in International Reading Literacy Study, m.fl. (2007). *PIRLS: norske elevers leseinnsats og leseferdigheter: resultater for fjerde og femte trinn i den internasjonale studien PIRLS 2006*. Stavanger: Lesesenteret, Universitetet i Stavanger.

- Spear-Swerling, L. (2006). Childrens Reading Comprehension and Oral Reading Fluency in Easy Text. *Reading and Writing, 19*(2), 199-220.
- Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology, 41*(1), 225-234.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72-110.
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Mahwah, N.J.: Lawrence Erlbaum.
- Stanovich, K. E. (2000). *Progress in understanding reading. Scientific foundations and new frontiers*. New York: Guilford Press.
- Statistisk sentralbyrå. (2004). *Statistisk årbok 2004* (123. årg. utg.). Oslo: Statistisk sentralbyrå.
- Stothard, S. E., & Hulme, C. (1992). Reading-Comprehension Difficulties in Children - the Role of Language Comprehension and Working Memory Skills. *Reading and Writing, 4*(3), 245-256.
- Stothard, S. E., & Hulme, C. (1996). A comparison of reading comprehension and decoding difficulties in children. I C. Cornoldi & J. Oakhill (red.), *Reading comprehension difficulties. Processes and intervention* (s. 93-112). Mahwah, NJ: Lawrence Erlbaum.
- Sundbye, N. (1987). Text Explicitness and Inferential Questioning - Effects on Story Understanding and Recall. *Reading Research Quarterly, 22*(1), 82-98.
- Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*(3), 261-285.
- 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.
- Tornes, J. (1996). *Kartleggingsprøve: Matematikk: M4, M5, M6, M7: Lærerveiledning*. Jaren: PP-tjenestens materiellservice.
- van den Broek, P., Kendeou, P., Kremer, K., Lynch, J., Butler, J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. I S. G. Paris & S. A. Stahl (red.), *Children's reading comprehension and assessment* (s. 107-130). Mahwah, NJ: Erlbaum.
- van den Broek, P., & Kremer, K. E. (2000). The mind in action: What it means to comprehend during reading. I B. M. Taylor, M. F. Graves & P. van den Broek (red.), *Reading for meaning. Forstering*

- comprehension in the middle grades* (s. 1-31). New York: Teachers college press.
- van Keer, H. (2004). Forstering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology*, 74(1), 37 - 70.
- Vellutino, F. R., & Schatschneider, C. (2004). Experimental and quasi-experimental design in literacy research. I N. K. Duke & M. H. Mallette (red.), *Literacy research methologies* (s. 114 - 148). New York: The Guilford Press.
- Warren, W. H., Nicoholas, D. W., & Trabasso, T. (1979). Events chains and inferences in understanding narratives. I R. O. Freedle (red.), *New directions in discourse processing* (vol. II, s. 23-52). Norwood: ABLEX publishing corporation.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89(3), 420-432.
- Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., m.fl. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, 45(5), 432-445.
- Williams, J. P. (2005). Instruction in reading comprehension for primary-grade students: A focus on text structure. *Journal of Special Education*, 39(1), 6-18.
- 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.
- Winne, P. H., Graham, L., & Prock, L. (1993). A Model of Poor Readers Text-Based Inferencing - Effects of Explanatory Feedback. *Reading Research Quarterly*, 28(1), 52-66.
- Yuill, N. (2007). Visiting joke city: How can talking about jokes foster metalinguistic awareness in poor comprehenders? I D. McNamara (red.), *Reading comprehension strategies. Theories, interventions, and technologies* (s. 325-345). New York Lawrence Erlbaum Associates.
- Yuill, N., & Oakhill, J. (1988). Effects of Inference Awareness Training on Poor Reading-Comprehension. *Applied Cognitive Psychology*, 2(1), 33-45.

*Referanser*

---

Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: an experimental investigation*. Cambridge: Cambridge University Press.

## Liste over tabeller

Tabell 1	Oversikt over studiene i utvalget	s. 53
Tabell 2	Oversikt over deltakerne i studiene	s. 54
Tabell 3	Oversikt over inferenstyper i studiene	s. 57
Tabell 4	Oversikt over eksperimentgrupper, subgrupper og kontrollgrupper	s. 58
Tabell 5	Oppsummering av kjennetegn ved studiene	s. 66
Tabell 6	Oppsummering av organisatoriske trekk ved inferenstreningen i studiene i utvalget	s. 67
Tabell 7	Oversikt over mål på effekt i studiene	s. 70
Tabell 8	Oppsummering av effektstørrelser (Cohen's d) for primær- primær overført og generalisert effekt	s. 73
Tabell 9	Vektete gjennomsnittlige effektstørrelser for studier av brobyggende inferens og utfyllende inferens	s. 75
Tabell 10	Effektstørrelser for primær overføring målt med utsatt prøve	s. 77
Tabell 11	Oversikt over eksperimentbetingelse og kontrollbetingelse	s. 87
Tabell 12	Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler for gruppe 1 og gruppe 2, samt forskjeller mellom gruppene	s. 95
Tabell 13	Gruppenes frekvens i prosent for instruksjonsspørsmålene for instruksjonsspørsmålene til tekstene i oppgavesett A og B	s. 96
Tabell 14	Oversikt over korrelasjoner	s. 97
Tabell 15	Korrelasjon mellom instruksjonsspørsmål og inferensspørsmål	s. 99
Tabell 16	Oversikt over innhold i leksjonene	s. 117
Tabell 17	Studiens design	s. 117
Tabell 18	Skoler i utvalget	s. 118
Tabell 19	Klasser i utvalget	s. 119
Tabell 20	Kjønnfordeling, tospråklighet og alder i eksperimentgruppe og kontrollgruppe	s. 120
Tabell 21	Oversikt over variabler og prøver	s. 127
Tabell 22	Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på pretest, samt forskjeller mellom gruppene	s. 132
Tabell 23	Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på posttest, samt forskjeller mellom gruppene	s. 134

*Liste over tabeller*

---

Tabell 24	Gjennomsnitt, standardavvik, skewness og kurtosis for alle variabler på utsatt posttest	s. 136
Tabell 25	Oversikt over korrelasjoner mellom kontrollvariabler og effektvariabler	s. 139
Tabell 26	Ikke-justerte og justerte gjennomsnitt for matematikk målt ved posttest	s. 141
Tabell 27	Enveis ANCOVA for posttestskårer i matematikk kontrollert for matematikk målt ved pretest	s. 141
Tabell 28	Gjennomsnitt og standardavvik for gutter og jenter, samt kjønnsforskjeller for variablene målt ved pretest	s. 142
Tabell 29	Gjennomsnitt, standardavvik og signifikansnivå hos enspråklige og tospråklige målt ved pretest	s. 143
Tabell 30	Eksperimentklassenes endringsskårer og opplysninger om klasser og lærere i eksperimentgruppen	s. 157
Tabell 31	Ikke-justerte og justerte gjennomsnitt for inferens ved posttest	s. 159
Tabell 32	Enveis ANCOVA for inferens målt ved posttest kontrollert for inferens ved pretest, ordforråd og ordavkoding	s. 159
Tabell 33	Ikke-justerte og justerte gjennomsnitt for leseforståelse ved posttest	s. 160
Tabell 34	Enveis ANCOVA for leseforståelse ved posttest, kontrollert for pretestskårer i leseforståelse, ordforråd, ordavkoding og nonverbale ferdigheter	s. 160
Tabell 35	Korrelasjoner mellom log odds endringsskårer i inferens og leseforståelse og kontrollvariabler	s. 161
Tabell 36	Sammenligning av log odds endringsskårer fra pretest til posttest hos gutter i eksperimentgruppe og kontrollgruppe	s. 163
Tabell 37	Sammenligning av log odds endringsskårer fra pretest til posttest hos jenter i eksperimentgruppe og kontrollgruppe	s. 163
Tabell 38	Sammenligning av log odds endringsskårer fra pretest til posttest hos enspråklige i eksperimentgruppen og kontrollgruppen	s. 167
Tabell 39	Sammenligning av log odds endringsskårer fra pretest til posttest hos tospråklige i eksperimentgruppen og kontrollgruppen	s. 167

## Liste over figurer

Figur 1	Inferenstyper	s.34
Figur 2	Gruppenes samlede skåre på inferensspørsmål i oppgavesett A og B	s. 98
Figur 3	Gruppenes samlede skårer på faktaspørsmål i oppgavesett A og B	s. 100
Figur 4	Log odds transformerte endringsskårer for eksperimentgruppe og kontrollgruppe	s. 147
Figur 5	Eksperimentgruppens og kontrollgruppens endringsskårer (pretest til posttest) basert på råskårer på faktaspørsmål og tolkningsspørsmål	s. 149
Figur 6	Eksperimentgruppens og kontrollgruppens endringsskårer (pretest til posttest) basert på råskårer på fiksjonstekster og nonfiksjonstekster	s. 153
Figur 7	Sammenligning av klassenes log odds transformerte endringsskårer fra pretest til posttest på inferensprøven, hvor signifikansnivå er basert på t-test	s. 155
Figur 8	Sammenligning av klassenes log odds transformerte endringsskårer fra pretest til posttest på prøven i leseforståelse, hvor signifikansnivå er basert på t-test	s. 155
Figur 9	Log oddsskårer på åpne spørsmål og understreknings spørsmål på inferensprøven hos gutter og jenter i eksperimentgruppe og kontrollgruppe	s. 165
Figur 10	Log oddsskårer på faktaspørsmål og tolkningsspørsmål på prøven i leseforståelse hos gutter og jenter i eksperimentgruppe og kontrollgruppe	s. 165
Figur 11	Log oddsskårer på fiksjonstekster og nonfiksjonstekster på prøven i leseforståelse hos gutter og jenter i eksperimentgruppe og kontrollgruppe	s. 166
Figur 12	Utviklingen i råskårer i inferens til de fire klassene som deltok på alle tre måletidspunkter	s. 169
Figur 13	Utviklingen i råskårer i leseforståelse til de fire klassene som deltok på alle tre måletidspunkter	s. 170





## **Oversikt over appendiks**

Appendiks 1	Eksempel på inferenstekst med støtte (kapittel 6)
Appendiks 2	Semantisk ordforrådsprøve (kapittel 6)
Appendiks 3	Fonologisk ordforrådsprøve (kapittel 6)
Appendiks 4	Instruksjon til forundersøkelsen (kapittel 6)
Appendiks 5	Inferensprøven (kapittel 7)
Appendiks 6	Instruksjon til inferensprøve (kapittel 7)
Appendiks 7	Ordforrådsprøve (kapittel 7)
Appendiks 8	Instruksjon til ordforrådsprøve (kapittel 7)
Appendiks 9	Eksempel på materiell fra inferenstreningen (kapittel 7)
Appendiks 10	Eksempel på leksjonsbeskrivelse (kapittel 7)
Appendiks 11	Eksempel på loggskjema (kapittel 7)
Appendiks 12	Observasjonsskjema (kapittel 7)

## **Appendiks 1: Eksempel på inferenstekst med støtte (kapittel 6)**



## Norske emigranter

På 1800- og 1900-tallet utvandret mer enn 800 000 nordmenn til Amerika. Den første skuten som fraktet norske emigranter heiste seil i 1825. De første emigrantskipene var relativt små, senere foregikk transporten av emigranter med større skip som fullriggere og barker. Overfarten til Amerika kunne være en påkjenning for passasjerene. Mangel på vann begrenset mulighetene for renhold, toalettene om bord var ofte veldig urenlige. Passasjerene sov svært tett, og lopper og lus spredde seg raskt fra køye til køye.

Ordene **seil**, **fullrigger** og **bark** forteller deg om skutene som ble brukt til å frakte emigranter til Amerika. Har du en fellesbetegnelse på skutene?

Ordene **begrenset mulighetene for renhold**, **urenlige**, **lopper og lus** forteller om hvorfor reisen kunne være en påkjenning for passasjerene. Hva besto denne påkjenningen i?

Det var vanlig at emigrantene hadde med, og tilberedte mat selv. Maten måtte tåle lagring så lenge reisen varte. Den vanligste provianten besto derfor av saltet eller tørket mat som for eksempel spekekjøtt og flatbrød. Emigrantene ble også oppfordret til å ha med gryter, bestikk og servise. Daglige rasjoner med vann og ved var vanligvis inkludert i billetten.

Under reisen var passasjerene opptatt av daglige gjøremål, det kunne være matlaging, strikking, lek eller spill. Søndager var det vanlig at kapteinen holdt gudstjeneste på dekket. Dårlig vær satte imidlertid begrensninger for aktivitetene.

Emigrantene fryktet storm og dårlig vær, skjønt de ville helst også unngå lange perioder med stille vær. Uvær gjorde selvfølgelig reisen ubehaglig, men i de første femti år som utvandringen til Amerika pågikk, var det faktisk ingen som omkom i forlis eller havari. Likevel var det en del emigranter som døde før de nådde Amerika.

Kilde: Solem, B. (2000). Med de største og hurtigste skibe. I: Det løfterike landet. Nasjonalbiblioteket & Norsk utvandrer museum (nedlastet 14.02.2007).

### Spørsmål:

1. Hvor vanskelig synes du at teksten var? Sett et kryss på linjen fra svært enkel til svært vanskelig.

Svært enkel \_\_\_\_\_ svært vanskelig

2. Hvorfor ville emigrantene unngå stille vær?

3. Hva tror du var årsaken til at del emigranter døde før de nådde Amerika?

4. Hvor vanskelig synes du spørsmål 2 og 3 var? Sett et kryss på linjen fra svært enkel til svært vanskelig.

Svært enkel \_\_\_\_\_ svært vanskelig

5. Hvem pleide å holde gudstjeneste om bord i emigrantskipene?

6. Hvor vanskelig synes du spørsmål 5 var? Sett et kryss på linjen fra svært enkel til svært vanskelig.

Svært enkel \_\_\_\_\_ svært vanskelig

**STOPP!**





## **Appendiks 2: Semantisk ordforrådsprøve (kapittel 6)**



## Øvingsoppgave

Seksløper	Pistol	Revolver	Gevær
Førerkort	Diplom	Identifikasjon	Sertifikat

1. Finnerlønn	Honorar	Gratiale	Dusør
2. Foredrag	Tale	Dialog	Diskusjon
3. Eksamensbevis	Vitnemål	Protokoll	Karakter
4. Altan	Karnapp	Pergola	Balkong
5. Oppheve	Benekte	Dementere	Annullere
6. Seilskip	Skonnert	Minesveiper	Snekke
7. Sammenheng	Betingelse	Forbindelse	Omstendighet
8. Fruktgrøt	Fromasj	Sorbé	Kompott
9. Samling av stjerner	Asteroide	Galakse	Meteor
10. Sverd	Bajonett	Spyd	Sabel
11. Lysbilde	Film	Fotografi	Dias
12. Plattform	Fundament	Grunnmur	Hjørnestein
13. Bestilling	Ultimatum	Rekvisisjon	Påbud
14. Enerett	Monopol	Egenrådlig	Analogi
15. Tohjulet hestekjøretøy	Ekvipasje	Karet	Sulky
16. Sannsynlighet	Skjebne	Håp	Probabilitet
17. Garantist	Kontorist	Kapitalist	Kausjonist
18. Misdannet	Deformert	Kaputt	Skadet
19. Motvilje	Fortvilelse	Sjenerthet	Aversjon
20. Drøfte	Instruere	Konferere	Irettesette
21. Antakelse	Dilemma	Hypotese	Problem
22. Avgjørende	Presserende	Tungtveiende	Utslagsgivende
23. Synsvinkel	Affinitet	Aspekt	Proporsjon
24. En som lager tønner	Garver	Bøkker	Fløter
25. Instrument med mange strenger	Cymbal	Fagott	Sitar
26. Kategori	Klasse	Kasus	Prototyp
27. Oppdra	Akseptere	Aktualisere	Sivilisere
28. Skape samtidighet	Influere	Sirkulere	Synkronisere
29. Blåøyd	Rigid	Misunnelig	Naiv
30. Bedøvelse	Rehabilitasjon	Anestesi	Affeksjon

## **Appendiks 3: Fonologisk ordforrådsprøve (kapittel 6)**



## Øvingsoppgave

Spisested	Reparasjon	Revolusjon	Restaurant
Hard bergart	Granitt	Granat	Graffiti

1. Spreng med et smell	Eksplodere	Eksponere	Eksportere
2. Pusse	Polere	Posjere	Posere
3. Ønske til lykke	Graduere	Gratulere	Granulere
4. Anbringe	Plagiere	Plassere	Planere
5. Kreve	Insistere	Inspisere	Inspirere
6. Forkle	Markere	Marsjere	Maskere
7. Resultat	Fasett	Falsett	Fasit
8. Iakttå	Obdusere	Operere	Observere
9. Fullkommenhet	Perfeksjon	Persepsjon	Permisjon
10. Kunstig	Syntaktisk	Syntetisk	Synoptisk
11. Penger	Kapital	Kapittel	Kapillar
12. Fremsi	Residere	Resignere	Resitere
13. Vanskeliggjøre	Komponere	Komplisere	Komprimere
14. Fasthet	Kontinens	Konsistens	Kontinent
15. Tilegne	Desertere	Dedikere	Designere
16. Beskyttelses plate	Visir	Visitt	Viser
17. Dempe	Moderere	Modellere	Modulere
18. Etterligne	Invitere	Integrere	Imitere
19. Adgang	Aksent	Aksept	Aksess
20. Oppstigende luftstrøm	Termin	Termitt	Termikk
21. Innstille	Nominere	Normere	Notere
22. Sammenblanding	Konklusjon	Konfusjon	Konsesjon
23. Kjennetegn	Monografi	Monogam	Monogram
24. Rang	Statutt	Statue	Status
25. Forargelse	Indikasjon	Indignasjon	Indiskresjon
26. Meningsmåling	Galla	Gallup	Galopp
27. Utmerket	Eminent	Emigrant	Emirat
28. Fratre	Abstrahere	Abdisere	Adusere
29. Sykehus	Hospitant	Hospital	Hospits
30. Tryllekunstner	Illusorisk	Illustratør	Illusjonist



## **Appendiks 4: Instruksjon til forundersøkelsen (kapittel 6)**



### **Instruksjon til forundersøkelsen**

To skoleklasser deles i fire grupper á 10 til 15 elever. Første hefte er felles for alle elever, hefte to (inferenstekster) rulleres på følgende måte:

Gruppe 1: Sett A med støtte, sett B uten støtte

Gruppe 2: Sett B uten støtte, sett A med støtte

Gruppe 3: Sett A uten støtte, sett B med støtte

Gruppe 4: Sett B med støtte, sett A uten støtte

Sett A:

Vennenes samfunn

Fabrikken

Gol stavkirke

Sett B:

Katland

Bensinstasjonen

Norske emigranter

ID:

Siffer: gruppe/ hefte nr (1, 2, 3, 4)

Bokstav: skole (L eller G)

Bokstav: Klasse (D eller B, A eller B)

Siffer: Elev nr

### **Generell instruksjon**

Elevene sitter slik at de ikke kan samarbeide, alle har kun blyant eller penn framme på pulten.

*I undersøkelsen som dere skal være med på er anonym, det vil si at ingen skal vite hvem som har svart hva. Dere skal derfor ikke skrive navet deres på arkene dere får utdelt.*

*Dere skal nå få et hefte med oppgaver som dere skal arbeide med. Heftet inneholder forskjellige leseoppgaver. I noen oppgaver skal dere kun lese ord, andre ganger skal dere arbeide med tekster med spørsmål til.*

*Når dere jobber med oppgavene skal dere velge svaralternativer og krysse av (demonstrer på tavle). Dere skal ikke bruke viskelær. Dersom dere krysser av feil, setter dere en ring rundt det gale svaret og krysser av på rett sted (demonstrer på tavle).*

*Jeg skal forklare dere alle oppgavene, noen ganger er en eller to prøveoppgaver. Etter en del av oppgavene finner dere ordet stopp. Det betyr at dere ikke skal bla videre før det blir gitt beskjed om det.*

*På de tre første oppgavene får dere en bestemt tid til å arbeide med oppgaven på. Det er ikke forventet at alle skal klare å lese alt eller svare på alle spørsmålene. Noen vil nok bli ferdige før de andre og blir nødt til å sitte og vente litt underveis. Det er da viktig at dere ikke blar videre i heftet. Det er lov å tegne på blanke sider, men det er ikke lov å forstyrre de som fremdeles arbeider.*

*Dersom dere ikke blir ferdige med en oppgave når tiden er ute, setter dere ei stjerne i margen for å vise hvor langt dere kom. Hvis dere leste da tiden gikk ut, setter dere ei stjerne til høyre for den siste linja dere leste. Hvis dere arbeidet med et spørsmål da tiden gikk ut, setter dere ei stjerne til høyre for spørsmålet (demonstrer).*

For hver oppgave får dere vite hva dere skal gjøre. Det er viktig at dere arbeider så godt dere kan og så raskt som dere kan.

**Kryss av for kjønn - gutt/ jente.**

### **Ordavkoding**

*Øverst på siden ser du ei tegning. Ved siden av ei rekke med ord som du skal lese. Ett av ordene passer til tegningen. Sett kryss på det ordet (demonstrerer).*

*Se nå på tegningen på neste linje. Les ordet i rekka ved siden av. Sett kryss på det ordet som passer til tegninga (sjekk at alle krysser rett).*

*Nå skal du gjøre på samme måte med alle oppgavene på denne siden og på de fem neste sidene – men vent til jeg sier i fra. Se først på tegningen, les deretter ordenen ved siden av og kryss av på det ordet du tror det er. Du kan jobbe helt til du kommer fram til ordet ”stopp”. Dersom du blir ferdig før tiden er ute, skal du ikke bla videre. Dersom du ikke er helt ferdig med oppgava når tiden er ute, lager du ei stjerne i margen på høyre side. Du får fem minutter til å gjøre alle oppgavene. Begynn!*

### **Tårer lagt i jorden**

I denne oppgaven skal du lese en beskrivelsestekst. Du skal først lese hele teksten så raskt du kan og så godt du kan. Så svarer du på spørsmålene til slutt.

*Les øvingsoppgaven og svar på spørsmålet nedenfor (sjekk at alle krysser rett).*

*Les videre hele teksten under og svar på spørsmålene på de tre neste sidene. Dersom du kommer til ordet ”stopp”, sitter du stille og venter. Dersom du ikke er helt ferdig når tiden er ute, lager du ei stjerne i margen på høyre side. Du får seks minutter til å gjøre denne oppgaven. Begynn!*

### **Kroppens klokke**

I denne oppgava skal du lese en tekst som heter ”Kroppens klokke”. Oppgaven starter med en kort tekst og noen spørsmål. Så kommer en lengre tekst og ti spørsmål til slutt.

*Les øvingsoppgaven og svar på spørsmålet nedenfor (sjekk at alle krysser rett).*

Les videre teksten nedenfor. Les først de to første avsnittene og svar på de fire første spørsmålene. Les så resten av teksten og svar på spørsmål til slutt. Når du kommer til ordet ”stopp”, vet du hva du skal gjøre, og husk å sette stjerne dersom du ikke blir ferdig. Du får seks minutter til å gjøre denne oppgaven. Begynn!

### **Semantisk ordforrådsprøve**

I første rekke finner du et ord. Du skal finne ut hvilket av de tre neste ordene som betyr det samme som det første. Når du har funnet det riktige ordet skal du sette en ring rundt det ordet som er riktig. Dersom du ringer rundt feil ord, setter du bare et kryss over ordet og ringer rundt det som er riktig. Vi begynner med en øvingsoppgave:

*Seksløper, der det en pistol, revolver eller gevær?  
Seksløper det er en revolver, sett ring rundt det andre ordet.*

*Førerkort, er det det samme som diplom, identifikasjon eller sertifikat?  
Førerkort og sertifikat er det samme som sertifikat, derfor ringer du rundt det siste ordet.*

*Bla over til neste side, der finner dere samme type oppgaver som dere skal gjøre når jeg sier i fra. Nå skal dere ikke si svaret høyt, bare krysse av på rett ord uten å snakke sammen. Det er bare lov å krysse av på ett ord på hver oppgave. Jeg kommer til å lese ordene høyt for dere. Hvis jeg leser for fort så rekker dere en hånd i været og sier i fra, så skal jeg si ordene en gang til.*

Les ordene høyt i rolig tempo på samme måte som i eksemplene, spør etter et par oppgaver om det går for fort eller sakte. Les hver oppgave en gang, gjenta oppgaven dersom noen ber om det.

Dersom elever kommenterer at de ikke vet svaret, eller at de er syns to av alternativene virker like riktige, si at de skal krysse av for det de mener er mest riktig – kun ett kryss er lov per oppgave.

### **Fonologisk ordforrådsprøve**

*I denne oppgaven skal dere gjøre som i den forrige. I første rekke finner du et ord. Du skal finne ut hvilket av de tre neste orden som betyr det samme som det første. Sett ring rundt det riktige ordet. Først en øvingsoppgave:*

*Spisested, er det det samme som reparasjon, revolusjon eller restaurant?  
Spisested det er det samme som restaurant derfor ringer du rundt det siste ordet*

*Hard bergart, er det det samme som granitt, granat eller graffiti?  
Granitt er en har bergart, så da setter du ring rundt det første ordet.*

*Bla over til neste side, der finner dere samme type oppgaver som dere skal gjøre når jeg sier i fra. Nå skal dere ikke si svaret høyt, bare krysse av på rett ord uten å snakke sammen. Det er bare lov å krysse av på ett ord på hver oppgave. Jeg kommer til å lese ordene høyt for dere. Hvis jeg leser for fort så rekker dere en hånd i været og sier i fra, så skal jeg si ordene en gang til.*

Les ordene høyt i rolig tempo på samme måte som i eksemplene, spør etter et par oppgaver om det går for fort eller sakte. Les hver oppgave en gang, gjenta oppgaven dersom noen ber om det.

Dersom elever kommenterer at de ikke vet svaret, eller at de er syns to av alternativene virker like riktige, si at de skal krysse av for det de mener er mest riktig – kun ett kryss er lov per oppgave.

### **Inferenstekster uten støtte**

*Dere skal nå få tre korte tekster med spørsmål til. Les først igjennom teksten nøye to ganger. Når du har lest to ganger blar du over til neste side og bretter siden du har lest under resten av arkene – slik (demonstrerer). Når dere først har bladd over, er det ikke lov å bla tilbake. Så skal dere svare på spørsmål til teksten.*

#### Dersom første uten støtte er første sett:

*Første spørsmål handler om hvor lett eller vanskelig du synes at teksten var. Under spørsmålet er det en strek. På den ene siden av streken står det svært enkelt, på den andre siden står det svært vanskelig. Du skal sette et kryss på streken for å markere hvor vanskelig du synes at den var. Dersom du synes teksten var lett, setter du krysset ganske nært der det står ”svært enkelt”, dersom du synes den var vanskelig setter du kryss nært der det står ”svært vanskelig”. Synes du teksten var middels vanskelig, setter du et kryss på midten. (demonstrer på tavla). Du får også tilsvarende spørsmål om hvor lett eller vanskelig du synes spørsmålene var, da gjør du på samme måte.*

#### Gjentas for alle:

*Når dere svarer på spørsmålene er det svært viktig at dere svarer med hele setninger. Svaret alene skal være en eller flere meningsfulle setninger, svaret skal kunne forstås at noen som ikke har lest teksten. Det kan være mer enn ett riktig svar – husk det.*

*Når dere er ferdig med å svare på spørsmålene legger dere ned blyanten og venter til jeg gir beskjed. Ikke bla videre. Nå kan dere begynne!*

### **Inferenstekster med støtte**

*Dere skal nå få lese tekster som er delt i to. Først skal dere lese nøye igjennom første avsnitt. Så skal dere svare på to spørsmål til det første avsnittet. Svar med hele setninger, husk at det kan være mer enn ett riktig svar. Så blar du over til neste side og leser resten av teksten nøye. Så blar du over til neste side hvor det er spørsmål til teksten, da bretter sidene du har lest under resten av arkene – slik (demonstrerer). Når dere først har bladd over, er det ikke lov å bla tilbake. Så skal dere svare på spørsmål til teksten.*

#### Dersom første uten støtte er første sett:

*Første spørsmål handler om hvor lett eller vanskelig du synes at teksten var. Under spørsmålet er det en strek. På den ene siden av streken står det svært enkelt, på den andre siden står det svært vanskelig. Du skal sette et kryss på streken for å markere hvor vanskelig du synes at den var. Dersom du synes teksten var lett, setter du krysset ganske nært der det står ”svært enkelt”, dersom du synes den var vanskelig setter du kryss nært der det står ”svært vanskelig”. Synes du teksten var middels vanskelig, setter du et kryss på midten. (demonstrer på tavla). Du får også tilsvarende spørsmål om hvor lett eller vanskelig du synes spørsmålene var, da gjør du på samme måte.*

#### Gjentas for alle:

*Når dere svarer på spørsmålene er det svært viktig at dere svarer med hele setninger. Svaret alene skal være en eller flere meningsfulle setninger, svaret skal kunne forstås at noen som ikke har lest teksten. Det kan være mer enn ett riktig svar – husk det.*

*Når dere er ferdig med å svare på spørsmålene legger dere ned blyanten og venter til jeg gir beskjed. Ikke bla videre. Nå kan dere begynne!*

## **Appendiks 5: Inferensprøve (Kapittel 7)**





---

# Leseprøve 1

**Øvingsoppgave:**

Jon var veldig flink med skolearbeid. Alle i klassen til Jon spilte fotball om ettermiddagen.

Jon hadde svært få venner.

**Spørsmål 1:**

a) Hvorfor hadde Jon så få venner?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

## **Jordskjelv**

Jordskjelv skyldes bevegelser i jordskorpen. Jordskjelv kan føre til at hus faller sammen, da kan mange mennesker bli skadet eller drept. Derfor er det viktig å sørge for at bygninger som bygges tåler jordskjelv. Selv om bygningene er solide, kan det være farlig å være innendørs på grunn av tunge møbler og løse gjenstander. Et skjelv kan for eksempel føre til at bokhyller fulle av bøker raser over ende.

To nesten helt like jordskjelv rammet Armenia i 1988 og San Francisco i 1989. I Armenia ble svært mange mennesker drept, mens i San Francisco var det heldigvis få mennesker som mistet livet.

I land hvor det ofte er jordskjelv, bør folk gjøre rommene i hjemmene sine sikre i tilfelle jordskjelv oppstår.

### **Spørsmål 1:**

a) Hvordan kan folk sikre rommene i hjemmene sine i tilfelle jordskjelv oppstår?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

**Spørsmål 2:**

a) Hva var grunnen til at så mange flere mennesker døde i forbindelse med jordskjelvet i Armenia enn i San Francisco?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 2 over eller ved siden av ordene du har streket under.

## **Norske utvandrere**

På 1800 og 1900 tallet utvandret mer enn 800 000 nordmenn til Amerika. Den første skuten som fraktet norske utvandrere heiste seil i 1825. Overfarten til Amerika tok opptil flere måneder, og kunne være en påkjenning for passasjerene. Mangel på vann begrenset mulighetene for renhold, toalettene om bord på seilskutene var særlig urenlige. Passasjerene sov svært tett, og lopper og lus spredde seg raskt fra køye til køye.

Det var vanlig at utvandrerne hadde med mat selv, de kunne ikke ta med hvilken som helst mat på reisen. Grønnsaker og melk spiste de lite av. Stort sett levde de av saltet eller tørket mat som for eksempel spekekjøtt og flatbrød.

Utvandrerne fryktet storm og dårlig vær, skjønt de ville helst også unngå lange perioder med stille vær. Uvær gjorde selvfølgelig reisen ubehaglig, men det var svært få som døde som følge av at skip havarerte og sank i havet. Likevel var det mange utvandrere som døde før de nådde Amerika.

### **Spørsmål 1:**

a) Hvorfor spiste utvandrerne lite grønnsaker og melk på turen?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

**Spørsmål 2:**

a) Hvorfor ville utvandrerne unngå stille vær?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 2 over eller ved siden av ordene du har streket under.

**Spørsmål 3:**

a) Hva tror du var årsaken til at mange utvandrere døde før de nådde Amerika?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 3 over eller ved siden av ordene du har streket under.

## **Søvn**

Visste du at vi mennesker vokser mens vi sover? De fleste voksne sover mellom 6 og 9 timer i døgnet. Barn sover vanligvis rundt 12 timer, mens babyer trenger minst 18 timers søvn i døgnet.

Jordens 24-timers rytme er innebygd i både mennesker og dyr som en slags biologisk klokke. Mennesker og dyr har vendt seg til å sove i den delen av døgnet hvor det er mest vanskelig å finne mat. Dyr som er kjøttetere trenger mer søvn enn de som bare spiser planter. De fleste dyr sover om natten når det er mørkt, men noen dyr, som for eksempel flaggermus, sover om dagen. Sjiraffer trenger bare 20 til 25 minutters søvn hvert døgn, mens dyr som tiger og løve sover mye mer.

### **Spørsmål 1:**

a) Hvorfor trenger babyer mer søvn enn barn og voksne?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

**Spørsmål 2:**

a) Hvorfor er sjiraffers og løvers søvnvaner så forskjellige?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 2 over eller ved siden av ordene du har streket under.

**Spørsmål 3:**

a) Hvorfor sover flaggermusen om dagen?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 3 over eller ved siden av ordene du har streket under.



## **Fugler**

En typisk fugl har en liten og lett kropp, det er helt nødvendig for at den skal klare å fly. Det som gjør at fuglen kan sveve i luften uten å bevege vingene, er at den har vinger som krummer oppover. Lufta har derfor lengre vei på oversiden enn på undersiden av vingen.

Dette gjør at lufttrykket over vingen blir lavere, og fuglen presses opp i lufta.

De fleste fugler har mange fiender, og i perioder med lite mat er det bare de sterkeste som overlever. Noen fugler finner små insekter i trær eller på bakken, mens andre lever av fisk. Det finnes også fugler som lever av nøtter. Formen på fuglenes nebb er like forskjellig som matvanene. Noen har store, krokete nebb, mens andre har tynne, spisse nebb.

Fugler føder ikke levende unger, de legger egg i stedet. Det ville ha vært helt umulig for fuglen å klare seg dersom den hadde vært gravid med barn i magen slik som oss mennesker.

Fugler overlever ikke så lenge ute i naturen, mange småfugler dør før de er ett år gamle. I fangenskap derimot, finnes det eksempler på fugler som er blitt over seksti år.

### **Spørsmål 1:**

a) Hvorfor kan lufttrykket presse fuglen opp i lufta?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

**Spørsmål 2:**

a) Hvorfor har noen fugler store og krokete nebb, mens andre har spisse og tynne nebb?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 2 over eller ved siden av ordene du har streket under.

**Spørsmål 3:**

a) Hvorfor ville ikke fuglen klare seg dersom den hadde vært gravid med barn i magen slik som oss mennesker?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 3 over eller ved siden av ordene du har streket under.

**Spørsmål 4:**

a) Hvorfor kan fugler i fangenskap bli eldre enn fugler som lever ute i naturen?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 4 over eller ved siden av ordene du har streket under.

## **I trafikken**

Motorveier er veier som er laget for at man skal kunne kjøre trygt i høy fart. På motorveier gjelder spesielle regler. Det er for eksempel ikke lov å snu eller rygge.

Lyskryss regulerer trafikken av lyssignaler. Rødt lys betyr at du må stoppe og vente, gult betyr at du må gjøre deg klar. Når det blir grønt lys, kan du kjøre.

I rundkjøringer kjører man rundt en sirkel i veibanen. Man må selv se etter så man ikke kjører på noen, og man må tilpasse farten i forhold til andre biler. I rundkjøringene er det plass til flere biler enn i lyskryss, og de er derfor mer effektive.. Ulykker blir som regel mindre alvorlige i rundkjøringer enn i lyskryss.

Hvert år er det mange alvorlige bilulykker. Likevel er svært få mennesker redde for å kjøre eller sitte på i bil. Langt flere er redde for å ta fly. Fly har mange flere passasjerer enn biler, men sjansen for å bli rammet av en flyulykke er veldig liten sammenlignet med å bli rammet av en bilulykke. Flyulykker får mer oppmerksomhet i aviser og på tv enn bilulykker.

### **Spørsmål 1:**

a) Hvorfor er det ikke lov å snu eller rygge på motorveier?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.

**Spørsmål 2:**

a) Hvorfor har rundkjøringene plass til flere biler enn lyskryss?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 2 over eller ved siden av ordene du har streket under.

**Spørsmål 3:**

a) Hvorfor blir ulykker som regel mindre alvorlige i rundkjøringer enn i lyskryss?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 3 over eller ved siden av ordene du har streket under.

**Spørsmål 4:**

a) Hvorfor får flyulykker mer oppmerksomhet i aviser og tv enn bilulykker?

b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 4 over eller ved siden av ordene du har streket under.

## **Appendiks 6: Instruksjon til inferensprøve (kapittel 7)**



## Instruksjon til inferensprøven

*Sjekk at det står riktig navn på arkene dere har fått utdelt. Riv av hjørne på arket der navnet deres står.*

*Dere skal nå lese til sammen fem tekster og svare på spørsmål til hver av tekstene. Hvert spørsmål består av en a – oppgave og en b – oppgave.*

*A – oppgaven er spørsmål om innholdet i teksten. Dere skal svare på spørsmålene ut i fra det dere har lest i teksten, ikke ut i fra det dere kan fra før. Det er lov å se på teksten når dere svarer på spørsmålene.*

*Vi begynner med en øvingsoppgave:*

*”Jon var veldig flink med skolearbeid<sup>1</sup>. Alle klassekameratene til Jon spilte  fotball<sup>1</sup> om ettermiddagen. Jon hadde svært få venner.”*

*”a) Hvorfor hadde Jon så få venner?”*

*Det riktige svaret er at han hadde få venner fordi han leste lekser mens de andre spilte fotball.*

*Det er jo også mulig at Jon er en ufyselig type, eller at klassekameratene er så slemme at de ikke lar Jon få være sammen med dem. Det kan også være mange andre svar på spørsmålet, men dere skal svare ut i fra hva som står i teksten – selv om svaret ikke står direkte i teksten.*

*I b – oppgavene skal dere streke under to ord i teksten som hjalp dere med å finne svar på spørsmål a. I tillegg skal dere skrive ett tall over eller ved siden av ordene dere har streket under.*

*”b) Strek under to ord i teksten som hjalp deg med å finne svaret på spørsmålet over. Skriv tallet 1 over eller ved siden av ordene du har streket under.”*

*Hvilke to ord var det som hjalp oss med å finne svar på hvorfor Jon har så få venner? Jo, det var ordene fotball og skolearbeid. Sett en strek under de ordene og sett et 1- tall over eller ved siden av ordet du har streket under. Husk at dere skal streke under i teksten, ikke i spørsmålet.*

*Noen ganger vil det være flere ord som hjelper dere med å finne rett svar, da må dere velge ut to ord som dere streker under. Dersom dere ikke har klart å finne svar på spørsmål a, skal dere likevel streke under ord som dere tror leder til rett svar. Dere får poeng både for å svare på a – spørsmålene og for b – spørsmålene hvor dere streker under ord.*

*Mange av spørsmålene er vanskelige, og dere må kanskje gruble litt for å finne svaret. Pass på at dere ikke blir sittende for lenge på hvert av spørsmålene. Hvis dere ikke finner svar på spørsmålet, setter dere et kryss på det spørsmålet dere ikke klarer og går dere videre til neste spørsmål. Altså: Kryss på oppgaver dere har prøvd å løse, slik at det er tydelig at dere ikke bare har glemt den. Hvis dere har mer tid igjen på slutten, så kan dere gå tilbake til de vanskelige spørsmålene.*

*Spørsmål?*

*Da har dere 40 minutter å jobbe.*

Etter 10 minutter sjekk at alle er ferdige med den første teksten. Dersom noen fremdeles grubler på de første spørsmålene, be de gå videre å heller gå tilbake hvis det blir tid mot slutten.

Si til alle: *Husk at dere skal svare på b- oppgavene selv om dere ikke har klart å finne svar på a – oppgaven.*

Etter 20 minutter sjekk at alle er ferdige med den andre teksten om norske utvandrere. Dersom noen fremdeles grubler på de spørsmålene til den andre teksten, be de gå videre å heller gå tilbake hvis det blir tid mot slutten.

Etter 40 minutter blir heftene samlet inn.



## **Appendiks 7: Ordforrådsprøve (kapittel 7)**



---

Øvingsoppgave

Hus	Bolig	Telt	Vogn
Drosje	Lastebil	Taxi	Buss

1. Flink	Arg	Dyktig	Fornøyd
2. Finnerlønn	Honorar	Gratiale	Dusør
3. Foredrag	Tale	Dialog	Diskusjon
4. Eksamensbevis	Vitnemål	Resultat	Karakter
5. Altan	Karnapp	Pergola	Balkong
6. Konge	President	Monark	Gardist
7. Kansellere	Avvise	Nekte	Avlyse
8. Sen	Snarrådig	Gretten	Langsom
9. Seilskip	Skonnert	Minesveiper	Snekke
10. Fruktgrøt	Fromasj	Sorbé	Kompott
11. Plattform	Fundament	Grunnmur	Hjørnestein
12. Bestilling	Ultimatum	Rekvisisjon	Påbud
13. Samling av stjerner	Asteroide	Galakse	Meteor
14. Enerett	Monopol	Egenrådig	Analogi
15. Valg	Skjebne	Sjanse	Avgjørelse
16. Misdannet	Deformert	Kaputt	Skadet
17. Forakte	Fortvile	Sjenere	Avsky
18. Drøfte	Instruere	Konferere	Irettesette
19. Sverd	Bajonett	Spyd	Sabel
20. Antakelse	Dilemma	Hypotese	Problem
21. Sport	Idrett	Interesse	Underholdning
22. Kategori	Klasse	Kasus	Prototyp
23. Avgjørende	Presserende	Tungtveiende	Utslagsgivende
24. Naiv	Sjenert	Godtroende	Misunnelig

## **Appendiks 8: Instruksjon til ordforrådsprøve (kapittel 7)**



## **Instruksjon til ordforrådsprøven**

Fortell at elevene skal gjennomføre en ordprøve. Hold opp et prøveark og pek på øvingsoppgaven:

*I første rekke finner du et ord. Du skal finne ut hvilket av de tre neste ordene som betyr det samme som det første. Når du har funnet det riktige ordet skal du sette et kryss på det ordet som er riktig. Dersom du krysser av feil ord, setter du bare en ring rundt ordet og krysser av det som er riktig. Dere trenger ikke å bruke viskelær. Vi begynner med en øvingsoppgave:*

*Hus, er det det samme som bolig, telt eller vogn?*

*(La elevene svare.) Hus det er det samme som bolig så da setter du kryss på det første ordet.*

*Drosje, er det det samme som lastebil, taxi eller buss?*

*(La elevene svare.) Drosje og taxi betyr det samme, derfor setter du kryss på ordet i midten.*

*Bla over til neste side, der finner dere samme type oppgaver som dere skal gjøre når jeg sier i fra. Nå skal dere ikke si svaret høyt, bare krysse av på rett ord uten å snakke sammen. Det er bare lov å krysse av på ett ord på hver oppgave. Jeg kommer til å lese ordene høyt for dere. Hvis jeg leser for fort så rekker dere en hånd i været og sier i fra, så skal jeg si ordene en gang til.*

Les ordene høyt i rolig tempo på samme måte som i eksemplene, spør etter et par oppgaver om det går for fort eller sakte. Les hver oppgave en gang, gjenta oppgaven dersom noen ber om det. Når du kommer til item 5 og 15 spør: *Er alle på nummer 5(15) nå?*

Dersom elever kommenterer at de ikke vet svaret, eller at de synes to av alternativene virker like riktige, si at de skal krysse av for det de mener er mest riktig – kun ett kryss er lov per oppgave.





## **Appendiks 9: Eksempel på materiell fra inferenstreningen (kapittel 7)**



## **Sjiraffen**

Sjiraffen kan bli over fem meter høy og er med sin langstrakte kropp verdens høyeste dyr. Den spiser blader fra trær og forsyner seg gjerne av bladene høyt oppe i trekronene. De lange beina og den lange halsen gjør det vanskelig for sjiraffen å drikke. Sjiraffen er mest sårbar når den drikker, da kan den lett bli angrepet av rovdyr.

Sjiraffer finnes bare i Afrika, hvor de er spredt over et stort område. De er et vanlig syn på de åpne slettene sør for Sahara.

Sjiraffer holder sammen i flokker som i noen tilfeller kan bestå av opptil hundre dyr. De går alltid i flokk når de skal ned til et vannhull for å drikke.

Hvorfor går sjiraffene i flokk når de skal drikke?  
Skriv svaret på spørsmålet inn i den tomme ruta i midten.

Sjiraffen er sårbar når den drikker.



Sjiraffer drikker i flokk.

## **Skjørbuk**

Skjørbuk er en sykdom som skyldes mangel på vitamin C. Sykdommen utvikler seg etter mange måneder med ensidig kost. Tegn på skjorbuk er at man begynner å blø i munnen og at man blir lett trett og får dårlig appetitt.

Her i Norge trenger vi ikke lengre å bekymre oss for denne sykdommen siden skjorbuk nå er blitt svært sjelden. Men på 1500 tallet og helt fram til begynnelsen av 1900 tallet var det en vanlig sykdom. Det var spesielt mange sjøfolk som ble rammet av skjorbuk.

**Spørsmål 1:**

Hvorfor var det særlig sjøfolk som ble rammet av skjørbuk?

Skriv svaret på spørsmålet inn i den tomme ruta i midten.

Skjørbuk skyldes mangel på vitamin C og ensidig kost.



Det var spesielt mange sjøfolk som ble rammet av skjørbuk.

**Spørsmål 2:**

Hvorfor er skjorbuk en sjelden sykdom i Norge i dag?  
Skriv svaret på spørsmålet inn i den tomme ruta i midten.

Skjorbuk skyldes mangel på vitamin C og ensidig kost.



Skjorbuk er en sjelden sykdom i Norge i dag.





## **Appendiks 10: Eksempel på leksjonsbeskrivelse til inferenstreningen (kapittel 7)**



### **Leksjon 3**

I denne økten skal elevene jobbe med grafiske modeller hvor informasjonen fra teksten allerede er fylt ut. Elevenes oppgave er å svare på spørsmålet og fylle dette inn i modellen.

#### Sjiraffen

Denne teksten og oppgaven gjennomgås i fellesskap. Gi forklaringer på eventuelle vanskelige ord. Teksten og spørsmålet leses høyt. Forklar elevene at svaret på spørsmålet ikke står direkte i teksten og at modellen eller figuren under skal hjelpe de med å finne rett svar. I to av rutene i figuren står det to opplysninger som er hentet rett fra teksten. Be elevene finne igjen disse to opplysningene i teksten og sette en strek under den setningen hvor opplysningene står. Hjelp elevene med å finne rett svar på spørsmålet (at Sjiraffene bytter på med å holde vakt når de drikker).

#### Skjørbuk

Elevene leser teksten stille hver for seg. Gi forklaring på eventuelle vanskelige ord i teksten. Oppgavene løses i par, eventuelt i små grupper. Be elevene om å finne ut hvor i teksten opplysningene som står i rutene er. Elevene skal sette strek under setningene der hvor opplysningene står i teksten. Deretter skal de finne rett svar på oppgavene Sjekk at alle har fått rett svar på oppgavene (spørsmål 1: Sjøfolkene hadde begrenset tilgang på (fersk) mat om bord. Spørsmål 2: I dag har vi variert kost og god tilgang på vitamin C.) Ta en felles gjennomgang av spørsmålene mot slutten av økten.



## **Appendiks 11: Eksempel på loggskjema (kapittel 7)**



## Loggskjema for leksjon 3

1) Dato for leksjon:

2) Hvor mange av elevene gjennomførte følgende oppgaver? (sett ett kryss i hver rad)

Materiell	Alle	Over halvparten av elevene	Under halvparten av elevene	Ingen
Sjiraffen				
Skjørbuk				

3) Tid brukt på leksjonen: ca \_\_\_\_\_ minutter

4) Elever som ikke var tilstede:

5) Eventuelle kommentarer:





## **Appendiks 12: Observasjonsskjema (kapittel 7)**



## Observasjon – klasse

**Dato:**

**0-5 minutter:**

**Skole/ klasse:**

**Hvilken leksjon/ type oppgaver:**

**Hvor mange elever tilstede?**

**Hvor mange lærere tilstede?**

**Hvor lang tid tar det å få ro i klassen før undervisningen kan begynne?**

**5-10 minutter:**

**Antall tilfeller av elever som avbryter lærer:**

**Antall tilfeller av elever som avbryter andre elever (avbrudd/ forstyrrelser som ikke angår oppgavene):**

**Hvor mange elever snakker samtidig som lærer gir felles beskjeder/ instruksjon:**

**10-15 minutter:**

**Hvor mye snakker hvem?**

**Lærer:**

**Elever:**

**15-20 minutter:**

**I hvilken grad er det snakk/ støy som ikke handler om oppgavene**

**Stor grad – middels grad – liten grad**

**Generelt:**

**I hvilken grad virker det som om elevene generelt synes arbeidet med tekstene og oppgavene lystbetont?**

**Liten grad – middels grad – stor grad**

**Inferenstreningen:**

**Hvor mange elever ble ferdig med alle tekstene/ oppgavene?**

**Hvor mye tid ble brukt til å gjennomføre leksjonen?**

**Spørsmål til læreren etter timen: var dette en representativ time (med hensyn til arbeidsro spesielt)?**

**Kommentarer:**