

# Vokabularets betydning for leseforståelsen

Hvilke sammenhenger er det mellom vokabular og  
leseforståelse blant 8. klassinger?

**Ragnhild Hustavenes og Lisbeth Svelstad**



Masteroppgave i spesialpedagogikk  
Det utdanningsvitenskapelige fakultet  
Institutt for spesialpedagogikk

**UNIVERSITETET I OSLO**

Juni 2010

## Sammendrag

Studien omhandler vokabularet betydning for leseforståelsen. Oppgaven er knyttet opp mot forskningsprosjektet Kunnskapsgenerering i det Spesialpedagogiske Praksisfeltet (KiSP – prosjektet) ved Institutt for Spesialpedagogikk ved Universitetet i Oslo. Prosjektet er longitudinelt og vil pågå over tre år. Målet med KiSP - prosjektet er å undersøke effekten av to metodisk forskjellige intervensjonsprogrammer i vokabularlæring.

Flere undersøkelser viser en klar sammenheng mellom tidlig vokabularinnlæring og dens positive prediksjon på senere leseforståelse. Ofte ser man i den tidligere forskningen at vokabular deles inn i bredde og dybde. Med vokabularbredde menes den mengde ord som en person kjenner og klarer å hente frem fra hukommelsen, når han/hun har bruk for dem, uten nødvendigvis å ha full forståelse for betydningen av alle ordene man bruker. Med vokabulardybde menes den mengde ord en person har klar forståelse for i et gitt språk (Bø og Helle, 2008). Enkelte undersøkelser har vist at breddevokabular har den største innvirkningen på leseforståelse (Tannenbaum, Torgesen & Wagner, 2005), mens andre mener at dybdevokabular har større effekt på leseforståelsen (Ouellette, 2006; Roth, Speece & Cooper, 2002; Snow, Tabors, Nicholson og Kurland, 1995). Andre undersøkelser igjen, deler ikke vokabular inn i bredde og dybde, men undersøker leseforståelsen opp mot vokabular generelt, ordavkodingsferdigheter og lesekunnskap (Mezynski, 1983; Petersen, 2008).

Denne oppgaven undersøker betydningen av vokabular knyttet opp mot en eventuell effekt på lesing hos elever i 8. klasse. Også i denne undersøkelsen deles vokabular inn i bredde – og dybdevokabular, grunnen til dette er for å finne ut om noen av disse i mer eller mindre grad har ulik innvirkning på leseforståelsen. Med utgangspunkt i dette ble følgende problemstilling formulert:

*Hvilke sammenhenger er det mellom vokabular og leseforståelse blant 8. klassinger?*

Med utgangspunkt i dette ble det videre formulert to ulike forskningsspørsmål:

*Hvilken innvirkning har vokabularbredde på leseforståelsen blant 8. klassinger?*

*Hvilken innvirkning har vokabulardybde på leseforståelsen blant 8. klassinger?*

Problemstillingen er belyst med en kvantitativ metodisk tilnærming og et deskriptivt ikke-eksperimentelt design. Dette vil si at undersøkelsen kartlegger og beskriver forhold mellom

vokabular og leseforståelse målt hos en gruppe elever på ett tidspunkt. Elevene i undersøkelsen er valgt ut fra prosjektgruppen i hovedprosjektet. Utvalget består av 112 elever i 8. klasse fra to ulike skoler i Oslo, alle elevene har norsk som morsmål. Ni mastergradsstudenter har deltatt i testingen av elevene som deltar i KiSP – prosjektet, og i forhold til arbeid med masteroppgaven har studentene fått tilgang til datamaterialet.

For å kartlegge vokabularbredde er British Picture Vocabulary Scale (BPVS II) (Dunn, Dunn, Whetton og Burley, 1997) benyttet. Vokabulardybde er målt med deltesten Ordforståelse fra Wechsler Intelligence Scale for Children (WISC III) (Wechsler, 1991). Nasjonale prøver i lesing ble brukt som mål på leseforståelse. For å undersøke om det er noen bakenforliggende variabler som er med og forklarer en eventuell sammenheng mellom vokabular og leseforståelse, benyttes testene Ordkjedetesten (Høien og Tønnesen, 2007) og Raven Standard Progressive Matriser (Raven) (Raven, 1986). De to sistnevnte testene måler henholdsvis avkodingsferdigheter og generell kognitiv fungering.

Det er benyttet både deskriptiv og analytisk statistikk i dataanalysen. Relevante analyser har vært korrelasjons- og regresjonsanalyser.

Resultater i denne undersøkelsen viser at vokabulartestene BPVS II og deltesten Ordforståelse i WISC III har en sterk korrelasjon (henholdsvis .46 og .45) med leseforståelse når den måles med Nasjonale prøver i lesing, og begge er satt til et signifikansnivå på 1 %.

Regresjonsanalysen viser at det er deltesten Ordforståelse i WISC III, som måler dybdevokabular, som har den største unike forklaringsvariansen på 3,2 %, sammenlignet med BPVS II, som måler breddevokabular, som har en unik forklaringsvariens på 2,1 % etter at det er kontrollert for Ordkjedetesten og Raven. Resultatene viser et signifikant forhold mellom leseforståelse og vokabulardybde med et signifikansnivå satt til 1 %. Når det gjelder forholdet mellom leseforståelse og breddevokabular er resultatet nær et signifikansnivå på 5 %.

## Forord

Endelig er vi i mål med masteroppgaven. Prosessen som nå avsluttes har vært spennende, lærerik og til tider krevende.

Stor takk til vår veileder Gunvor Dalby Vea, for nyttige og konstruktive tilbakemeldinger under hele prosessen. Vi må også få takke hverandre for gode diskusjoner og samarbeid. En takk også til Kristin og Kari som har tatt seg tid til å lese korrektur.

Vi vil også takke forskergruppen i prosjektet Kunnskapsgenerering i det Spesialpedagogiske Praksisfeltet. Det har vært veldig lærerikt å kunne delta i dette prosjektet.

Til slutt vil vi takke familie og venner for støtte og oppmuntring underveis!

Oslo, 25. mai 2010

Ragnhild Hustavenes og Lisbeth Svelstad

# Innhold

<b>Sammendrag</b> .....	2
<b>Forord</b> .....	4
<b>Innhold</b> .....	<b>5</b>
<b>1 Innledning</b> .....	<b>8</b>
1.1 Bakgrunn, tema og formål .....	8
1.2 Oppgavens struktur .....	10
<b>2 Teoretisk og empirisk bakgrunn</b> .....	<b>11</b>
2.1 Vokabular .....	11
2.1.1 <i>Bredde – og dybdevokabular</i> .....	14
2.2 Leseforståelse .....	15
2.2.1 <i>Språkforståelse</i> .....	16
2.2.2 <i>Forkunnskaper</i> .....	17
2.2.3 <i>Å danne inferenser</i> .....	18
2.2.4 <i>Sjangerkjennskap og tekststruktur</i> .....	18
2.2.5 <i>Kognitive evner</i> .....	19
2.2.6 <i>Leseforståelsesstrategier og metakognisjon</i> .....	20
2.3 Sammenhengen mellom vokabular og leseforståelse .....	21
2.4 Tidligere empiriske undersøkelser .....	24
2.4.1 <i>Sammenhengen mellom bredde- og dybdevokabular og leseforståelse</i> .....	24
2.4.2 <i>Sammenhengen mellom generelle kognitive evner og leseforståelse</i> .....	27
2.5 Problemstillinger .....	29
<b>3 Metode</b> .....	<b>30</b>
3.1 Design .....	30
3.2 Utvalg .....	31
3.3 Gjennomføring .....	32
3.4 Instrumenter .....	33
3.4.1 <i>Nasjonale prøver i lesing</i> .....	33

3.4.2	<i>Deltesten Ordforståelse i WISC - III</i> .....	33
3.4.3	<i>British Picture Vocabulary Scale II</i> .....	34
3.4.4	<i>Ordkjedetesten</i> .....	34
3.4.5	<i>Raven</i> .....	35
3.5	Analyse.....	35
3.6	Validitet.....	37
3.6.1	<i>Statistisk validitet</i> .....	37
3.6.2	<i>Indre validitet</i> .....	38
3.6.3	<i>Begrepsvaliditet</i> .....	39
3.6.4	<i>Ytre validitet</i> .....	40
3.7	Reliabilitet .....	40
3.8	Etiske betraktninger.....	41
<b>4</b>	<b>Resultater</b> .....	<b>42</b>
4.1	Deskriptiv statistikk.....	42
4.2	Reliabilitetsanalyse.....	44
4.3	Korrelasjoner.....	45
4.4	Regresjonsanalyse .....	46
4.5	Oppsummering av resultater .....	48
<b>5</b>	<b>Drøfting</b> .....	<b>50</b>
5.1	Hvilken innvirkning har vokabulardybde på leseforståelsen blant 8. klassinger?.....	50
5.2	Hvilken innvirkning har vokabularbredde på leseforståelsen blant 8. klassinger? .....	53
5.3	Leseforståelse og generelle kognitive evner.....	55
5.4	Leseforståelse, vokabular og ordavkoding .....	56
5.5	Resultatene drøftet i lys av validitet .....	58
5.5.1	<i>Statistisk validitet</i> .....	58
5.5.2	<i>Reliabilitet</i> .....	60
5.5.3	<i>Indre validitet</i> .....	61
5.5.4	<i>Begrepsvaliditet</i> .....	63
5.5.5	<i>Ytre validitet</i> .....	65

5.6 Oppsummering .....	66
<b>Kildeliste.....</b>	<b>69</b>

### **Tabelloversikt**

Tabell 1 Deskriptiv statistikk.....	43
Tabell 2 Korrelasjonsanalyse.....	45
Tabell 3 Regresjonsanalyse.....	47

# 1 Innledning

## 1.1 Bakgrunn, tema og formål

Denne undersøkelsen vil bli gjennomført i tilknytning til forskningsprosjektet Kunnskapsgenerering i det Spesialpedagogiske Praksisfeltet (KiSP). Prosjektet vil undersøke effekten av to metodiske forskjellige intervensjonsprogrammer i vokabularlæring, og er i regi av Institutt for Spesialpedagogikk (ISP) ved Universitetet i Oslo (UiO). I forbindelse med denne masteroppgaven vil vi først og fremst undersøke betydningen av vokabular knyttet opp mot en eventuell effekt på lesing hos elever i 8. klasse.

Det er forsket mye på vokabular og leseforståelse. Enkelte undersøkelser har vist at dybdevokabular har større effekt på leseforståelsen (Ouellette, 2006; Roth, Speece og Cooper, 2002; Snow, Tabors, Nicholson og Kurland, 1995), mens andre mener at breddevokabular har den største innvirkningen på leseforståelse (Tannenbaum, Torgesen og Wagner, 2005). Det finnes også undersøkelser som ikke legger vekt på å dele vokabular inn i bredde og dybde. Disse undersøker leseforståelsen opp mot vokabular generelt, ordavkodingsferdigheter og kunnskaper i lesing (Mezynski, 1983; Petersen, 2008).

Temaet er interessant både fra et empirisk og et teoretisk perspektiv. I et empirisk perspektiv ser man at tidligere forskning gir varierende resultater angående om det er bredde – eller dybdevokabular som har størst effekt på leseforståelsen. I et teoretisk perspektiv er temaet interessant for å se om vokabularinnlæring har en positiv innvirkning på senere leseforståelse. Mesteparten av teorien tilsier dette, derfor kan det være spennende å finne ut om resultatene i denne undersøkelsen samsvarer med teorien.

I denne studien deles vokabular inn i bredde og dybde. Både når det gjelder teori og tidligere forskning deles ofte vokabular inn i disse to begrepene. Det ble derfor interessant for denne undersøkelsen å gjøre det samme, for å finne ut om det er noe forskjell på breddevokabularet og dybdevokabularet innvirkning på leseforståelsen hos 8. klassinger. I en studie der både vokabularmessige forhold og forhold angående lesing kartlegges samtidig, vil det ikke være mulig å avdekke hvordan ulike ferdigheter påvirker hverandre. Men det vil være mulig å si noe om graden av sammenheng mellom ulike variabler. Forholdet mellom vokabular og



leseforståelse kan imidlertid drøftes ut fra et teoretisk perspektiv og i forhold til hvordan funn som gjøres passer inn i et slikt perspektiv.

Formålet med denne undersøkelsen er å få en dypere kunnskap om forholdet mellom vokabular og leseforståelse hos norske 8. klassinger. Utvikling av leseforståelse henger tett sammen med omfanget av barnas vokabular, og selv om de har knekt ordavkodingen vil den videre leseutviklingen lide hvis barnet ikke har et ordforråd som kan støtte forståelsen. Selv om et dårlig vokabular kan påvirke ordavkodingen, vil barn som møter skolen med et dårlig utviklet ordforråd først og fremst få vansker med å utvikle en god leseforståelse. Lesing er en god kilde for vokabularutviklingen, men hvis vokabularet i utgangspunkt er dårlig kan skoletekster fort bli for vanskelige. Skolespråket er avgjørende for å forstå fagbøker. Noe av dette ordforrådet må elevene lære gjennom undervisningen. Det er viktig at læreren forsikrer seg om at elevene har det ordforrådet som trengs for å forstå en fagtekst. Det gir det nødvendige grunnlaget for at videre læring skal kunne finne sted (Frost 2003; Frost, 1999). Kort oppsummert kan man si at det som kjennetegner gode lesere er at de har et godt ordforråd, de kan "gå inn" i nye ord og forstå dem mens de leser, de kan finne deler av ord som støtter forståelsen av ordet, de kan bruke sammenhengen i teksten som støtte i ordforståelsen, de har et godt utviklet skolespråklig ordforråd og de er bevisste på når et ord ikke er forstått.

Forståelse av forholdet mellom vokabular og leseforståelse kan være nyttig gjennom et barns leseutviklingsprosess. Barn med dårlig vokabular kan være i fare for å få leseforståelsvansker, som igjen kan føre til innvirkning på deres senere skolegang. Et dårlig vokabular kan påvirke ulike forhold ved leseutvikling, som avkoding, leseflyt og leseforståelse. Ved å skille mellom bredde – og dybdevokabular kan man få et innblikk i hva som er viktig å legge vekt på i vokabularinnlæringen. Ut fra teorien kan man anta at kvantiteten på vokabularet spiller størst rolle for leseforståelsen i de tidligste skoleårene, mens ettersom barnet blir eldre og tekstene øker i vanskelighetsgrad vil kvaliteten på vokabularet være viktigere for leseforståelsen. Kunnskap om tema kan også være viktig for å sette inn hensiktsmessig intervensjon. På denne måten forebygges senere leseforståelsvansker gjennom en tidlig vektlegging av en grunnleggende vokabularinnlæring, der både bredden og dybden blir viktig.

## 1.2 Oppgavens struktur

I kapittel 2 presenteres teori og sentrale empiriske funn. Innledningsvis gjøres det rede for begrepene vokabular og leseforståelse. Deretter beskrives sammenhengen mellom disse to begrepene og deres gjensidige påvirkning på hverandre. Empirien beskriver tidligere forskning innenfor dette feltet. Problemstillingen og de to forskningsspørsmålene presenteres mot slutten av kapittel 2, på bakgrunn av den foregående framstillingen av teori og empiri.

I kapittel 3 blir det redegjort for den metodiske tilnærmingen. Avslutningsvis beskrives sentral validitetsteori før etiske betraktninger blir kommentert.

I kapittel 4 presenteres undersøkelsens resultater. Først blir de deskriptive dataene beskrevet. Deretter redegjøres det for de analytiske resultatene ved hjelp av korrelasjons – og regresjonsanalyse.

I kapittel 5 diskuteres undersøkelsens resultater i forhold til tidligere presentert teori og empiri. Deretter drøftes resultatene i lys av validitetsspørsmål. Avslutningsvis oppsummeres resultatene og implikasjoner for eventuell videre forskning beskrives.

## 2 Teoretisk og empirisk bakgrunn

Først i dette kapitlet avklares begrepene vokabular og leseforståelse ut fra teoretiske rammer. Deretter diskuteres sammenhengen mellom disse to begrepene og deres gjensidige påvirkning på hverandre. Til slutt vises det til tidligere forskning som er gjort innenfor dette feltet.

### 2.1 Vokabular

Ord er så viktige for menneskene at vi ofte glemmer å reflektere over verdien og makten som ligger i begrepet "ord". Vokabularet vårt, kan oftere enn dialekt, avsløre vår sosiale bakgrunn og vår utdannelsesbakgrunn. Ord er et verktøy og sentrale kognitive språkenheter vi bruker for å uttrykke vår bakgrunnskunnskap, ideer og for å lære nye ord (Drum og Konopak, 1987). Ord kan ha flere betydninger. Ords betydning varierer i ulike kontekster, men kjernebetydningen er motstandsdyktig til kontekstuell variasjon. Ord er en konvensjon formet i et gitt språk i en gitt kultur. Dette gjør at ord kan brukes som kommunikasjonsmiddel innenfor en gitt språkkultur (ibid).

Bishop (1997) definerer vokabular som de mentale representasjoner av ord som inneholder informasjon om både den fonologiske formen av ord og representasjon av deres mening. Vokabularet blir så lagret som sekvenser av familier representert i minnet eller "leksikon", som språklyder og disse assosieres så med spesifikke meninger. Fonologiske og semantiske representasjoner av ord er dermed de byggesteinene som danner en persons vokabular. Ut fra definisjonen til Bishop (1997) kan man si at ord er et fenomen som har minst to sider; en fonologisk form og et semantisk innhold. Ordets form viser seg å ha mer enn bare en fonologisk form, den har også en morfologisk form. Den morfologiske formen består av minste meningsbærende, men ikke selvstendige språkenhet - morfemer (Rommetveit, 1972), som er faste kombinasjoner av språklyder, og disse kombinasjonene har en eller flere bestemte betydninger.

Vokabular er et teoretisk begrep som kan defineres på ulike nivå. Vokabular kan sees på som et sett av ord som er basis i byggingen og forståelsen av en setning (Anglin, 1993). Vokabular er det ordtilfanget som man kan forestille seg er organisert i leksikon i form av ordets fonologiske form, forskjellig fra, men likevel i kontakt med de semantiske representasjonene

av ordmening (Coleman, 1998). En kan da trekke et skille mellom nummer av leksikalske enheter (breddevokabular) og omfanget av semantiske representasjoner (dybdevokabular) (Ouellette, 2006).

Vokabulartilegnelse består av to aspekter: lagring og raffinering av fonologiske representasjoner, og lagring og elaborering av assosierte semantiske kunnskaper (Ouellette, 2006). Ved lagring og raffinering av fonologiske representasjoner er talepersepsjon og relaterte ferdigheter, samt det fonologiske kortidsminnet av betydning (Gathercole og Baddeley, 1993). Ved tilegnelsen av ordbetydninger, må man bearbeide tilgjengelig informasjon for å finne ut hva ordet betyr. Så lagres betydningen i minnet og etter hvert videreutvikles og nyanseres betydningen (Ouellette, 2006).

Men hvilke faktorer har betydning for utviklingen av vokabular? Anglin (1993) mener at barn har to typer komplekse evner de bruker for å tilegne seg vokabular. Den første er at barna lærer ord ved å trekke slutninger om ordets betydning ut i fra muntlige og skriftlige kontekster. På denne måten utvikler de representasjoner for ordet og dets betydning i sitt langtidsminne. Den andre evnen er å analysere ordstrukturer og på den måten avlede deres betydning.

Nagy og Herman (1987) antyder at det er muntlig kontekst som er hovedkilden for tilegnelse av vokabular, spesielt i førskolealder. Denne konteksten er rik på slikt som tilstedeværelse av selve objektet/fenomenet, intonasjon og gester. Dette er med på å støtte forståelse og inferens. Sternberg (1987) mener også at kontekst er hovedkilden til øking av vokabularkunnskaper. Gjennom sammenhengen mellom kontekst og kontekstuell læring får man evnen til å trekke slutninger eller inferens. En annen faktor som har betydning for vokabulartilegnelse er arbeidsminnet. Arbeidsminnets kapasitet bestemmer hvor effektivt individer bruker kontekst for å forstå og produsere ord (Daneman og Green, 1986). Bates og Goodman (1997) hevder det er en sterk sammenheng mellom vokabularstørrelse og grammatisk kompetanse. De fremhever at vokabularutviklingen og grammatikkutviklingen aldri distanserer seg fra hverandre. Dette medfører at grammatikk også er en viktig faktor for utvikling av vokabular. Tekst er også en rik kilde til vokabulartilegnelse, ved å utvikle gode ferdigheter i lesing er dette en av de mest effektive strategiene i tilegnelsen av vokabular (Baker, Simmons & Kameenui, 1998). Stanovich (1986) mener at ved å utvikle gode ferdigheter i lesing er dette med på å øke vokabularutviklingen som igjen er med på å bygge opp lesingen. Denne

vekselvirkningen mellom vokabular og lesing vil fortsette i uforminsket styrke å utvikle seg gjennom og utover barndommen (Bishop, 1997).

Jo flere ord vi har, desto lettere lærer vi nye ord. Ouellette (2006) nevner i denne sammenhengen inferens; at de elevene som klarer å lese mellom linjene klarer ved dette å komme frem til betydningen av nye ord. Undersøkelser viser også at elever med et fattig ordforråd utvikler mindre ord enn de elevene som har et rikt ordforråd. Det blir en slags Mattheus effekt. I løpet av et år vil en elev lære seg ca 3 000 ord. Gode elever kan lære seg opp til 5 000 ord per år, mens en med et dårlig vokabular kanskje ikke lærer mer en 1 000 ord. Gjennom vokabular øving i skolen vil ikke elevene lære mer enn 300 ord. Den største læringen av nye ord skjer i det daglige (Carlisle og Rice, 2002). Kontekst er viktig i læring av nye ord. Forskning viser at etter fjerde klasse skjer det meste av innlæringen av nye ord gjennom lesing (ibid.).

Stahl og Nagy (2006) har laget en pyramide for utvikling av vokabular. Pyramiden er tenkt som en læringsveileder i læring av vokabular. Denne modellen er tatt med i denne sammenhengen fordi den illustrer også utviklingen i vokabulartilegnelse. Modellen sier noe om hvordan vokabularet utvikler seg fra bredde til dybde.

Pyramiden er delt inn i tre nivåer, hvor variasjon og mengde er viktige faktorer. Det er også viktig hvor mye innsats og tid en bruker på innlæringen på de ulike nivåene. Pyramiden er organisert slik at den går fra breddevokabular til dybdevokabular. På bunnen av pyramiden er det innlæring av mange høyfrekvente ord. Lenger opp i pyramiden er antall ord mindre, men elevene skal få økt kunnskap om hvert enkelt ord. Alle tre nivåene er like viktige. Nedenfor presenteres de ulike nivåene.

Nivå I er på bunnen av pyramiden og er det bredeste nivået. Et viktig fokus på dette nivået er at innlæring av ord ikke skal være hovedaktiviteten, men at denne prosessen skal komme gjennom daglige gjøremål. Det kan for eksempel være gjennom samtaler eller lesing. Nivå I skiller mellom muntlig og skriftlig språk i innlæring av vokabular. For de yngste barna vil muntlig språk være den viktigste kilden for vokabular, men for eldre barn vil etter hvert skriftspråket bli den største kilden til vokabulartilegnelse. Men for de barna som strever med lesing og har dårlige avkodingsferdigheter er ikke skriftspråket en like sterk kilde. Biemiller (1999) gjorde en sterk sak ut av at så lenge barna strevde med avkodingsferdighetene sine har de fortsatt behov for å bli lest til. Så lenge barna ikke har gode nok avkodingsstrategier vil

utviklingen av vokabular skje gjennom den muntlige bruken av språket. Etter hvert som barna blir gode lesere vil de lære mest gjennom å lese selv (Stahl og Nagy, 2006). Nivå I er det nivået som dekker det meste av innlæringen av høyfrekvente ord som vi trenger i den daglige tale. Dette nivået skal gjøre barna nysgjerrige på å lære nye ord.

Nivå II er det nivået som skal gi elevene økt kunnskap om det spesifikke ord. Det vil si at elevene skal lære ordene godt, slik at de blir trygge på å bruke ordet på sin egen måte og forstå ordet i ulike sammenhenger. Det som er viktig på dette nivået er at læringen skjer i små steg og ikke alt læres samtidig. Det er ulike måter å tilegne seg kunnskap om et ord på; elevene kan bruke oppslagsverk eller at de kan bli fortalt hva ordets betydning er. Andre ganger kan en forstå ordet gjennom konteksten i en tekst eller sammenhengen i en setning. Det kan ikke forventes at elevene skal kunne bruke ordet korrekt med det første. Etter hvert som de blir kjent med ordet vil nok bruken av ordet bli mer korrekt. For at elevene skal få en økt kunnskap om ord er det viktig at de lærer seg ulike strategier i forhold til hvordan de skal få til dette og videreutvikle vokabularferdighetene sine.

På nivå III skal elevene føres til det nivået som skal gi den komplette forståelse av ordet. Elevene skal kunne bruke ordet skriftlig, og forstå ordet automatisk i lesing av en tekst. Det er mange viktige ord elever skulle ha lært, noe som vil ta mye tid fra undervisningen. I denne sammenhengen mener Stahl og Nagy (2006) det er to faktorer som skal være avgjørende i valg av utvidelsen av ordforståelsen: viktigheten av et ord og anvendeligheten til et ord. I valg av viktigheten av et ord, kommer det for eksempel an på hva som er det sentrale i en tekst som skal gjennomgås i timen. Det er viktig at denne utvelgelsen er gjennomtenkt slik at det ikke blir for mange ord, men at de ordene som blir valgt er sentrale. Det andre som er viktig er anvendbarheten til ordet. Dette kan vi dele inn i to dimensjoner; det første er hvor brukbart dette ordet er for elevene; det bør være ord som elevene kan bruke i andre sammenhenger også. Det andre som er viktig er at det er anvendbart for eleven i det ordet blir innlært. Ord som blir valgt er ord som elevene ikke kan fra før, men som ikke er for vanskelig å forstå.

### **2.1.1 Bredde – og dybdevokabular**

Vokabularet til ulike mennesker varierer både når det gjelder antall ord, som er vokabularbredden, og ordbetydningen man kan og kvaliteten på ordkunnskaper, som da er ens vokabulardybde. Ved tilegnelsen av vokabular involverer det lagringen av representasjoner av

ordets fonologiske form i vokabular, sammen med semantiske representasjoner av ordmening (Levelt, Roelofs og Meyer, 1999). Man kan da se på vokabularet som et organisert lager av fonologiske ordformer som er forskjellige fra, men som man likevel kan si er i sterk sammenheng med semantiske representasjoner av ordmening (Coleman 1998; Levelt et al.1999). En følge av dette er at man kan trekke et skille mellom leksikalske enheter (vokabularbredde) og omfanget av semantiske representasjoner (vokabulardybde) (Ouellette, 2006).

Dette skillet stemmer med hva som er kjent om barns språkutvikling (Ouellette, 2006). Barn kan lagre et ord i sitt vokabular uten helt å forstå fullt ut betydningen (eller betydningene) av ordet. Dette er med på å utvide barns vokabularbredde, men ikke nødvendigvis vokabulardybden (Lahey, 1988). Over tid vil ordbetydningen bli videreutviklet; ordet vil bli forbedret, mer presist og utdypet, noe som vil bidra til økning i barnets vokabulardybde.

Noen vil se på bredde og dybde som to motsetninger, men er det mulig å se det på en annen måte? En kan anta at jo dypere kunnskap man har av et ord, desto flere ord kan man, eller det motsatte; at jo flere ord man kan, desto større vil nettverket være og jo dypere vil kunnskapen være. Ifølge Nagy og Herman (1987) mener de at jo tettere nettverket er rundt et ord, desto større er individets kunnskap om dette ordet. Det er dette som ofte menes som vokabulardybde. Vermeer (2001) fremhever at det enda er for lite kjent om forholdet mellom breddevokabular og dybdevokabular. Han ser på denne sammenhengen som en bekreftelse på at det ikke er konseptuelle forskjeller mellom disse to dimensjonene, vokabularbredde og -dybde er to dimensjoner av et fenomen. En mulig årsak til hvorfor bredde- og dybdevokabular korresponderer betydelig er måten dybdekunnskaper operasjonaliseres på i undersøkelser. Oppgaver som måler dybdevokabular baseres ofte på orddefinisjoner, noe som krever en viss mengde av breddevokabular kunnskaper (ibid).

## **2.2 Leseforståelse**

I Læreplanverket for Kunnskapsløftet (2006) beskrives lesing både som en grunnleggende ferdighet og som kulturell kompetanse. Gjennom lesing får elevene erfaringer som gir mulighet for læring og opplevelse og til å forstå seg selv og samfunnet. På de høyeste klassetrinnene i grunnskolen forutsettes det at elevene skal kunne bruke leseferdighetene i kunnskapstilegnelsen i de fleste av skolefagene.

Lesing er ofte i faglitteraturen blitt definert slik: Avkoding x Forståelse = Lesing. Med avkoding menes ofte teknisk leseferdighet som bygger på flere delferdigheter. Disse er nødvendige for at leseren skal kunne kjenne igjen ord eller nonsensord og sette dem sammen til et ”lydbilde” (Molander og Skauge, 2009). Den tekniske ferdigheten alene gir ikke mening til teksten. Det er først når teksten blir forstått, at eleven får utbytte av avkodingen (ibid.). Leseren bør helst være i stand til å avkode eller identifisere de enkelte ordene i en tekst uten for mye unøyaktighet eller strev hvis han eller hun skal oppnå en god leseforståelse (Bråten, 2007). Det er godt dokumentert at det er en sammenheng mellom ordavkoding og leseforståelse (Vellutino, 2003). I barneskolealderen er denne sammenhengen spesielt sterk, hvor problemer med å oppnå flytende ordavkoding ser ut til å være den viktigste og vanligste årsaken til dårlig leseforståelse (Bråten, 2007). På senere alderstrinn er imidlertid ikke sammenhengen mellom ordavkoding og leseforståelse like sterk. Da ser det ut til at andre komponenter, som for eksempel ordforråd, har større betydning for leseforståelsen (ibid.). Siden denne undersøkelsen tar utgangspunkt i elever på 8. klassetrinn og vokabularets betydning for leseforståelsen, vil vi videre ha hovedfokus på forståelsesaspektet ved lesingen.

Leseforståelse er en kompleks prosess og kan defineres som ”å utvinne og skape mening ved å gjennomføre og samhandle med skrevet tekst” (Bråten, 2007, s.45). Elevene skal på den ene siden finne fram til tekstens mening og på den andre siden skape mening ved å trekke slutninger som går utover den bokstavelige meningen i teksten (ibid.). I denne prosessen foregår det et kontinuerlig samspill mellom leseren og teksten. Leseren finner mening i teksten på bakgrunn av hans eller hennes tidligere kunnskap, hvordan teksten er og forfatterens budskap, hva som er hensikten med lesingen og den konteksten der lesingen finner sted (Carlisle og Rice, 2002). Å forstå en tekst er en utfordrende og krevende aktivitet som er avhengig av samspillet mellom en rekke faktorer for å lykkes (Brudholm, 2002). Som nevnt tidligere er en automatisert ordavkoding en nødvendig, men ikke en tilstrekkelig betingelse for en god leseforståelse. Videre vil vi gå mer i dybden på andre faktorer som kan ha betydning for en god leseforståelse.

### **2.2.1 Språkforståelse**

For å forstå en tekst må man ha en god språkforståelse. Elevenes muntlige språkkompetanse har også stor betydning for leseforståelsen. Dersom det skorter på den muntlige språkkompetansen, er det ikke mulig å lese med stor forståelse. (Bråten, 2007). Bråten (1997)



fremhever at leseren må anvende tre forskjellige ”standarder” for å evaluere forståelsen av en tekst:

### **Den leksikalske standard**

Det første språklige området som har betydning for en god leseforståelse er den leksikalske standard; det vil si ordforrådets betydning for leseforståelsen. God leseforståelse impliserer at leseren anvender den leksikalske standard, dette innebærer at leseren avgjør om han/hun har forstått enkeltord i teksten. Å kunne forstå et ord krever mer enn at eleven kan avkode ordet, eleven skal også kjenne begrepet bak ordet (Bråten, 1997 og Brudholm, 2002). Lesere med større ordforråd ser ut til å forstå det de leser bedre enn lesere med mindre ordforråd (Bråten, 2007).

### **Den syntaktiske standard**

En syntaktisk standard forutsetter en viss forståelse for grammatikalske regler og regler for setningsoppbygning (Bråten, 1997). Syntaktisk bevissthet handler om elevens sensitivitet for hva som er grammatisk akseptabelt, noe som setter dem i stand til å registrere om en setning er i overensstemmelse med språkets regler for bøyning og setningsbygning (Bråten, 2007). Leserene kan anvende denne standard for å avgjøre om den leste setning er korrekt utformet og således gir mening. Den syntaktiske kompleksitet i en setning kombinert med lange setninger kan være et problem i leseforståelsen (Brudholm, 2002).

### **Den semantiske standard**

Den semantiske standard tar hensyn til setninger og teksten som helhet. Denne standard benyttes bl.a. for å vurdere tekstens sammenheng og om for eksempel informasjonen i en setning eller et avsnitt er i overensstemmelse med informasjonen i en annen setning eller et annet avsnitt (Bråten, 1997). Det er den semantiske standard som krever den grundigste bearbeidelse av teksten og er den viktigste for god leseforståelse (ibid.).

## **2.2.2 Forkunnskaper**

Hva en elev vet om tekstens innhold fra før, er med på å bestemme hva eleven får ut av å lese teksten (Bråten, 2007). Lesere som har et omfattende nettverk av generell kunnskap og

ekspertise på spesifikke områder, forstår lettere og kan holde på mer informasjon enn dem som har mindre kunnskap (Carlisle og Rice, 2002). Det er særlig to slags forkunnskaper som har blitt studert av leseforskere; bredden av kunnskaper på et bestemt innholdsområde og dybden av kunnskaper om enkelte emner eller temaer innenfor et innholdsområde (Bråten, 2007). ”Både bredde- og dybdekunnskap som en elev bringer med seg inn i lesesituasjonen, kan ha avgjørende betydning for hans eller hennes læring og forståelse” (Bråten, 2007, s.61). Mange svake lesere har et mer begrenset vokabular og mindre bakgrunnskunnskaper enn normale lesere. Et annet vanlig problem er at svake lesere kan inneha forkunnskaper, men klarer ikke å nyttiggjøre seg disse mens de leser (Carlisle og Rice, 2002).

### **2.2.3 Å danne inferenser**

Leseforståelse innebærer evnen til å forstå mer enn det som direkte står i teksten. Leseren skal ikke bare kunne lese ordene som står på linjen, men også kunne lese mellom linjene.

Forståelse innebærer derfor en evne til å lese ting i teksten som ikke står der, men som er viktige for vår forståelse, det kalles å danne inferens. Å inferere vil si å dra slutninger som ikke eksplisitt er nevnt i teksten (Brudholm, 2002). Inferens er ikke en egenskap ved teksten, men noe som oppstår i leseren, avhengig av hans/hennes bakgrunnskunnskaper. Inferens krever at man sammenholder det språklige materiale med konteksten, fyller ut ”hull” i teksten og på den måten etablerer sammenhenger i teksten (ibid.).

### **2.2.4 Sjangerkjennskap og tekststruktur**

Å ha en god sjangerkjennskap har også betydning for leseforståelsen, for den respektive sjangeren er avgjørende for hva leseren kan forvente seg av teksten; ikke bare hva slags tekst det er, men også noe om måten den er skrevet på (Brudholm, 2002). En sjanger angår både tekstens form og dens innhold. Den angir hvordan saken som teksten forteller om, bør forstås. Sjangerkjennskap kan være et slags arbeidsredskap, og utgjør derfor en generell forståelsesramme som ”tilbys” leseren før lesingen av teksten (ibid.). Med struktur menes hvordan teksten er bygd opp og tekststruktur viser vanligvis til to typer tekster: narrative (fortellende) og forklarende (sakprosa) (Strømsø, 2007). I arbeidet med å forstå en tekst er interaksjonen mellom tekstens struktur og leserens kjennskap til denne strukturen viktig (ibid.). Strukturen i en tekst er helt avgjørende for hvordan den blir forstått. Tekstens struktur har betydning for hvilke strategier leseren tar i bruk for å forstå innholdet; leseren kan lettere

overse unødvendig eller tilfeldig informasjon og dermed lettere å velge ut den sentrale informasjonen. Leseren kan også enklere se den mer generelle betydningen av innholdet og konstruere noen hovedideer som sammenfatter teksten (van Dijk og Kintch, 1978).

### **2.2.5 Kognitive evner**

Det er visse ikke – språklige kognitive evner som har åpenbar betydning for leseforståelsen. I følge Vellutino (2003) gjelder dette først og fremst oppmerksomhet, visuell forestillingsevne og generell intelligens.

Elever som ikke lar seg avlede av verken indre eller ytre forstyrrelser, som for eksempel tanker og følelser eller fysiske og sosiale omgivelser, og som kan arbeide konsentrert over tid, har et klart fortrinn når de skal forsøke å forstå innholdet i det de leser, og dess mer krevende teksten er, dess større fortrinn vil de ha (Bråten, 2007). Faktorer som kan skyldes oppmerksomhetsproblemene kan være manglende interesse for innholdet i bestemte tekster, lite motivasjon for lesing generelt, mangelfullt ordforråd, svake avkodingsferdigheter etc.(ibid.).

Leserens evne til å danne, lagre og framhente visuelle forestillinger kan påvirke leseforståelsen. Gode lesere reagerer ofte på teksten ved å danne seg indre bilder av de meningene som teksten inneholder. Dette kan føre til en dypere forståelse av det leste, på grunn av de personlige tolkningene teksten representerer (Bråten, 2007).

Elever med god og dårlig leseforståelse skiller seg ofte fra hverandre med hensyn til generell intelligens (Bråten, 2007). Å kunne tenke logisk, analysere informasjon kritisk og trekke slutninger som går utover det som er direkte uttrykt i teksten er viktig for å kunne forstå det man leser. Dette er ferdigheter som vanligvis blir forsøkt kartlagt ved hjelp av intelligenstester, som for eksempel Wechsler – batteriene eller Ravens matrisetester (ibid.). Mens sammenhengen mellom generell intelligens og leseforståelse virker ganske klar, er det liten sammenheng mellom målt intelligens og ordavkodingsferdigheter. ”Dette bekrefter at god leseforståelse avhenger av komponenter som favner videre enn feilfri og flytende ordavkodning” (Bråten, 2007 s.60).

## 2.2.6 Leseforståelsesstrategier og metakognisjon

”Leseforståelsesstrategier kan defineres som mentale aktiviteter som leseren velger å iverksette for tilegne seg, organisere og utdype informasjon fra tekst, samt for å overvåke og styre sin egen tekstforståelse.” (Bråten 2007, s.67). Samuelstuen og Bråten (2005) deler leseforståelsesstrategier inn i fire hovedkategorier; hukommelsesstrategier, organiseringsstrategier, elaboreringsstrategier og overvåkingsstrategier. For å repetere eller gjenta informasjon i teksten, brukes hukommelsesstrategier. Dette kan gjøres ved å lese et avsnitt flere ganger, gjentelle eller skrive deler av teksten for å memorere den.

Organiseringsstrategier brukes for å binde sammen, gruppere eller ordne informasjon eller ideer som blir presentert i teksten, for eksempel ved å tegne begrepskart, skrive sammendrag eller skaffe seg oversikt over innholdet i teksten. For å gjøre teksten mer meningsfull ved at den nye informasjonen bearbeides, belyses, utbroderes, berikes eller forandres i lys av den kunnskapen leseren har fra før, brukes elaboreringsstrategier. Dette innebærer at leseren konstruerer relasjoner mellom noe han eller hun kan fra før og det som skal læres.

Overvåkingsstrategier bruker leseren for å sjekke eller evaluere sin egen forståelse under lesingen, for eksempel ved å spørre seg selv om han eller hun har forstått det som leses. Da kan leseren enten bekrefte at forståelsen av teksten er god, eller hvis forståelsen er ufullstendig kan leseren øke den strategiske innsatsen. Dette kan gjøres ved eventuelt å sette inn andre strategier som kan hjelpe leseren med å organisere og elaborere informasjonen. Å overvåke sin egen forståelse er en form for metakognisjon (Bråten, 2007).

”Metakognisjon, dvs.: viten om og regulering av egne kognitive prosesser” (Brudholm, 2002 s. 57). Forståelsen av teksten kan foregå mer eller mindre automatisk, men den kan også være gjenstand for bevisste valg og bevisst styring fra leserens side. Bevisstheten om forståelsen kalles metakognisjon. Metabevissthet innebærer at leseren er seg bevisst både på hva han forstår og hva han ikke forstår, samt at han kan styre forståelsesprosessen. Den viktigste metakognitive strategi er å være seg bevisst om hvorfor forståelsen bryter sammen, enten fordi teksten inneholder et ord eller flere man ikke forstår, at den har en komplisert syntaks, at teksten forutsetter en viten om verden som leseren ikke har, osv. (ibid.).

Gode lesere er aktive i leseprosessen, de bygger kunnskap, leser med et spesifikt formål og handler ut fra dette. De er kort sagt strategiske før, under og etter at de har lest en tekst. Strategisk lesing dreier seg ikke bare om å bruke lesestrategier effektivt, men også kunne bruke dem fleksibelt, dvs. å vite i hvilke situasjoner de ulike strategiene egner seg best

(Bråten, 2007). Svake lesere har en tendens til å ha mindre kunnskap om strategier og hvordan de skal brukes. De er også mindre bevisste på hvilke krav ulike leseoppgaver stiller, mindre bevisste sine egne styrker og svakheter, dårligere til å avgjøre vanskelighetsgraden til en tekst og klarer dårligere å overvåke sin forståelse og regulere bruken av strategier (Carlisle og Rice, 2002).

## 2.3 Sammenhengen mellom vokabular og leseforståelse

En av det viktigste ferdighetene unge elever lærer seg er kunnskapen til å forstå en skrevet tekst, som i de fleste sammenhenger blir kalt leseforståelsen. Det å klare å hente ut informasjon fra en tekst er den ultimate hensikten med lesing. Vokabularet er fundamentet til leseforståelsen. En kan ikke forstå en tekst uten å vite hva de fleste ord betyr (Nagy, 1988). Helt fra tidlige barneår er alle former for erfaringer vi gjør oss og mengde av ord vi øver oss på, om det skjer i hjemmet eller senere på skolen, viktig. Dette er med på å øke vokabularet, som senere vil være med på å påvirke leseforståelse (Dickinson og Tabors, 2001; Hart og Risley, 1995). Ved denne opplagte sammenhengen skulle man tro at all form for vokabularøvelser er med på å øke leseforståelsen. Ifølge flere studier viser det seg at mange vanlige metoder feiler i forsøket på å øke leseforståelsen gjennom øving av ord. Eksempel på tradisjonelle metoder som er ineffektive er for eksempel å memorere definisjoner og skrive en setning der en bruker ordet (Mezynski 1983; Pearson og Gallagher 1983; Stahl og Fairbanks 1986 referert fra Nagy, 1988).

Anderson og Freebody (1981) var blant de første som så på at et stort vokabular kunne ha sammenheng med leseferdigheter. De kom fram til tre hypoteser som kan være med på å forklare denne sammenhengen. Den første kalte de for instrumenthypotesen. Den er enkel; jo flere ord man kan, desto bedre leser blir man. Den andre hypotesen er kunnskapshypotesen. Med dette menes at det ikke er den individuelle forståelsen av hvert ord som øker forståelsen, men det er selve hovedinnholdet i teksten som vil avgjøre hvor godt man vil forstå teksten. Kunnskap om et enkelt ord er bare toppen av isfjellet, det er en god og sammensatt kunnskap om et tema som virkelig øker forståelsen. En annen måte å si det på er ved at en øker forbindelsen mellom allerede eksisterende ordmeninger. Ordene er da lagret i det semantiske nettverket hvor hvert ord har en forbindelse til andre ord, og til andre konsept, fakta og

spesifikke minner. Så når et ord blir gjenkjent vil også forbindelsen til andre ord i nettverket bli fremkalt. Dette er med på å øke elevens assosiasjoner når de leser et ord, så jo sterkere disse assosiasjonene er, jo mer vil eleven få ut av en tekst. Så ved at elevene kan noe om innholdet i teksten, er det slik at jo mer kunnskap de har, desto mer vil de klare å få ut av teksten. Den tredje hypotesen de har er om forholdet mellom vokabular og leseforståelse og evnen elevene har til å bruke kunnskapene sine. I følge denne hypotesen så er det korrelasjon mellom vokabular og leseforståelse, ikke fordi den ene er med på å øke kunnskapen i den andre, men på grunn av at både vokabular og leseforståelse reflekterer til en mer generell underliggende verbal evne. Det vil si at elever med gode verbale evner har bedre læring av nye ord, de vil da ende opp med et større vokabular og av den grunn vil de forstå en tekst bedre.

Sternberg og Powell (1983) mener denne hypotesen fokuserer på evnen elevene har til inferens, at dette er en måte å forklare noe av sammenhengen mellom vokabularkunnskap og leseforståelse på. Inferens er evnen eleven har til å lese mellom linjene, de forstår mer enn akkurat det som står i teksten. Det kan også være andre evner eleven har som kan være med på å forklare forholdet mellom vokabularkunnskaper og leseforståelse. En kan tenke seg at metalingvistisk bevissthet er en slik evne. Fonologisk bevissthet, bevissthet om språklydene, er kjent for å være en metalingvistisk evne som spiller en viktig rolle i tidlig lesing. Det viser seg at elever med et stort vokabular også har en høyere fonologisk bevissthet, som igjen gir suksess i tidlige leseferdigheter. Gode ferdigheter i lesing er igjen med på å øke leseforståelse. Andre aspekter ved metalingvistiske evner er å ha en bevissthet om strukturen av ordet i form av morfem som er den minste meningsbærende enhet. Elever med god morfologisk bevissthet viser seg å ha et større vokabular, og omvendt. Morfologisk bevissthet er også med på å øke elevens leseforståelse ved å hjelpe eleven i avkoding av lange ord med mer nøyaktighet og flyt.

En annen type av metalingvistisk kunnskap som etter hvert vil bli viktig for elevene, er kunnskap om ordenes flertydige mening. At de forstår at et ord kan ha flere meninger, det blir viktig at de forstår sammenhengen ordet er brukt i inni teksten. Det en kan slå fast er at med et større vokabular ikke bare kan mange ord, men at man også vet meningen av mange flere ord (Qian, 1999; Vermeer, 2001).

I forbindelse med disse tre hypotesene til Anderson og Freebody har Mezynski (1983) referert i Carlisle & Rise (2002) lagt til en hypotese til. Den har hun kalt tilgangshypotesen. Med det

mener hun hvor fort man har tilgang til ordene; jo større vokabular en har, desto større vil tilgangen til ordene være. Dette gir en bedre flyt og leseforståelse. Har man mindre tilgang til ord kan dette være med på å senke flyten og man kan av den grunn gå glipp av viktig informasjon i teksten.

Tilslutt har Stanovich (1986) og andre foreslått den resiproke hypotesen. Det at relasjonen mellom vokabular og leseforståelse går begge veier, i en slags vekselvirkning. En kan se på denne hypotesen som en sirkel. Å ha et større vokabular gjør deg til en bedre leser, å være en bedre leser fører til at du leser mer og ved å lese mer får du et større vokabular. Denne utviklingen vil bare øke etter hvert. Dette er den positive siden, bedre lesere leser mer og øker vokabularet sitt, som igjen fører til at man blir en bedre leser. Men som nevnt tidligere har dette også en negativ side. Dårligere lesere leser mindre og vil ikke utvikle noe stort vokabular. De finner lesing for vanskelig. Etter hvert vil tekstene bli vanskeligere og kreve et større vokabular. Det er ikke bare vokabularkunnskapene som blir utviklet i en slik gjensidighet. Å ha mye kunnskap om verden gjør at man blir en bedre leser, dette igjen fører til at man leser mer, og ved å lese mer øker man sine bakgrunnskunnskaper.

Kan man så si noe om at noen av disse hypotesene er mer sanne enn de andre? Anderson og Freebody (1981) poengterte at disse hypotesene ikke er avhengige av hverandre. Om den ene er rett fører det ikke til at de andre er feil. Forholdet mellom vokabular og leseforståelse er et komplekst forhold. Dette gjør at ingen av hypotesene er helt feil.

Oppbygging av ordforråd og bakgrunnskunnskaper er viktig for å oppnå en flytende ordgjenkjennelse. Pikulski (2006) fremhever at det er viktig at leseren har hørt ordene i teksten før og vet hva de betyr. I tidenes løp har det vært forskjellige vurderinger av hvor mye elever på de enkelte klassetrinn kan forventes å utvide deres ordforråd. Snow, Burns & Griffin (1998) har for eksempel anført at elever bør utvide deres ordforråd med 2 500 – 3 000 ord hvert år. For den enkelte leser er det å utvide sitt ordforråd langt mer krevende enn å lære sammenhengen mellom lyder og bokstaver. Selv om ordforråd og bakgrunnskunnskaper ikke er helt det samme, er det meget tett forbundet (Pikulski, 2006).

Snow et al.(1998) peker på at tidlige lesevaner kan vise seg først i 4. klasse, grunnen til det er at da jobber elevene oftere og mer dyptgående med fagprosa innenfor de ulike fagområdene. Betydningen av leserens bakgrunnskunnskaper for leseforståelsen er bredt og anerkjent dokumentert. Det ser ut til at manglende kjennskap til de anvendte ord og

forkunnskaper spesielt innenfor de fagområder det arbeides med i skolen er en barriere for leseforståelsen. Disse faktorer kan også spille en rolle for generelt dårlige prestasjoner hos elever, som klarer seg bra i lesing på de tidlige klassetrinn, men som kommer i vanskeligheter, når innholdet i tekster de møter senere i sin skolegang stiller større krav til deres ordforråd (Pikulski, 2006).

## 2.4 Tidligere empiriske undersøkelser

Når det gjelder tidligere forskning som undersøker sammenhengen mellom vokabular og leseforståelse, har denne studien i størst mulig grad tatt utgangspunkt i Ouellette (2006) og Tannenbaum, Torgesen & Wagner (2005). Grunnen til det er at de er mest like undersøkelsen i denne studien, ved at det stort sett er brukt de samme variablene der vokabular deles inn i bredde og dybde. Videre er undersøkelsene til Petersen (2008), Carlisle & Rice (2002) og Wagner, Muse & Tannenbaum (2007) sett på, fordi disse også undersøker forholdet mellom vokabular og leseforståelse.

### 2.4.1 Sammenhengen mellom bredde- og dybdevokabular og leseforståelse

Ouellette (2006) undersøkte 85 fjerdeklassinger med hensyn til avkoding, visuell ordgjenkjennelse, leseforståelse, bredde – og dybdevokabular. I sin undersøkelse skiller han mellom reseptivt og ekspressivt breddevokabular. I denne forbindelse velges det her å ta utgangspunkt i det reseptive breddevokabularet, grunnen til det er at Ouellette (2006) har brukt testen Peabody Picture Vocabulary Test (Dunn & Dunn, 1981) for å måle reseptivt breddevokabular. Dette er den amerikanske versjonen av testen BPVS II, som denne studien bruker som mål på breddevokabular. Ouellette (2006) fant i sin undersøkelse at dybdevokabular har en unik forklaringsvarians på 8 % etter at det er kontrollert for alder, non – verbal intelligens, avkoding, visuell ordgjenkjennelse og breddevokabular og signifikansnivået er satt til 1 %. Breddevokabular har en unik forklaringsvarians på 2,5 % etter at det er kontrollert for alder, non – verbal intelligens, avkoding, visuell ordgjenkjennelse og dybdevokabular og resultatet er ikke signifikant (ibid.). Når bredde – og dybdevokabular ikke er kontrollert for hverandre er den unike forklaringsvariansen til dybdevokabular 12,1 % og til breddevokabular 6,1 % og disse har et signifikansnivå satt til henholdsvis 0,1 % og 5 %.



Som resultatene i denne undersøkelsen viser forklarer dybdevokabularet den største variansen på leseforståelsen sett i forhold til breddevokabular.

Det delte synet blant forskere når det gjelder vokabularets rolle i prediksjonen av leseforståelse, er at dybdevokabular, i større grad enn breddevokabular, predikerer leseforståelse i en positiv retning. Det vil si at fonologiske faktorer er mindre relevante her, enn semantisk kunnskap og organisering (Muter, Hulme, Snowling og Stevenson, 2004; Nation og Snowling, 2004; Roth, Speece og Cooper, 2002; Sènèchal, Ouellette og Rodney, 2006; Share og Leikin, 2004; Snow, Tabors, Nicholson og Kurland, 1995). Bishop (1997) definerer vokabular som de mentale representasjoner av ord som er lagret i individets minne og disse mentale representasjonene inneholder informasjon om både fonologisk form av ord og representasjon av deres betydning. Basert på Bishop (1997) sin definisjon av vokabular, er ord et fenomen som har minst to sider; fonologisk form og semantisk innhold.

Som nevnt tidligere menes det med vokabularbredde den mengde ord som en person kjenner og klarer å hente frem fra hukommelsen, når han/hun har bruk for dem, uten nødvendigvis å ha full forståelse for betydningen av alle ordene man bruker. Med vokabulardybde menes den mengde av ord en person har klar forståelse for i et gitt språk (Bø og Helle, 2008).

Vokabularutvikling forutsetter både fonologisk og semantisk vekst, men ut fra tidligere studier ser vi viktigheten av å ha fokus på vokabulardybde, da dette viser seg å ha størst effekt på utviklingen av en god leseforståelse (Ouellette, 2006).

Viktigheten av vokabular når det gjelder leseforståelse har mange ganger vist seg i en rekke longitudinelle studier; Muter et al. (2004) fant en moderat korrelasjon mellom nivået på vokabularet hos barn i barnehagen og deres nivå på leseforståelse to år senere. Snow et al. (1995) fant samme korrelasjoner på resultater fra vokabulartester hos barn i barnehagen og leseprøver som inkluderte forståelse i 1. klasse. Regresjonsanalyser fra 2. klassingers leseforståelse indikerer at vokabular kunnskap målt i barnehagen var en sterkere prediktor på leseforståelse enn ordavkodingsferdigheter (Roth et al. 2002). Disse resultatene viser at vokabular spiller en prediktiv rolle for senere tilegnelse av leseforståelse. Når det gjelder forskjellen mellom bredde – og dybdevokabular, har de nevnte undersøkelser evaluert barnas evne til muntlig å forklare ord; altså barna fikk oppgaver som krevde semantisk kunnskap. Dette viser at dybdevokabular, i en sterkere grad enn breddevokabular, har en større effekt på leseforståelse. Studier relatert til breddevokabular, der barnas reseptive vokabular måles gjennom gjenkjennelse av enkeltord, for å si noe om størrelsen på barnets leksikon; altså hvor

mange ord barnet kan, har ikke like stor effekt på leseforståelsen som tidligere nevnte studier, der dybdevokabularet har blitt brukt som mål (Roth et al.2002 og Snow et al.1995).

Petersen (2008) har gjort en undersøkelse på hva som har betydning for elevers leseforståelse. Undersøkelsen tok utgangspunkt i elever som gikk i 7. til 10. klasse ved ni forskjellige skoler. Elevene ble undersøkt innen fem kategorier: leseforståelse, vokabular, avkoding, kognitive ferdigheter og kjennskap til teksttyper. Testene tar utgangspunkt i leseforståelse og vokabularkunnskaper innen sjangeren allmenkunnskap, historie og biologi. Resultatene fra analysen av de tre vokabulartestene viser at de forskjellige testene uavhengig av hverandre kan bidra til forklaring av variasjoner i leseforståelsen. Vokabulartestene som hun brukte var en allmenn ordkjennsprøve, der elevene skulle velge det ordet, blant tre mulige ord, som betyr det samme som eksempelordet. Denne testen måler elevenes faglige og allmenne ordkjennskap. Så brukte hun en faglig ordkjennsprøve med historie og biologi som utgangspunkt. Denne testen er utformet etter samme metode som den allmenne ordkjennsprøven. I den tredje testen skulle elevene finne det ordet som ikke hører til, eksempel på dette er: sko, frakk, belte, teppe. Ved denne testen kan man si noe om elevenes semantiske kategoriseringsevne. Videre i analysen så hun på hva som bidro mest til variasjon av leseforståelsen, avkoding eller vokabular. Tallene viste at vokabular forklarer mer av variasjonene enn avkoding. Men vokabular og avkoding utgjør hele 45 % av forklaringen i variasjonen i leseforståelsen.

En faktor som er med på å øke effekten mellom forholdet til vokabular og leseforståelse er å øke barns grunnleggende lesekunnskaper. Dette er med på å gi tilgang til ideer og informasjon i teksten. Mezynski (1983) sin undersøkelse, som Carlisle og Rice (2002) refererer til, viser at dårlige avkodere kan ha et godt vokabular, men med dårlige lesekunnskaper klarer de ikke å bruke vokabularkunnskapen sin for å forstå teksten. Så selv om man er god i vokabular må man også være en god avkoder, for å få en sammenheng mellom vokabular og leseforståelse.

Tannenbaum, Torgesen og Wagner (2005) har gjort en undersøkelse på forholdet mellom ordkunnskap og leseforståelse hos 3. klassinger. I undersøkelsen ble bredde – og dybdevokabular og ordgjenkjennelse sett på opp mot leseforståelse som tre underliggende variabler. De fant i denne studien at breddevokabular hadde en unik forklaringsvarians på 19 % til leseforståelse, mens dybdevokabular/ordgjenkjennelse hadde en unik forklaringsvarians på 2 %. Men begge variablene var signifikante i forhold til leseforståelsen.

En annen studie som har blitt gjort, testet fire alternative modeller av utviklingsforholdet mellom vokabular og leseforståelse. Dette er en longitudinell studie av elever fra andre til femte klasse (Wagner, Muse og Tannenbaum, 2007). Den første modellen viser at individuelle forskjeller i leseforståelse utøver en kausal innflytelse på vokabularutvikling, fordi mange nye ord læres via konteksten fra en lest tekst bedre, enn gjennom å bli fortalt meningen av et ord. Dette viser at gode lesere har lettere for å finne ut hva et ord betyr, enn det dårlige lesere har. Her konkluderes det med at individuelle forskjeller i leseforståelse fører til individuelle forskjeller i vokabular; altså jo bedre leseforståelse man har, desto større vokabularkunnskap vil man ha. Den andre modellen forteller oss individuelle forskjeller i vokabular utøver en kausal innflytelse på utviklingen av leseforståelse. Et større vokabular er med på å styrke opplevelsen av lesingen, som over tid vil øke leseforståelsen. Den tredje modellen sier at både modell 1 og 2 foregår parallelt, og har en gjensidig påvirkning på hverandre. Den fjerde modellen sier at vokabular og leseforståelse korrelerer fordi de har en felles tredjevariabel (ibid.). Eksempler på en tredjevariabel kan være Anderson og Freebody (1981) sin instrumenthypotese, kunnskapshypotese og hypotesen om evnene eleven har til å bruke kunnskapene sine, som vi har beskrevet tidligere. Metalingvistisk bevissthet kan også være en slik tredjevariabel (Wagner et al.,2007).

#### **2.4.2 Sammenhengen mellom generelle kognitive evner og leseforståelse**

Som nevnt tidligere er der visse ikke – språklige kognitive evner som har en åpenbar betydning for leseforståelsen. Dette gjelder først og fremst oppmerksomhet, visuell forestillingsevne og generell intelligens (Vellutino, 2003). Her legges hovedvekten på sammenhengen mellom generelle kognitive evner og leseforståelse. Elever med god og dårlig leseforståelse skiller seg ofte fra hverandre med hensyn til generell intelligens (Bråten, 2007). Å kunne tenke logisk, analysere informasjon kritisk og trekke slutninger som går utover det som er direkte uttrykt i teksten er viktig for å kunne forstå det man leser. Videre presenteres undersøkelser som ser på sammenhengen mellom leseforståelse og generelle kognitive evner målt gjennom testen Raven.

To ulike hypoteser har blitt fremlagt; at relasjonen mellom leseferdigheter og generell kognitiv fungering målt med Raven er liten og ikke – signifikant. Og at denne relasjonen er sterkere på de høyere klassetrinnene sammenlignet med de lavere klassetrinnene (Stanovich,

Cunningham & Feeman, 1984). Det er disse to hypotesene som ligger til grunn for undersøkelsen som Craver (1990) gjennomførte. Han brukte Raven som mål på generell kognitiv fungering og National Reading Standard Test som mål på leseferdigheter. Sistnevnte test er en kombinasjon av mål på nøyaktighet og hastighet, som igjen sier mye om en persons leseforståelse. Testen fungerer slik at testpersonen for hvert femte ord får en valgmulighet på to ord, der bare det ene skal settes inn i teksten. Det riktige ordet er det som passer inn i resten av setningen (Craver, 1977). Utvalget bestod av 486 elever fra 2. – 12. klasse. Resultatene fra undersøkelsen viser at korrelasjonene mellom skårene på Raven og leseforståelse ligger mellom .40 og .60 på 8 av 11 klassetrinn. Dette funnet er i sterk motsetning til det Stanovich et.al (1984) fant i sin undersøkelse; at relasjonen mellom leseferdigheter og generell kognitiv fungering målt med Raven er liten og ikke – signifikant. Craver (1990) fant heller ingen indikasjon på at korrelasjonen mellom Raven og leseforståelse øker nevneverdig med alder. Gjennomsnittet på korrelasjonene mellom 2. – 6. klasse var på .48 og gjennomsnittet på korrelasjonene mellom 7. – 12. klasse var på .49. Dette motbeviser også den andre hypotesen til Stanovich et al. (1984) om at på de høye klassetrinnene er relasjonen mellom Raven og leseforståelse sterkere enn på de lave klassetrinnene.

Finbak (1998) har gjort en studie der sammenhengen mellom avkodingsvansker og stavevansker hos elever med dysleksi og mer generelle lese – og stavevansker er undersøkt. Undersøkelsen er delt inn i to faser. Alle elevene deltok i fase en, der målet var å identifisere en gruppe elever med lese – og stavevansker ved å benytte en prosedyre som kunne gi et entydig grunnlag for å skille ut elever som hadde lese – og stavevansker. 586 elever gjennomgikk en leseprøve, to orddiktater, en setningsdiktat og to intelligensprøver. De resultatene elevene oppnådde på leseprøven og på en av diktatene i fase en, dannet grunnlaget for å velge ut de elevene som hadde både lese – og stavevansker, til sammen 46 elever (ibid.). I denne sammenhengen her blir i hovedsak resultatene fra elevgruppen i fase en og sammenhengen mellom leseprøven og den ene intelligenstesten diskutert nærmere.

Leseprøven Finbak (1998) brukte i sin undersøkelse er Carlstens leseprøve for 7. klasse (Carlsten, 1993), denne prøven består av en tekst på til sammen 998 ord og den er en kontroll for leseforståelse ved at enkelte ord er utelatt fra teksten og erstattet med en parentes med tre ord, der det ene ordet er det utelatte. Eleven skal sette strek under det ordet som passer best til teksten. Prøven har 36 slike parenteser. Den ene intelligenstesten som blir brukt i Finbak (1998) sin undersøkelse er tilsvarende som i denne undersøkelsen; Ravens standard

progressive matriser. Resultatene viser en korrelasjon på .15 mellom leseforståelse målt med Carlstens leseprøve for 7. klasse og generelle kognitive evner målt med testen Raven. Som vi ser er denne korrelasjonen lavere enn det som både ble funnet i Craver (1990) sin undersøkelse og foreliggende undersøkelse.

## 2.5 Problemstillinger

Som vi har sett i denne delen viser tidligere forskning på dette feltet variasjoner i sine resultater når det gjelder sammenhengen mellom vokabular og leseforståelse. Disse tidligere studiene har i hovedsak undersøkt yngre elever, vi ønsker derfor med denne undersøkelsen å si noe om hvilke betydninger vokabular har på leseforståelsen hos eldre elever og se om disse samsvarer med forskningen som har blitt gjort tidligere. Vi har valgt å skille mellom bredde – og dybdevokabular. Grunnen til det er fordi det har blitt gjort i andre studier og med dette vil man finne ut om noen av disse i mer eller mindre grad har ulik innvirkning på leseforståelsen. Med utgangspunkt i dette er det formulert følgende problemstilling:

*Hvilke sammenhenger er det mellom vokabular og leseforståelse blant 8. klassinger?*

Med utgangspunkt i dette ble det videre formulert to ulike forskningsspørsmål:

*Hvilken innvirkning har vokabularbredde på leseforståelsen blant 8.klassinger?*

*Hvilken innvirkning har vokabulardybde på leseforståelsen blant 8.klassinger?*

## 3 Metode

I følgende kapittel beskrives den fremgangsmåten som er benyttet for å få svar på problemstillingen, altså denne undersøkelsens metode. Det vil bli redegjort for oppgavens design, utvalget av informanter som er testet for å innhente data, prosedyrer ved datainnsamling, testene brukt i denne oppgaven og analysen. Deretter kommer en beskrivelse av validitet og reliabilitet knyttet til denne typen undersøkelse. Drøftingen av ulike problemstillinger knyttet til validitet og reliabilitet vil vi komme nærmere inn på i kapittel 5. Til slutt i dette kapittelet belyses det hvordan forskningsetiske hensyn er ivaretatt i undersøkelsen.

### 3.1 Design

Problemstillingen i denne undersøkelsen er belyst med en kvantitativ metodisk tilnærming. Undersøkelsen har et ikke – eksperimentelt design, som kjennetegnes av fravær til å kunne manipulere variabler (Kleven, 2002a). Dataene som danner grunnlag for studien er hentet inn på et gitt tidspunkt og studien kan betraktes som en tverrsnittsstudie. Ved å bruke en kvantitativ forskning har det blitt gjort måling av sammenhengen mellom ulike variabler (ibid.).

I denne oppgaven er problemstillingen delt inn i to forskningsspørsmål som undersøker bredde- og dybdevokabularets innvirkning på leseforståelsen blant 8. klassinger. Denne sammenhengen blir beskrevet gjennom korrelasjonsstatistikk, og deretter drøftet nærmere (Gall, Gall og Borg, 2007). Gjennom en slik analyse søker vi etter kunnskap om hvilke variabler som samvarierer, og hvilke variabler som tilsynelatende er uavhengige (Befring, 2007). I denne studien undersøkes vokabularets betydning for leseforståelsen. Dette designet har gitt mulighet til å studere hvor stor innvirkning variabelen vokabular har på leseforståelsen. Gjennom dette er det ønskelig å undersøke videre hvorvidt ulike variabler kan forklare sammenhengen til den avhengige variabelen, noe som også kan betegnes som en prediksjonsstudie (Gall et. al, 2007). I denne undersøkelsen er vokabular prediktorvariabel, som også blir kalt den uavhengige variabel. Prediktorvariabelen forklarer variasjon i kriterievariabelen. Variabelen vokabular deles inn i to, bredde - og dybdevokabular. Leseforståelsen blir i denne studien kriterievariabelen som er den avhengige variabelen. Dette

blir da den variabelen som man ønsker og predikere fremtidige hendelser, betingelser, og/eller prestasjoner som er målt på et tidligere tidspunkt (ibid.). To tredjevariabler blir brukt i denne undersøkelsen; den ene er mål på generell kognitiv fungering, i denne forbindelsen brukes testen Raven. Den andre er mål på avkodingsferdigheter, i denne sammenhengen brukes ordkjedetesten. Disse variablene er med på å danne grunnlag for statistisk korrelasjons- og samvariasjonsanalyser (Befring, 2007).

## 3.2 Utvalg

Data i denne studien ble samlet fra 112 elever i 8. klassetrinn på to større skoler i Oslo. Utvalget var knyttet til forskningsprosjektet KiSP, og omfatter de eldste elevene i prosjektet. Valg av 8.klassinger i denne studien begrunnes med at på senere alderstrinn er ikke sammenhengen mellom ordavkodning og leseforståelse like sterk, i motsetning til i barneskolealderen, der denne sammenhengen er spesielt sterk og ser ut til å være den viktigste og vanligste årsaken til dårlig leseforståelse (Bråten, 2007). I 8. klasse ser det ut til at andre komponenter, som for eksempel vokabular, har større betydning for leseforståelsen (ibid.). Jo eldre barnet blir, desto større krav stilles det til å ha en forståelse for det man leser. For at man skal ha en forståelse for det man leser, er det viktig at man har et vokabular som danner grunnlaget for økt innsikt og dermed også økt forståelse. I denne studien vil vi først og fremst undersøke betydningen av vokabular knyttet opp mot en eventuell effekt på lesing hos elever i 8.klasse med en alder fra 12 – 13 år.

Ved begge skolene er det et stort innslag av minoritetsspråklige, vårt utvalg består av de elevene som er definert som majoritetsspråklige elever. Med dette menes de elever som har norsk som morsmål. Grunnen til at minoritetsspråklige elever er valgt bort, er fordi det kan være store spredning i elevenes norskkunnskaper. Årsaken til dette er at noen elever har bodd i Norge lenge og er flinke i norsk, mens andre har vært her i kortere tid og har lite kunnskaper i norsk.

Det er ikke tatt hensyn kjønnsfordeling i denne undersøkelsen. Det er 54 jenter og 58 gutter med i utvalget. Ut fra problemstillingen og forskningsspørsmålene er det ikke lagt opp til å se på forskjeller ved kjønn. Det som ønskes undersøkt er i hovedsak vokabularets betydning for leseforståelsen på tvers av kjønn.

### 3.3 Gjennomføring

Gjennomføringen ble utført i regi av forskergruppa i KiSP – prosjektet i samarbeid med skolene. Prosjektet er longitudinelt og vil pågå over tre år. Siden dette er begynnelsen på prosjektet, har vi vært med i gjennomføringen av pre – testfasen.

De individuelle testene ble gjennomført av mastergradsstudenter ved Institutt for spesialpedagogikk ved Universitetet i Oslo, med spesifikke lærevansker eller logopedi som fordypningsfelt og to forskningsassistenter tilknyttet KiSP – prosjektet, i løpet av oktober og november 2009. Ordkjedetesten og Nasjonale prøver i lesing ble gjennomført av lærerne ved de to skolene. Forskningsassistentene gjennomførte Raven og masterstudentene gjennomførte de individuelle prøvene British Picture Vocabulary Scale (BPVS II) og deltesten Ordforståelse i Wechsler Intelligence Scale for Children (WISC III). Lærerne, forskningsassistentene og mastergradsstudentene har alle fått opplæring i administreringen av testene av de ansvarlige for KiSP - prosjektet. Grunnen til dette er for å sikre lik gjennomføring av testene og like prosedyrer for skåring av testene.

Alle testene ble gjennomført på elevenes respektive skoler; gruppetestene i elevenes ordinære klasserom og de individuelle testene i avskjermede grupperom. De individuelle testene ble gjennomført på ca. 30 – 40 minutter per elev. Dette skaper trygghet for eleven. Elevene var også informert på forhånd om at de ville bli testet. Det var også samme student som administrerte begge de individuelle testene på den samme eleven. Dette for å sikre reliabiliteten til testene, og at eleven skal ha minst mulig personer å forholde seg til. Alt dette vil mest sannsynlig være med på å legge forholdene til rette for en maksimal yting fra testpersonen.

Gjennomføringen av testene ble gjort i henhold til testmanual og skåringen av testene ble utført av mastergradsstudentene, forskningsassistentene og lærerne. Tilslutt ble resultatene lagt inn i dataprogrammet SPSS av i hovedsak masterstudenter og forskningsassistenter tilknyttet KiSP – prosjektet.



## 3.4 Instrumenter

I det følgende beskrives testene som er brukt i undersøkelsen. Testene blir her kun presentert. Hvorvidt de er egnet til å frembringe data som kan belyse oppgavens problemstillinger, blir diskutert i avsnittet om validitet.

### 3.4.1 Nasjonale prøver i lesing

Nasjonale prøver gjennomføres for å få informasjon om elevenes grunnleggende ferdigheter er i samsvar med målene som står oppført i læreplanen. Der defineres de grunnleggende ferdigheter som: å kunne uttrykke seg muntlig og skriftlig, å kunne lese, regne og bruke digitale verktøy. Disse ferdighetene er en del av alle fag, de har stor betydning for hvordan elevene klarer seg på skolen og senere i livet og de er viktige for læring og utvikling i alle fag.

Nasjonale prøver i lesing for 8. klasse måler leseforståelsen gjennom å undersøke elevens ferdighet i å hente ut informasjon av en tekst, tolke og sette sammen informasjon, samt reflektere selvstendig omkring tekstens form og innhold. Testen inneholder to typer tekstsjangre; skjønnlitterære tekster og fagprosa. Prøvene er obligatoriske og eleven har 90 minutter på seg til å gjennomføre prøven.

Nasjonale prøver i lesing består av to oppgavetyper; flervalgsoppgaver og en åpen oppgave. En flervalgsoppgave består av et spørsmål og fire svaralternativer, der bare ett svar er riktig. En annen type flervalgsoppgave inneholder påstander som eleven må ta stilling til, der skal det krysses av for ett av to svaralternativer, for eksempel sant eller usant. En åpen oppgave består av et spørsmål som skal svares på med egne ord. Noen spørsmål ber bare om et kort svar, for eksempel et tall eller et navn. Mens andre spørsmål ber om en forklaring eller en begrunnelse. Prøven har én oppgave på sidemål og resten av oppgavene er på hovedmål. Maksimumsskåre for Nasjonale prøver i lesing er 46 poeng (Utdanningsdirektoratet).

### 3.4.2 Deltesten Ordforståelse i WISC - III

For å måle dybdevokabular eller produktivt vokabular hos elevene ble de testet med deltesten Ordforståelse fra Wechsler Intelligence Scale for Children - WISC III (Wechsler, 1991).

Hensikten med testen er å kartlegge barnets ordforståelse ved at barnet må forklare ulike ord

muntlig, dette kan gi informasjon om barnets læreevne, informasjonstilgang, idèrikdom, minne, begrepsdannelse og språkutvikling. Testen er for aldersgruppen seks til 16 år og administreres individuelt. Testleder spør eleven hva et ord betyr og noterer så ned elevens fullstendige svar. Ett godt og presist svar får en skåre på to poeng og et tilstrekkelig svar, som har et fattigere meningsinnhold enn et to poengsvar, gis en skåre på ett poeng. Testen har skåringskriterier for ett godt og ett tilstrekkelig svar, og testen avbrytes etter fem nullpoengsvar på rad. Oppgavene har en stigende vanskelighetsgrad og en maksimumsskåre på 60 poeng. Etter den standardiserte normeringen skal reliabiliteten i aldersgruppen 12 – 13 år ligge på .85 - .95.

### **3.4.3 British Picture Vocabulary Scale II**

For å måle breddevokabular eller reseptivt vokabular hos eleven ble testen British Picture Vocabulary Scale – BPVS II (Dunn, Dunn, Whetton og Burley, 1997) brukt. Testen er nylig omarbeidet fra engelsk til norsk ved Institutt for spesialpedagogikk ved Universitetet i Oslo. Dette er en individbasert test som kartlegger ordforståelse, reseptivt vokabular i aldersgruppen tre til 18 år. Materialet består av en rekke plansjer med fire nummererte tegninger. For hver plansje sier testlederen et ord, og barnet skal peke eller si nummeret på det bildet som passer til begrepet. Vanskegraden er økende, og verbal respons er ikke nødvendig. Testen er delt inn i blokker på tolv og tolv items, der hver blokk har noenlunde samme vanskegrad. Det skåres ett poeng for hvert riktige svar og testen avbrytes etter åtte eller flere feil i en blokk. Maksimumsskåre for BPVS II er 144 poeng. Etter den standardiserte normeringen skal reliabiliteten i aldersgruppen 12 – 13 år ligge på .89.

### **3.4.4 Ordkjedetesten**

Ordkjedetesten er en gruppeprøve som måler ordavkodningsferdigheter; altså de tekniske sidene ved lesing. Testen er standardisert for elever på 3. – 10. trinn og er en screeningprøve som ”fanger opp” de elevene som har vansker med å avkode ord. Testen består av 90 ordkjeder, et eksempel på en slik ordkjede er: ”trepilhvemku”. Ordene i testen er på mellom to og syv bokstaver og fra ulike ordklasser. Oppgaven består i å markere mellomrommet med tre streker slik at man ender opp med fire ord i hver ordkjede. Deltakerne skal på fire minutter markere så mange ordgrenser som mulig. I løpet av høsten 2007 ble det foretatt en ny

normering av testen. Maksimumsskåre for Ordkjedetesten er 90 poeng (Høien og Tønnesen, 2007). Reliabiliteten på denne testen er ikke oppgitt.

### 3.4.5 Raven

Raven Standard Progressive Matrices er en anerkjent og klassisk test for å måle analytisk intelligens. Analytisk intelligens referer til evnen til å kunne resonnerer og løse problemer med ny informasjon, uten i stor grad å være avhengig av en eksplisitt base av deklarativ kunnskap, som stammer fra enten skolegang eller tidligere erfaring. For å løse oppgavene i testen må man tilpasse sin tenkning til et nytt kognitivt problem (Carpenter, Just & Shell, 1990).

Hensikten med testen er å vurdere problemløsningskapasitet uten å ta i bruk verbalt språk.

Korrekt oppgaveløsning forutsetter logisk tenkning, blant annet ved å sammenlikne mønstre og mønsterkombinasjoner. Prøven er inndelt i fem serier a tolv oppgaver, disse 60 testleddene består hver av en figur hvor én del mangler, og testpersonens oppgave er å identifisere det manglende elementet blant et sett på seks til åtte valgmuligheter. Det er stigende vanskelighetsgrad både innad og på tvers av de fem seriene. Testen er i hovedsak beregnet på alderen seks til 17 år, men kan også brukes på voksne. Testen kan gjennomføres både individuelt og i gruppe. I denne undersøkelsen brukes Raven som en gruppetest.

Maksimumsskåre for Raven er 60 poeng. Etter den standardiserte normeringen skal reliabiliteten ligge på . 86 (Raven, 1986).

## 3.5 Analyse

Datamaterialet har blitt analysert gjennom statistikkprogrammet ”Statistical Packages for the Social Sciences” (SPSS), dette for å kunne gjennomføre deskriptive og analytiske statistiske beregninger. I analysen brukes parametrisk statistikk, grunn til det er at dataene er på intervall/forholdstallnivå og populasjonen kan antas å være normalfordelt (Sørensen, 2006).

Under avsnittet om deskriptiv statistikk beskrives utvalgsskårene på variablene ut i fra gjennomsnitt, standardavvik og bredden på oppnådde skårer på hver test. Fordelingene ut i fra normalitet vurderes videre gjennom statistiske metoder ved skjevhet og kurtosis, som knytter seg til spørsmålene om symmetri og spissitet av fordelingen (Tabachnick & Fidell, 2007). Det er gjennomført reliabilitetsanalyser på BPVS II og deltesten Ordforståelse i WISC III med Cronbach`s alfa koeffisient ( $\alpha$ ), grunnen til det er for å få innsyn i den indre konsistensen på

de ulike målene. Når Cronbach's Alpha har et mål som viser verdien en, viser det høy indre reliabilitet. Ved et mål som viser null, viser det en lav indre reliabilitet (Kleven, 2002b). I følge Gall et al. (2007) betraktes reliabilitet over .80 som tilstrekkelig akseptabelt i de fleste sammenhenger. Forskningsspørsmålene gjelder sammenhengen mellom undersøkte variabler, disse sammenhengene blir utforsket og drøftet ved hjelp av bivariat korrelasjonsstatistikk. Størrelsen på relasjonen mellom to variabler forklares gjennom en slik type statistikk (ibid.). I denne sammenhengen har det blitt brukt Pearson r korrelasjon, som kan regnes ut når begge variablene en ønsker å undersøke er uttrykt som kontinuerlige variabler (ibid.).

Videre brukes regresjonsanalyse for å analysere sammenhengen mellom den avhengige variabelen og de uavhengige variablene. Det er en statistisk analyseteknikk som kan uttrykke hvor mye av variansen i en avhengig variabel som kan forklares av verdier på et utvalg uavhengige variabler (Lewis-Beck, 1980). Denne analysen blir forklart med uttrykket  $R^2$  som sier noe om hvor stor del av variansen som er felles for X og Y, der X er den uavhengige variabelen og Y er den avhengige variabelen.  $R^2$  kan også være et uttrykk for forklart varians (Kleven, 2002a).

Tabachnick & Fidell (2007) har utarbeidet en formel ( $N=50+8 \times m$ ) for mål på hvor mange uavhengige variabler det er optimalt å ha i en regresjonsanalyse ut fra antall informanter (N), dette er med på å øke den statistiske kraften. I regresjonsanalysen i denne studien ble det lagt inn fire uavhengige variabler, ifølge formelen må man i denne undersøkelsen ha et minimum på 82 informanter. Siden det er 112 informanter i undersøkelsen, kan man anta at den statistiske kraften er sterk.

I den deskriptive statistikken brukes råskårer og i korrelasjonsanalysen og regresjonsanalysen er det brukt Z-skårer. Da er råskårene for hver enkelt test regnet om til standardiserte skårer i forhold til normalfordelingskurven. For å regne om råskårer til Z – skårer ble det brukt Blom's Formula korrigert for alder, denne formelen ligger i dataprogrammet SPSS. Gjennomsnittet i en normalfordelingskurve er null og standardavviket varierer fra +/- 1, men formen blir ikke endret. Z – skåren viser hvor mange standardavvik en skåre ligger over eller under gjennomsnittet i fordelingen (Christophersen, 2009). Disse to gir vanligvis en god forklaring på resultatet til hver enkelt testperson sett i forhold et større utvalg. Z – skårer gir mer nøyaktige tall og er korrigert for alderen (Gall et. al., 2007). I denne undersøkelsen varierer alderen fra 12 – 13 år.

I denne analysen ønskes det å undersøke hvor sterk sammenheng det er mellom de ulike uavhengige variablene i prediksjon på leseforståelsen, altså elevenes prestasjoner i leseforståelse som kan forklares av forskjeller i ferdigheter innen vokabular.

## 3.6 Validitet

Det er nødvendig å vurdere ulike typer validitet og reliabilitet for å kunne tolke resultater og trekke slutninger fra en undersøkelse. Hensikten med forskning er å finne en vei til sannhet som kan dokumentere at vi har forsøkt å eliminere feilkilder i forskningen på en kritisk og systematisk måte (Lund, 2002a). Anvendelsen av validitet og reliabilitet i forhold til egen undersøkelse blir drøftet i kapittel 5.

Validitet handler om hvor gyldig måleresultatet er, og graden av tillit og troverdighet uttrykkes ved å referere til høy og lav validitet (Befring, 2007). Validitet angår kvaliteten på de slutninger som trekkes fra en undersøkelse; "...the approximate truth of an inference" (Shadish, Cook og Campbell, 2002, s. 34). Cook og Campbell (1979) har utarbeidet et validitetssystem for å vurdere gyldigheten til vitenskaplige studier. Essensen i dette systemet blir her presentert, videre brukes dette som utgangspunkt for å vurdere validiteten i foreliggende undersøkelse. Opprinnelig ble dette systemet utviklet for kausale undersøkelser, men brukes nå vanligvis som en metodologisk referanseramme innen kvantitativ forskning (Lund, 2002a). De deler validitet inn i fire validitetstyper: statistisk validitet, indre validitet, begrepsvaliditet og ytre validitet (Shadish et al., 2002).

### 3.6.1 Statistisk validitet

I hvilken grad sammenhengen mellom variabler er signifikant, og i hvilken grad den er sterk, er to statistiske slutninger som angår statistisk validitet (Lund, 2002a). I denne undersøkelsen refererer statistisk validitet til gyldigheten i slutningene som trekkes angående de statistiske sammenhengene blant variablene i undersøkelsen, som er i hvilken grad det er sammenheng mellom vokabular og leseforståelse. God statistisk validitet fremheves som en nødvendig forutsetning for de andre validitetstypene (Shadish et al., 2002). Trusler mot statistisk validitet øker sannsynligheten for å gjøre en type I eller type II feil. Ved en type I feil kan vi feilaktig konkludere med at det er en sammenheng mellom variablene når det ikke er det, vi forkaster en sann nullhypotese. Ved type II feil oppstår det motsatte; vi kan konkludere med at det ikke

er en sammenheng mellom variablene når det faktisk er det, vi aksepterer en gal nullhypotese (ibid.). Ved god statistisk validitet blir statistisk signifikans avgjørende, samt at sammenhengen er sterk nok til å ha teoretisk betydning (Lund, 2002a). Det som skal til for å oppnå god statistisk validitet i en undersøkelse, er at fordelingen for variablene ikke har brudd på forutsetningen om normalitet, og at det ikke er lav statistisk styrke (ibid.). Muligheten for type II feil øker ved lav statistisk styrke. Det som er med på å svekke denne styrken er strengere signifikansnivå og større populasjonsvarians (ibid.). For å øke den statistiske styrken kan det gjøres gjennom god reliabilitet, stor utvalgsstørrelse, normalfordelt utvalg og gode måleinstrumenter (ibid.). Dette er med på å gi god statistisk validitet.

### 3.6.2 Indre validitet

Indre validitet angår slutninger om den observerte sammenhengen mellom uavhengig og avhengig variabel reflekterer en kausal sammenheng. Spørsmålet om kausalitet er den viktigste trusselen mot indre validitet (Shadish et al., 2002). Et ikke – eksperimentelt design gir ingen mulighet til å trekke slutninger om årsakssammenhenger. Den indre validiteten blir da en utfordring og kan ikke oppfylles, siden det er vanskelig å avgjøre hva som er virkning og hva som er årsak av variablene som undersøkes (Kleven, 2002a). I forbindelse med denne studien vil man for eksempel ikke vite noe om retningen i en relasjon der to variabler korrelerer. Grunnen til dette er at bak en bivariat korrelasjonskoeffisient, kan det skjule seg ulike kausalrelasjoner (Tabachnick & Fidell, 2007). En kvadrert korrelasjonskoeffisient ( $r^2$ ) kan for eksempel si at A forklarer 15 % varians av B, likevel er det umulig å vite om det er A som påvirker B, B som påvirker A eller om de påvirker hverandre (ibid.). Det finnes ikke noe endelig bevis for kausalitet i et ikke – eksperimentelt design, men man kan styrke tilliten til årsakstolkning gjennom å vise at mulige alternative tolkninger er lite sannsynlige. Letingen etter mulige skjulte variabler som kan tenkes å fungere som ”tredjevariabler” blir sentralt i jakten på ulike tolkninger som kan forenes med data (Kleven, 2002a). Hierarkisk multippel regresjonsanalyse er en statistisk metode som kan være med på å bedre den indre validiteten i et ikke – eksperimentelt design. I forklaringen av en avhengig variabel kan forskeren selv velge hvilke av de uavhengige variablene som skal prioriteres foran en annen (Tabachnick & Fidell, 2007). Når dette er sagt, hevder Kleven (2002b), at det fortsatt vil være usikkert hvorvidt det mangler relevante uavhengige variabler i analysen som kunne ført til andre resultater for de uavhengige variablene som allerede er i analysen.

Er det andre forklaringer på det statistiske resultatet? Dette spørsmålet er viktig å stille seg når det gjelder indre validitet og i denne forbindelse blir de ulike truslene mot indre validitet relevante. Trusler mot indre validitet vil være andre mulige årsaksfaktorer (Shadish et al., 2002).

### 3.6.3 Begrepsvaliditet

Det er nødvendig å arbeide både med teori og data for å kunne undersøke begrepsvaliditet, dette på grunn av at begrepsvaliditet er definert som grad av samsvar mellom begrepet slik det er definert teoretisk og begrepet slik vi lykkes med å operasjonalisere det (Kleven, 2002b). Begrepsvaliditet gjelder om de operasjonaliserte variablene måler de relevante begrepene. Begrepsvaliditet kan avgjøres for avhengig og uavhengig variabel hver for seg (Lund 2002a).

Shadish, Cook og Campbell (2002) har en videre definisjon av begrepsvaliditet; "The validity of inferences about the higher order constructs that represent sampling particulars" (s. 38).

Denne definisjonen inkluderer alt som er spesielt med utvalget, det er altså ikke bare snakk om undersøkelsens variabler. Begrepsvaliditet referer ikke bare til i hvilken grad begrepene er representative for den atferden som faktisk observeres, men også i hvilken grad personene og settingene er representative (ibid.).

Å vurdere om de elementene som registreres gjennom måleinstrumentene er representative for begrepene, altså i hvor stor grad de testene som er valgt måler de begrepene som ønskes å måles, er et uttrykk for begrepsvaliditet (Lund, 2002a). I denne studien reises spørsmålet om hvilken sammenheng det er mellom 8. klassingers vokabular og leseforståelse.

Validitetsspørsmålet som da kan stilles; er om det er benyttet gode indikatorer for vokabular og leseforståelse, altså kan vi gå ut fra at de skårene som elevene oppnår på de aktuelle testene er valide uttrykk for deres vokabular og leseforståelse?

Begrepsvaliditeten reduseres gjennom systematiske målefeil, det vil si feil som fører til konsistent skjevhet i målingen. Grunnen til det er at testene da avdekker bare enkelte sider ved det begrep som måles eller fordi irrelevante forhold blander seg inn (Kleven, 2002b).

### 3.6.4 Ytre validitet

Ytre validitet dreier seg om i hvor stor grad resultatene i undersøkelsen kan generaliseres ut over de personer, situasjoner og tider undersøkelsen gjelder (Lund, 2002b). En måte å styrke ytre validitet på er å sikre at utvalget er representativt ved å bruke et tilfeldig utvalg i populasjonen, det vil si at alle i populasjonen har lik sjanse til å bli med i utvalget; jo skjevere utvalget er i forhold til populasjonen, jo større er risikoen for at generaliseringer kan være ugyldige (Lund, 2002a.). Shadish et al. (2002) deler ytre validitet i to: generalisering til de som har vært med i utvalget og de som ikke har vært med, altså populasjonen det kan generaliseres til.

Vi vil i kapittel 5 komme nærmere inn på og drøfte de ulike validitetstypene i forhold til vår egen undersøkelse.

## 3.7 Reliabilitet

Reliabilitet viser til spørsmålet om målepresisjon (Befring, 2007). I hvilken grad data er frie for tilfeldige målefeil er et uttrykk for reliabilitet (Kleven, 2002b). Testreliabilitet handler om hvor konsistente, stabile og presise testskårene er (Gall et al., 2007). Målefeil vil alltid forekomme, og vi ville aldri fått nøyaktig samme testskårer om vi gjentok testene på de samme barna (Gall et al., 2007; Kleven, 2002b). Mange faktorer kan påvirke hvordan et barn skårer på en test, og disse faktorene kan ha lite med hvordan testen skal brukes og hvordan den utføres. Hvor lett barnet distraheres, om det er uopplagt og trøtt, i hvor stor grad svarene som barnet gir er ren gjetting og barnets følelsesmessige tilstand på det gitte tidspunktet for testen påvirker målepresisjonen. I tillegg bør det legges vekt på reliabilitetssvikt hos testadministrator. Feil i administrering eller skåring av testen fører også til tilfeldige målefeil og svekkelse av reliabiliteten.

God reliabilitet er en nødvendig, men ikke tilstrekkelig, betingelse for begrepsvaliditet og for statistisk validitet (Kleven 2002b).

Spørsmålet om reliabilitet knyttet til vår egen studie vil, i likhet med validitet, bli drøftet i kapittel 5.



## 3.8 Etske betraktninger

NESH (2006) har utviklet normer for god forskningsskikk, disse er utarbeidet for å hjelpe forskere og forskersamfunnet med å reflektere over etiske oppfatninger og holdninger. I alt vitenskapelig arbeid er det viktig at etiske hensyn er ivaretatt, og i all forskning skal forskeren arbeide ut fra en grunnleggende respekt for menneskeverdet.

Siden vår studie er knyttet opp mot KiSP - prosjektet, ivaretas etiske hensyn som anonymitet, frivillighet, informert samtykke og konfidensialitet. Når barn opptil 15 år skal delta i forskning må man ha samtykke fra foreldre/foresatte, dette for å hindre krenking av personlig integritet (Befring, 2007). Foreldre/foresatte ble på forhånd forespurt om de ønsket at deres barn skulle være med i prosjektet, og det ble skrevet under på en samtykkeerklæring. All informasjon som ble mottatt i forbindelse med prosjektet har blitt behandlet konfidensielt og testprotokoller er blitt oppbevart i nedlåst arkiv. Elevenes navn er blitt omgjort til ID – nummer, dette for å sikre anonymitet.

Barn og unge som deltar i forskning har særlig behov for og krav på beskyttelse i tråd med deres alder og behov, og man må se på barnet som et individuelt subjekt (ibid). Før undersøkelsen ble igangsatt ble deltakerne informert om prosjektet; hensikten med prosjektet, følgene av å delta og metodene som skulle benyttes, i denne forbindelse ble barna gitt informasjon som er forståelig sett i forhold til det alderstrinnet det befinner seg på.

Gall et al. (2007) påpeker at mange personer kan føle ubehag i testsituasjoner. I dette prosjektet er det mange ulike personer som har vært med på å teste elevene, i den forbindelse kan det tenkes at noen ungdommer synes det er vanskelig å forholde seg til så mange forskjellige mennesker. Det er da viktig at vi som voksne tenker igjennom hvordan vi forholder oss til testpersonene, at vi tar oss tid til å se hver enkelt elev og lar de føle seg ivaretatt. Testingen av elevene medførte ingen ulempe, skade eller risiko for dem. Som testledere har vi lagt til rette for at barna skulle få en følelse av mestring og få en god opplevelse av testsituasjonen.

## 4 Resultater

I dette kapitlet presenteres resultatene fra foreliggende undersøkelse. Først beskrives resultatene ut fra deskriptiv statistikk, dette for å få en oversikt over faktiske funn. Den deskriptive statistikken beskrives ut fra dataenes råskåre. Videre brukes korrelasjonsanalyse, for å se styrken på forholdet mellom variabler. Deretter gjøres en regresjonsanalyse for å avdekke det unike bidraget hver enkelt uavhengige variabel har på den avhengige variabelen. I den analytiske statistikken brukes dataenes  $Z$  – skåre.

### 4.1 Deskriptiv statistikk

Deskriptiv statistikk er teknikker som brukes for å få oversikt over og kartlegge hovedtendenser i et tallmateriale. Deskriptiv statistikk beskriver den informasjonen vi rent faktisk har funnet i vår undersøkelse – vårt utvalg (Sørensen, 2006). I det neste presenteres og kommenteres de deskriptive resultatene fra testene som er gjennomført; antall personer som har gjennomført testene på hver variabel, gjennomsnitt, standardavvik, alfaverdi, skjevhet, kurtosis, minimum - og maksimumsverdier.

Resultatene presenteres med grunnlag i råskårer, dette på grunn av at disse har mer informasjonsverdi enn  $Z$  – skårer og tallene er enklere å forstå.

Tabell 1 viser at gjennomsnittet på Nasjonale prøver i lesing er på 33, med et standardavvik på 8.39 og minimums – og maksimumsverdiene varierer fra fire til 46. Med en skjevhetsverdi på  $-1.23$  og en kurtosisverdi på  $1.55$  er det noe avvik fra normalfordelingen. Én elev har skåret 46 poeng, som er maksimumskåre på denne testen, dette fører ikke til noen tak - effekt.

Deltesten Ordforståelse i WISC III har en gjennomsnittsverdi på  $36.57$ , der minimums – og maksimumsverdiene varierer fra  $13 - 56$  og standardavviket er på  $7.18$ . Gjennomsnittet er forholdsvis høyt og skårene er likt fordelt rundt den mest typiske verdien; med en skjevhetsverdi på  $-0.13$ . Kurtosisverdien er på  $.48$ , dette viser en liten forekomst av ekstremverdier i forhold til normalfordelingen. Reliabilitetsanalyse er gjennomført på deltesten Ordforståelse i WISC III, denne viser en alfaverdi på  $.80$ , som er noe lavere en alfaverdien oppgitt i normeringsgruppa, men likevel innenfor det akseptable. Ifølge normeringen av testen i alderen  $12 - 13$  år, skal alfaverdien ligge på henholdsvis  $.85$  og  $.90$ .

**Tabell 1 Deskriptive resultater for målte variabler**

Variabel	N	M	SD	$\alpha$	Skew	Krt	Min	Max
Np i les	108	33	8.39		-1.23	1.55	4	46
WISC	108	36.57	7.18	.80	-.13	.48	13	56
BPVS	108	121.82	9	.87	-1.70	4.67	81	137
Ordkj	98	46.8	15.28		.68	.32	18	90
Raven	111	44.88	7.78		-1.43	3.00	14	56

*Note.* N= Antall personer; M = Gjennomsnitt; SD = Standardavvik;  $\alpha$  = Alfaverdi; Skew = Skjevhet; Krt = Kurtosis; Min = Minimum skåret verdi; Max = Maksimum skåret verdi; S40 = Setningsleseprøve; NP i les = Nasjonale prøver i lesing; Deltesten Ordforståelse i WISC III = Wechsler Intelligence Scale for Children – deltest: ordforståelse; BPVS II = British Picture Vocabulary Scale; Ord kj = Ordkjedetesten.

Gjennomsnittet på BPVS II er 121.82, minimums – og maksimumsverdiene varierer fra 81 – 137 og standardavvik ni. Kurtosisverdien er på 4.67, noe som avviker sterkt fra normalfordelingen, dette forteller oss om en relativt stor hyppighet av ekstremverdier sett i forhold til en normal fordelt kurve. Skjevhetsverdien på  $-1.70$  viser oss små avvik tatt i betraktning at enhetene er likt fordelt rundt den mest typiske verdien. De fleste elevene har skåret bra på denne testen, men på grunn av noen uteliggere, som har ekstreme verdier i forhold til resten av utvalget, blir kurtosisverdien såpass høy, og normalfordelingen påvirkes negativt på grunn av dette. Med en alfaverdi på  $.87$  viser dette målet en tilfredsstillende reliabilitet.

Ordkjedetesten har en gjennomsnittsverdi på 46.80, et standardavvik på 15.28 og minimums – og maksimumsverdiene varierer fra 18 – 90. Skjevhets – og kurtosisverdier på henholdsvis  $.68$  og  $.32$ . Disse verdiene er godt innenfor hva Christophersen (2006) betegner som lite avvik fra normalfordelingen. Siden det bare er én elev som har oppnådd høyest mulig poengsum (90), viser testen ingen takeffekt.

Gjennomsnittet på Raven er 44.88, har et standardavvik på 7.78 og variasjonen i minimums – og maksimumsverdier er 14 – 59. Skjevhetsverdien er på – 1.43, dette viser små avvik i forhold til at enhetene er likt fordelt rundt den mest typiske verdien. En kurtosisverdi på tre forteller oss om en sterk hyppighet av ekstremverdier i forhold til normalfordelingen. Vi ser det samme her som på BPVS II; at noen uteliggere med ekstremverdier i forhold til resten av utvalget, gjør at råskårefordelingen ikke vil bli normal.

Skjevhet - og kurtosisverdier lik null indikerer en perfekt normalfordeling. Er verdiene fra – 1 til 1 er det små avvik fra normalfordelingen, mens – 3 til 3 kan betraktes som sterke avvik (Christophersen, 2006). Tabellen viser høye/lave verdier på skjevhet og kurtosis, på grunn av at det i den deskriptive statistikken er tatt utgangspunkt i råskårer. Når Z – skårer brukes blir tallene mindre og de er korrigert for alderen, dette er igjen med på å minimalisere uteliggerne. Ved bruk av Z – skårer blir verdien på skjevhet -.018 og verdien på kurtosis -.251 på Nasjonale prøver i lesing. BPVS II får skjevhet- og kurtosisverdi på henholdsvis .005 og -.284. Når det gjelder Raven får den en skjevhetsverdi på .004 og en kurtosisverdi på -.204.

## 4.2 Reliabilitetsanalyse

Det er gjort en reliabilitetstest for to av variablene i denne undersøkelsen, deltesten Ordforståelse i WISC III og BPVS II. Grunnen til dette er for å finne ut i hvilken grad dataene er frie for tilfeldige målefeil. Det er da brukt Cronbach's Alpha som reliabilitetsmål. Når Cronbach's Alpha har et mål som viser verdien en, viser det høy indre reliabilitet. Ved et mål som viser null, viser det en lav indre reliabilitet (Kleven, 2002b). Resultatene i denne undersøkelsen viser at Cronbach's Alpha i deltesten Ordforståelse i WISC III ligger på .80 og i BPVS II på .87.

Dette er to standardiserte tester og i skåringsmanualen har deltesten Ordforståelse i WISC III en reliabilitet .85-.95 og BPVS II har en reliabilitet på .89. Skåring av testene er blitt utført og skåret av flere personer tilknyttet KiSP – prosjektet. BPVS II er en enkel test å skåre og små muligheter til å gjøre feil. Men ved skåring av deltesten Ordforståelse i WISC III gir det et lite rom for vurderinger som kan føre til feil i resultatene.

### 4.3 Korrelasjoner

Pearson korrelasjoner brukes når man vil se på styrken på forholdet mellom to variabler (Pallant, 2007). Ingen sammenheng mellom variablene betegnes ved en r skåre på null, mens en r skåre på  $-1$  eller  $+1$  indikerer en perfekt forutsigelse av en skåre når man kjenner til den andre.  $-1$  betyr komplett negativ sammenheng, mens  $+1$  betyr komplett positiv sammenheng (Gall et al. 2006).

I forhold til problemstillingen i denne undersøkelsen der man ønsker å se på sammenhengen mellom vokabular og leseforståelse, vil vi ha hovedfokus på korrelasjonen mellom disse to. Nasjonale prøver i lesing tester leseforståelse, variabelen ordavkodning måles med ordkjedetesten, BPVS II og deltesten Ordforståelse i WISC III tester henholdsvis bredde – og dybdevokabular og generell kognitiv fungering blir målt med testen Raven. I denne korrelasjonsanalysen er det brukt Z-skårer som gir mer nøyaktige tall og er korrigert for alderen.

**Tabell 2 Korrelasjoner mellom alle målte variabler.**

Variabel	1	2	3	4	5
1 NP i les	-				
2 Raven	.64**	-			
3 Ordkj	.31**	.13	-		
4 BPVS	.46**	.46**	.14	-	
5 WISC	.45**	.38**	.05	.49**	-

*Note.* S40 = Setningsleseprøve; NP i les = Nasjonale prøver i lesing; Deltesten Ordforståelse i WISC III = Wechsler Intelligence Scale for Children – deltest: ordforståelse; BPVS II = British Picture Vocabulary Scale; Ordkj = Ordkjedetesten.

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

Tabell 2 viser at Nasjonale prøver i lesing korrelerer signifikant med både deltesten Ordforståelse i WISC III og BPVS II med et signifikansnivå på 1 %. Dette forklarer at leseforståelse ut fra de Nasjonale prøvene i lesing viser en god sammenheng med både bredde- og dybdevokabular.

Ordkjedetesten som måler avkodingsferdigheter viser en signifikant korrelasjon med Nasjonale prøver i lesing med et signifikansnivå på 1 %. Dette er i samsvar med viktigheten av å ha gode avkodingsferdigheter, som igjen er en forutsetning for en god leseforståelse. Ordkjedetesten korrelerer svakt, men ikke signifikant, med BPVS II og deltesten Ordforståelse i WISC III, det kan være grunn til å anta at Ordkjedetesten ikke tapper vokabularkunnskap på samme måte som Nasjonale prøver i lesing.

Raven, som måler generell kognitiv fungering, korrelerer sterkt med Nasjonale prøver i lesing og vokabulartestene BPVS II og deltesten Ordforståelse i WISC III. Raven brukes i denne undersøkelsen som en test på generell kognitiv fungering, og gjennom sterke korrelasjoner viser resultatene samsvar og er reelle sett i forhold til elevenes kognitive evner. Det er ingen signifikant korrelasjon mellom Raven og Ordkjedetesten. En grunn til dette kan være at Ordkjedetesten i mindre grad enn de Nasjonale prøvene tapper kognitive evner, derfor er det bare en svak korrelasjon mellom disse to variablene.

## 4.4 Regresjonsanalyse

Multiple regresjonsanalyser er i hovedsak beregnet på analyse av to eller flere uavhengige variabler opp mot den avhengige variabel (Pedhazur, 1997). Men, som Pedhazur (1997) understreker, kan det i et ikke – eksperimentelt design være et problem at de uavhengige variablene korrelerer med hverandre, dette kan gjøre det vanskelig å avdekke det unike bidraget hver av dem har på den avhengige variabel. For å finne fellesvarians og unikheter fra de uavhengige variablene, kommer dette an på størrelsen på korrelasjonen mellom variablene, er det stor korrelasjon mellom variablene vil fellesvariansen av forklaringen på kriterievariabelen være større og det unike bli mindre (ibid.).

I denne undersøkelsen er leseforståelse den avhengige variabel, den vil bli analysert i forhold til bredde – og dybdevokabular, som er de uavhengige variablene. Dette gjøres for å finne ut om bredde – og dybdevokabular har ulik innvirkning på leseforståelsen. I tabell 3, modell 1 brukes Nasjonale prøver i lesing som avhengig variabel, denne blir vurdert opp mot Raven,

Ordkjedetesten, BPVS II og deltesten Ordforståelse i WISC III. I modell 2 brukes samme avhengige variabel, vurdert opp mot Raven, Ordkjedetesten, deltesten Ordforståelse i WISC III og BPVS II.

I en hierarkisk regresjonsanalyse må rekkefølgen på de uavhengige variablene bestemmes på forhånd før de legges inn. Etter hvert som flere prediktorvariabler føres inn i analysen, viser R square hvor mye de uavhengige variablene til sammen forklarer av den avhengige variabel.  $R^2$  viser variansen hver prediktorvariabel forklarer av kriterievariablen knyttet til det steget som prediktorvariablen er ført inn i analysen på. Hvor mange uavhengige variabler som inngår i regresjonsmodellen, tar dermed  $R^2$  hensyn til. I denne undersøkelsen ble denne rekkefølgen valgt: Raven, Ordkjedetesten, BPVS II og deltesten Ordforståelse i WISC III i modell 1 og Raven, Ordkjedetesten, deltesten Ordforståelse i WISC III og BPVS II i modell 2. Grunnen til denne rekkefølgen er at man kan anta at den generelle kognitive fungeringen og avkodingsferdigheter har betydning for leseforståelse

**Tabell 3 Hierarkiske regresjonsanalyser**

Steg	Prediktor	R square	$R^2$	<i>P</i>
Modell 1				
1	Raven	.458	.458	.000
2	Ordkjeder	.492	.045	.005
3	BPVS	.508	.021	.051
4	WISC	.520	.018	.069
Modell 2				
3	WISC	.520	.032	.014
4	BPVS	.520	.006	.295

*Note.* N= Antall personer;  $R^2$  = Rsquare Change; *P*= Significance F Change.

I tabell 3 viser modell 1 varians av Raven, Ordkjedetesten, BPVS II og deltesten Ordforståelse i WISC III som prediksjon på leseforståelse målt ved Nasjonale prøver i lesing. Modell 2 viser varians av Raven, Ordkjedetesten, deltesten Ordforståelse i WISC III og BPVS II som prediksjon på leseforståelse målt ved Nasjonale prøver i lesing.

I regresjonsanalysen brukes også Z – skårer, på samme måte som i korrelasjonsanalysen. Resultatene viser at Raven og ordkjedetesten i tabell 3, modell 1 viser lave  $p$  verdier, med et signifikansnivå satt til 1 %. BPVS II er på grensen til signifikant og deltesten Ordforståelse i WISC III viser et ikke-signifikant resultat. Resultatene for prediksjon av leseforståelse testet med Nasjonale prøver i lesing, viser at Raven, som tester generell kognitiv fungering, forklarer 45,8 % av variansen. Ordkjedetesten, som tester avkoding, har en unik forklaringsvariens på 4,5 % for leseforståelse etter at det er kontrollert for Raven. Mens BPVS II, som måler breddevokabular, unikt forklarer 2,1 % etter at det er kontrollert for Raven og Ordkjedetesten. Når deltesten Ordforståelse i WISC III er kontrollert for Raven, Ordkjedetesten og BPVS II har den en unik forklaringsvariens på 1,8 %.

I modell 2 får deltesten Ordforståelse i WISC III et signifikansnivå på under 5 % når den kommer før BPVS II i regresjonsanalysen. Den unike forklaringsvariensen for deltesten Ordforståelse i WISC III, som måler dybdevokabular, er 3,2 % når den er kontrollert for Raven og Ordkjedetesten. BPVS II har en unik forklaringsvariens på 0,6 % etter at den er kontrollert for Raven, Ordkjedetesten og deltesten Ordforståelse i WISC III.

Resultatene viser at dybdevokabular er signifikant og breddevokabular er tilnærmet signifikant, dybdevokabular har en større forklaringsvariens til leseforståelsen enn det breddevokabular har. Når BPVS II og deltesten Ordforståelse i WISC III er kontrollert for hverandre, viser disse ingen signifikans.

## 4.5 Oppsummering av resultater

Resultatene fra undersøkelsen viser at for å måle vokabularets betydning for leseforståelse, kommer det her frem at vokabulartestene BPVS II og deltesten Ordforståelse i WISC III har sterk korrelasjon til leseforståelsen når den testes med Nasjonale prøver i lesing. Samtidig viser regresjonsanalysen at det er deltesten Ordforståelse i WISC III, som måler



dybdevokabular, som har den største unike forklaringsvariansen for leseforståelsen sett i forhold til breddevokabular målt med BPVS II. Resultatene forteller oss også at ordavkoding, målt med Ordkjedetesten, har en større unik forklaringsvarians for leseforståelsen målt med Nasjonale prøver i lesing, enn det dybdevokabular, målt med deltesten Ordforståelse i WISC III og breddevokabular målt med testen BPVS II har. Resultatene viser en sterk sammenheng med Nasjonale prøver i lesing og Raven, som måler elevens kognitive evner. Dette forteller oss at Nasjonale prøver har et reelt samsvar med elevens generelle kognitive fungering.

## 5 Drøfting

I dette kapittelet vil problemstillingen, forskningsspørsmålene og resultater diskuteres i lys av teori og tidligere forskning. Resultatene vil deretter bli vurdert i forhold til aktuelle aspekter ved validitetstypene til Cook & Campbell (1979). Reliabilitet vil inngå i dette.

### 5.1 Hvilken innvirkning har vokabulardybde på leseforståelsen blant 8. klassinger?

Ut fra det første forskningsspørsmålet var det ønskelig å finne ut hvilken innvirkning vokabulardybde har på leseforståelsen blant elever i 8. klasse. I denne undersøkelsen viser resultatene at når leseforståelse ble testet med Nasjonale prøver i lesing viser regresjonsanalysen at deltesten Ordforståelse i WISC III, som tester dybdevokabular, har en unik forklaringsvarians på 3,2 % etter at den har blitt kontrollert for Raven og Ordkjedetesten, og et signifikansnivå satt til 1 %. Breddevokabular, målt med BPVS II, har en unik forklaringsvarians på 2,1 %, etter at det er kontrollert for Raven og Ordkjedetesten, og er nær, men ikke signifikant med. Siden det her tas utgangspunkt i resultatene før bredde – og dybdevokabular er kontrollert for hverandre, vil det samme bli gjort når det gjelder drøftingen av denne undersøkelsen opp mot Ouellette (2006) sin undersøkelse

At dybdevokabular hadde en større forklaringsvarians for leseforståelsen enn breddevokabular er ikke overraskende i forhold til det som tidligere er nevnt i teorien og empirien. Når det gjelder tidligere forskning på feltet viser flere studier at dybdevokabular i større grad enn breddevokabular predikerer leseforståelse i en positiv retning (Muter, Hulme, Snowling og Stevenson, 2004; Nation og Snowling, 2004; Roth, Speece og Cooper, 2002; Sènèchal, Ouellette og Rodney, 2006; Share og Leikin, 2004; Snow, Tabors, Nicholson og Kurland, 1995). En av de nyeste undersøkelsene innenfor dette feltet er den Ouellette (2006) gjennomførte. Han fant i sin undersøkelse at dybdevokabular har en unik forklaringsvarians på 12,1 % etter at det er kontrollert for alder, non – verbal intelligens, avkoding og visuell ordgjenkjenning, og signifikansnivået er satt til 0,1 %. Breddevokabular har en unik forklaringsvarians på 6,1 % etter at det er kontrollert for alder, non – verbal intelligens, avkoding og visuell ordgjenkjenning og signifikansnivået er satt til 5 % (ibid.). Dette samsvarer i stor grad med de resultatene denne undersøkelsen viser.

Begge undersøkelsene har i forhold til måling av dybdevokabularet evaluert elevenes evne til muntlig å forklare ord, og breddevokabularet måles gjennom enkeltordgjenkjennelse (ibid.). Forskjellen på Ouellette (2006) sin undersøkelse og denne studien her er at hans utvalg består av 4. klassinger, mens utvalget i denne undersøkelsen er 8. klassinger. Likevel ser vi at det er dybdevokabular som har sterkest forklaringsvarians på leseforståelsen uavhengig av alder i disse to undersøkelsene. Grunnen til dette kan være at elever i 4. klasse begynner og nå et såpass høyt nivå i skolesammenheng at mesteparten av breddevokabularet er lagt til grunn. Lesing blir i større og større grad brukt som et redskap i kunnskapstilegnelse og oppgaveløsning. Dybdevokabularet blir dermed mer og mer utfordret, med tanke på at i 4. klasse begynner fagtekstene å øke i vanskelighetsgrad og det stilles større krav til elevenes leseforståelse.

Logisk sett vil man si at kunnskap innen dybdevokabular øker i takt med barnets alder. Barn kan lagre et ord i sitt vokabular uten helt å forstå betydningen av ordet. Dette utvider barnets vokabularbredde, men ikke nødvendigvis vokabulardybden. Ettersom barnet blir eldre vil ordbetydningen bli videreutviklet; ordet vil bli presisert, utdypet og forbedret, noe som vil være med på å øke barnets dybdevokabular (Lahey, 1988). Den unike forklaringsvariansen på dybde - og breddevokabular i denne undersøkelsen samsvarer i forhold til dette. Altså; i forhold til vokabularutviklingen er det naturlig at dybdevokabularet har en mer positiv effekt på leseforståelse enn det breddevokabularet har. Det samme kommer frem gjennom Stahl og Nagy (2006) sin ”vokabularpyramide”, der utviklingen av vokabularet deles inn i tre nivåer og er organisert slik at den går fra breddevokabular til dybdevokabular. På nivå en er det innlæring av høyfrekvente ord det legges størst vekt på, på det neste nivået skal elevene gis økt kunnskap om det spesifikke ordet og videre skal eleven føres til det siste nivået, der de skal gis den komplette forståelsen av ordet. En fullstendig forståelse av et ord vil igjen være til stor hjelp når tekster øker i vanskelighetsgrad og kravet til leseforståelse blir høyere. Alle disse nivåene er like viktige når det gjelder utvikling av vokabular (ibid.).

En annen forskjell på Ouellette (2006) sin undersøkelse og denne undersøkelsen kan være forskjellen i vanskelighetsgraden på leseprøvene som brukes for å teste leseforståelse og hva de er ment å måle. Ouellette (2006) bruker deltesten The Passage Comprehension i Woodcock Reading Mastery Tests – Revised (Woodcock, 1998). I denne testen skal eleven lese en kort setning stille, deretter verbalt gi et svar på en skrevet setning som er basert på innholdet i den setningen eleven først leste (ibid.). Denne undersøkelsen bruker Nasjonale prøver i lesing som

mål på leseforståelse. Her stilles det krav til at eleven skal hente ut informasjon av en tekst, tolke og sette sammen informasjon, samt reflektere selvstendig omkring tekstens form og innhold. Prøven består av to oppgavetyper; flervalgsoppgaver og åpne oppgaver. Ut fra dette kan man mest sannsynlig trekke den slutning at den testen Ouellette (2006) bruker, er en enklere leseprøve enn det Nasjonale prøver i lesing er. Denne undersøkelsen bruker testen Raven som mål på non – verbal intelligens, Ouellette (2006) bruker Test of Nonverbal Intelligence (Brown, Sherbenou & Johnsen, 1997). Sistnevnte test gjennomføres ved at eleven blir presentert for ulike mønstre som inneholder geometriske figurer, og de skal plukke ut den figuren som gjør mønsteret komplett (ibid.). Denne testen skiller seg ikke nevneverdig fra Raven's progressive matriser, som går ut på det samme prinsippet; der hvert testledd består av en figur hvor én del mangler, elevens oppgave er å identifisere det manglende element (Raven, 1986). Disse to testene er forholdsvis like når det gjelder både utførelse og oppgavetype. Undersøkelsen til Ouellette (2006) viser at den unike forklaringsvariansen på dybdevokabular er 12,1 % etter at det er kontrollert for alder, non – verbal intelligens, avkoding og visuell ordgjenkjennelse og den unike forklaringsvariansen på non – verbal intelligens er på 10,5 % etter at det er kontrollert for alder.

Når man i denne undersøkelsen vurderer den unike forklaringsvariansen til deltesten Ordforståelse i WISC III som er på 3,2 %, etter at det er kontrollert for Raven og ordkjedetesten, opp mot den unike forklaringsvariansen på Raven som er på 45,8 %, ser man at det er en stor forskjell sammenlignet med undersøkelsen til Ouellette (2006). En forklaring på dette kan være, som nevnt tidligere, at Nasjonale prøver i lesing sannsynligvis er mer krevende enn den testen Ouellette (2006) brukte som mål på leseforståelse. Når leseforståelse måles med Nasjonale prøver i lesing krever disse så mye av elevenes kognitive evner som igjen fører til at dybdevokabularet får en såpass liten forklaringsvarians på leseforståelsen sett i forhold til Raven. I Ouellette (2006) sin undersøkelse er ikke forskjellen i variansen mellom dybdevokabular og non – verbal intelligens så stor, dette kan trolig forklares med at leseprøven som Ouellette (2006) bruker i mindre grad, enn Nasjonale prøver i lesning, stiller krav til elevenes generelle kognitive fungering. Forholdet mellom leseforståelse målt ved hjelp av Nasjonale prøver i lesing og generell kognitiv fungering, målt med Raven, vil bli drøftet senere i kapittelet.

Utvalget i denne undersøkelsen er 8. klassinger og jo eldre barnet blir, desto større krav stilles det til lesingen. Manglende kjennskap til anvendte ord og forkunnskaper innenfor de

fagområder det jobbes med i skolen er en barriere for leseforståelsen. Man kan anta at de fleste barn klarer seg bra i lesing på de tidlige klassetrinn, men kan komme i vanskeligheter når innholdet i tekster de møter senere i sin skolegang stiller større krav til deres vokabular (Pikulski, 2006). Tidlige lesevansker kan vise seg etter 4. klasse, grunnen til det er at da jobber elevene mer dyptgående og oftere med fagprosa innenfor de ulike fagområdene (Snow et al., 1998). Som nevnt tidligere kan det tyde på at Nasjonale prøver i lesing stiller høye krav til den kognitive funksjonen, men tatt dette i betraktning skårer utvalget i denne undersøkelsen i overkant av gjennomsnittet sett i forhold til landsnormeringen på Nasjonale prøver i lesing.

## **5.2 Hvilken innvirkning har vokabularbredde på leseforståelsen blant 8. klassinger?**

I det andre forskningsspørsmålet vil vi finne ut av breddevokabularet sin innvirkning på leseforståelsen blant 8. klassinger. Resultatene fra denne undersøkelsen viser at dybdevokabular har en større forklaringsvarians til leseforståelse enn breddevokabular. Det viser seg i regresjonsanalysen at BPVS II, som måler breddevokabular, har en unik forklaringsvarians på 2,1 % og er nær et signifikansnivå på 5 %. I regresjonsanalysen er BPVS II kontrollert for Raven, som er en generell kognitiv evnetest, og for ordkjetdetesten, som tester avkoding. Ser man så på denne sammenhengen ut i fra korrelasjonsanalysen, korrelerer BPVS II ganske høyt med leseforståelse og får da et signifikansnivå på 1 %. Men både breddevokabular og dybdevokabular har en signifikant sammenheng med leseforståelsen ut fra korrelasjonsanalysen.

I undersøkelsen til Tannenbaum, Torgersen og Wagner (2005) fant de at breddevokabular hadde størst varians til leseforståelsen hos 3. klassinger, men at både bredde- og dybdevokabular var signifikante til leseforståelsen, men den unike variansen til dybdevokabular var liten. Den unikt høye forklaringsvariansen på breddevokabular målt opp mot leseforståelse i undersøkelsen til Tannenbaum et al. (2005), er mye større enn hva som er funnet i andre undersøkelser. Dette kan indikere at forholdet mellom breddevokabular og leseforståelse er av større betydning enn hva tidligere forskning har funnet før. De mener at for å få en best mulig leseforståelse vil det bli best når breddevokabular og dybdevokabular/ordgjennkalling er til stede (Tannenbaum et al., 2005). Konklusjonen fra

undersøkelsen til Tannenbaum et al. (2005) var at leseforståelsen økte mest når breddevokabularet økte.

Som nevnt tidligere er det i denne undersøkelsen funnet at det er dybdevokabular som har den største forklaringsvariansen. Det kan være ulike årsaker til dette, en av de kan være de ulike klasstrinnene som undersøkelsene er blitt utført i. I undersøkelsen til Tannenbaum et al. (2005) viser breddevokabular en sterk forklaringsvariens til leseforståelsen målt i 3. klasse. Man kan av den grunn anta at breddevokabular har en sterkere innvirkning på lesing på lavere klasstrinn. I vår undersøkelse finner vi at dybdevokabular har noe større betydning enn breddevokabular, men ingen av delene har en sterk forklaringsvariens i forhold til leseforståelse. På høyere klasstrinn kan man anta at det kan være andre faktorer som spiller en viktigere rolle enn vokabular. En variabel som er tatt med i denne undersøkelsen er generell kognitive evner, målt med testen Raven. Denne testen har en større forklaringsvariens til leseforståelse, målt med Nasjonale prøver i lesing, enn hva breddevokabular har til leseforståelsen. En forklaring på dette kan være at testen Nasjonale prøver i lesing tapper mye av eleven sine generelle kognitive evner, derfor får testen Raven en så stor forklaringsvariens.

I denne undersøkelsen er testen BPVS II brukt for å måle vokabularbredde, men i undersøkelsen til Tannenbaum et al. (2005) brukte de PPVT – III (Dunn & Dunn, 1981) som er den amerikanske versjonen av BPVS II, og deltesten Ordforståelse i WISC III for å måle vokabularbredden. Ved at disse to undersøkelsene bruker tester ulikt til å måle breddevokabular, kan det antas at det har vært med på å gi ulikt utslag på resultatet i denne undersøkelsen og undersøkelsen til Tannenbaum et al.(2005). For å måle dybdevokabular brukte Tannenbaum et al. (2005) testene Multiple Meanings and Attributes subtest of the Language Processing Test-Revised (Richard & Hanner, 1995). Multiple Meanings er en test hvor elevene skal huske og definere ord som blir presentert i ulike kontekster, et ord skal passe til tre setninger. I testen Attributes skal elevene benevne så mange ord de klarer på ti sekunder innenfor ulike kategorier som for eksempel farger, dyr osv. Mens i denne undersøkelsen er det brukt deltesten Ordforståelse i WISC III for å måle av dybdevokabular Der skal elevene forklare eller gi synonymer for et ord muntlig. Ved måling av vokabulardybde måler også her testene noe ulikt, man kan anta at dette kan være med på at man i undersøkelsen til Tannenbaum et al. (2005) har fått resultater der dybdevokabular har en liten unik variens til leseforståelsen.

Breddevokabular er viktig selv om det ikke i denne undersøkelsen har en stor unik forklaringsvarians til leseforståelse, men i korrelasjonsanalysen korrelerer breddevokabular sterkt til leseforståelsen, det kan da antas at breddevokabular spiller en viktig rolle for leseforståelsen. Breddevokabular er byggesteinen til dybdevokabular. Barn lagrer et ord i sitt vokabular uten helt å forstå fullt ut betydningen/betydningene av ordet. Over tid vil betydningen bli videreutviklet; ordet vil bli forbedret, mer presist og utdypet (Lahey, 1988). Dette er med på å bidra til økning i barnets vokabulardybde.

### 5.3 Leseforståelse og generelle kognitive evner

Korrelasjonsanalysen i denne undersøkelsen viser at Raven, som er antatt å måle generell kognitiv fungering, korrelerer sterkt (.64) med leseforståelse når den er målt med Nasjonale prøver i lesing, og signifikansnivået er satt til 1 %. Den sterke korrelasjonen viser at resultatene samsvarer og er reelle sett i forhold til elevenes kognitive evner.

Craver (1990) fant i sin undersøkelse at korrelasjonen mellom skårer på Raven og høye reliable mål på leseforståelse varierte mellom .40 - .60 i 2. – 12. klasse, dette er i samsvar med det som er funnet i denne undersøkelsen. Ved å bruke Raven så eliminerer den faktoren lesing for å måle intelligens, testen måler non – verbal intelligens. En grunn til at Raven korrelerer sterkt med leseforståelsen hos 8. klassingene i denne undersøkelsen, kan være at leseforståelse på et såpass høyt nivå forutsetter mer av de samme ferdighetene som ligger til grunn for lytteforståelse i tidlige år; bruk av bakgrunnskunnskap, ferdigheter i inferens, hukommelsesstrategier og vokabular (Sticht, 1979).

Regresjonsanalysen i denne undersøkelsen viser at Raven har en unik forklaringsvarians på 45,8 % på leseforståelsen når den måles med Nasjonale prøver i lesing. Raven er lagt først inn i regresjonsanalysen, og av den grunn er den ikke kontrollert for verken avkoding eller bredde – og dybdevokabular. Det foreligger ingen regresjonsanalyse i forbindelse med Craver (1990) sin undersøkelse. Denne sterke forklaringsvariansen og den høye korrelasjonen, kan fortelle oss at Nasjonale prøver i lesing tapper i stor grad elevenes generelle kognitive fungering. Det finnes mange ulike tester som kan brukes som mål på leseforståelse, disse er bygd opp på forskjellige måter; noen er omfattende og krevende, mens andre igjen er enklere og i mindre grad vanskelige. Testen Craver (1990) bruker i sin undersøkelse som mål på leseforståelse, National Reading Standard Test, vil man kunne anta er enklere i sin form og

mindre krevende enn Nasjonale prøver i lesing. Førstnevnte test går ut på at testpersonen for hvert femte ord i en tekst får en valgmulighet på to ord, der bare det ene skal settes inn i teksten og stemme overens med resten av setningen (Craver, 1977). Testen Finbak (1998) bruker i sin undersøkelse for å måle leseforståelse er Carlstens leseprøve (1993). Denne testen går ut på stort sett det samme prinsippet som den testen Craver (1990) bruker i sin undersøkelse; enkelte ord er utelatt fra teksten og erstattet med en parentes med tre ord, der det ene ordet er det utelatte (Carlsten, 1993). Finbak (1998) fant i sin undersøkelse en korrelasjon på .15 mellom leseforståelse målt med Carlstens leseprøve og generell kognitiv fungering målt med Raven. Denne korrelasjonen er betraktelig lavere enn korrelasjonen i foreliggende undersøkelse. Nasjonale prøver i lesing er mer omfattende; disse måler leseforståelsen gjennom å undersøke elevens ferdighet i å hente ut informasjon av en tekst, tolke og sette sammen informasjon, samt reflektere selvstendig omkring tekstens form og innhold. Prøven består av to oppgavetyper; flervalgsoppgaver og åpen oppgave (Utdanningsdirektoratet). Ut fra dette kan være grunn til å tro at Nasjonale prøver i lesing krever mye av elevens kognitive evner og at Raven sin unike forklaringsvarians er reell.

## 5.4 Leseforståelse, vokabular og ordavkodning

Vokabularet er fundamentet til leseforståelsen, det er vanskelig å forstå en tekst om man ikke vet hva de fleste ord betyr (Nagy, 1988). Avkodning er også en viktig ferdighet i forhold til lesing. Den har en mer teknisk side som bygger på flere delferdigheter. Disse er viktig for at leseren skal kunne kjenne igjen ord eller nonsensord og setter dem sammen til et "lydbilde" (Molander og Skauge, 2009). Just og Carpenter (1987) fastslår at det er leserens bredde- og dybdevokabular som har effekt på leseforståelse. Vellutino (2003) sier det samme om ordavkodning og leseforståelse, og at det er en godt dokumentert sammenheng mellom de to. Er det også mulig å si at det er en sammenheng mellom leseforståelse, vokabular og ordavkodning, og ikke bare mellom leseforståelse og en av variablene ordavkodning eller vokabular?

I denne undersøkelsen har Ordkjedetesten en unik forklaringsvarians på 4,5 % til leseforståelsen målt med Nasjonale prøver i lesing etter at den var kontrollert for Raven. Det er med på å vise at ordavkodning har et bidrag til leseforståelsen. I korrelasjonsanalysen korrelerer Ordkjedetesten noe lavere enn testene for vokabular, BPVS II og deltesten Ordforståelse i WISC III og ordkjedetesten har et signifikant resultat. En grunn til det kan



være at denne undersøkelsen er utført på 8. klasse elever, og ved et så høyt klassetrinn kan en anta at vokabular har en større sammenheng med leseforståelsen og en kan forvente at ordavkodingsferdighetene er automatisert.

Petersen (2008) har gjort en studie på hva som hadde betydning for leseforståelsen. Det som bidro mest i hennes undersøkelse til leseforståelsen var vokabular og ordavkoding. Tallene viste at vokabular forklarte mer av variansen enn ordavkoding, men at disse to til sammen hadde den unike forklaringsvariansen på hele 45 % til leseforståelse. Dette viser at det er en sammenheng mellom leseforståelse, vokabular og ordavkoding. Det som kan ha betydning for hvor stor denne sammenhengen er, kommer an på hvor i leseopplæringen eleven er. I denne undersøkelsen er det brukt elever fra 8. klasse og i undersøkelsen til Petersen (2008) er det brukt elever fra 7. – 10. klasse. Det viser seg i begge undersøkelsene at ordavkoding har en korrelasjon til leseforståelsen, men den er ikke like sterk som vokabular. En grunn til dette kan være at når elever kommer opp på høyere klassetrinn, så vil det forventes at ordavkodingen i størst grad er automatisert og denne sammenhengen ikke er like sterk. Da ser det ut til at andre komponenter, som for eksempel ordforråd, har større betydning for leseforståelsen (Bråten, 2007).

I undersøkelsen til Mezynski (1983) referert i Carlisle & Rice (2002) viste det at dårlige ordavkodere kunne ha et godt vokabular, men ved dårlig avkoding klarte de ikke å hente frem vokabularkunnskapene sine slik at de fikk forståelse av innholdet i teksten. Det som det ble konkludert med var at man kunne ha et godt vokabular, men man må også være en god ordavkoder for å få en god leseforståelse.

Det viser seg at det er en sammenheng mellom de tre variablene; leseforståelse, vokabular og ordavkoding. Der sammenhengen til ordavkoding viser seg å være litt ulik. Mezynski (1983) referert i Carlisle & Rice (2002) viser at ordavkoding er viktig for å klare å bruke vokabularkunnskapene sine og for å få en forståelse ut av en tekst. Mens Petersen (2008) og i denne undersøkelsen ser man at ordavkoding har en sammenheng til leseforståelsen, men at den er størst på lavere klassetrinn. På høyere klassetrinn er det vokabular som har den sterkeste sammenhengen til leseforståelsen.

## 5.5 Resultatene drøftet i lys av validitet

### 5.5.1 Statistisk validitet

God statistisk validitet er en nødvendig betingelse for indre-, ytre- og begrepsvaliditet (Shadish et al., 2002). Når det i denne undersøkelsen skal foretas slutninger som angår sammenhengen mellom vokabular og leseforståelse, kan det diskuteres om sammenhengen skyldes tilfeldigheter. Dette kommer helt an på utvalgsstørrelsen og designet (Lund, 2002a). Statistisk validitet forutsetter at sammenhengen mellom variablene i en undersøkelse er statistisk signifikant eller rimelig sterk (ibid). For å oppnå statistisk validitet i denne undersøkelsen må det være sammenheng mellom leseforståelse, ordavkoding, vokabular og generelle kognitive evner. At en undersøkelse har en rimelig sterk sammenheng varierer i de enkelte undersøkelsene. Dette relateres vanligvis til statistisk signifikans. Det kan sies å være signifikant dersom det er liten sjanse (under 5 %) for at resultatene er tilfeldige (ibid). Når det gjelder den foreliggende undersøkelsen er det vist signifikante resultater. Det viser en sammenheng mellom leseforståelse, ordavkoding, vokabular og generelle kognitive evner.

Mulige trusler mot statistisk validitet kan ifølge Lund (2002a) deles inn i to: lav statistisk styrke og brudd på statistiske forutsetninger. Under vil truslene bli presentert og det vil bli vist hvordan de gjør seg gjeldende i denne undersøkelsen.

#### Lav statistisk styrke

Ved statistisk styrke snakker man om sjansen for å forkaste en gal nullhypotese (type I-feil). Det vil altså si sannsynligheten for at man konkluderer med at det statistisk sett er sammenhenger, uten at slike sammenhenger eksisterer (Shadish et al, 2002). Ut fra variabelen leseforståelse målt med Nasjonale prøver i lesing viser ordkjedetesten  $p = .005$ , deltesten Ordforståelse i WISC III  $p = .014$  og Raven  $p = .000$ . Dette er lave  $p$  verdier og man kan da anta at det er et reelt funn, at det er en sammenheng mellom leseforståelse, avkoding, dybdevokabular og generelle kognitive evner. Samtidig kan man ikke utelukkende være sikker på at man ikke gjør en type I – feil. Signifikansverdien for deltesten Ordforståelse i WISC III på 1,4 % indikerer at det fremdeles kan være sjanse for at funnet ikke er reelt. Det kan da sies at det er 1,4 % sjanse for å gjøre en type I – feil ved å si at det er sammenheng mellom variablene, uten at det i virkeligheten er det.

Resultatene vise at det er lave  $p$  verdi knyttet til forskningsspørsmålet om hvilke innvirkning vokabulardybde har på leseforståelsen blant 8. klassinger, og den statistiske validiteten er på den måten innfridd i forhold til dette. Utvalgsstørrelsen er også med på å øke undersøkelsens statistiske validitet. Utvalgsstørrelsen i denne undersøkelsen kan regnes som tilstrekkelig stor nok ( $N= 112$ ). Selv om det alltid vil være en liten sjanse for type I – feil.

Som nevnt før kan det også være fare for type II feil, at en konkluderer med at det ikke er sammenheng når det er det. I denne studien ble det vist et ikke-signifikant resultat i regresjonsanalysen mellom vokabularbredde og leseforståelse, men likevel burde dette bli konkludert med forsiktighet. Dette på grunn av at forholdet mellom breddevokabular og leseforståelse viser signifikante resultater i tidligere studier, og det er rimelig å anta at på et eller annet tidspunkt i leseutviklingen vil breddevokabular ha en signifikant forklaringsvarians til leseforståelsen.

Dårlig testreliabilitet er en trussel mot statistisk validitet, fordi dette vil redusere den statistiske styrken som igjen kan føre til type II-feil (Lund, 2002a). I denne undersøkelsen er det brukt kjente og godt standardiserte tester som har høy reliabilitet. Imidlertid kan det være mange andre faktorer som kan være med på å påvirke testsituasjonen. Det kan være faktorer som barnets dagsform, motivasjon, konsentrasjon og gjennomføring av mange tester. Dette kan gjøre at målingene gir muligheter for feil, slik at undersøkelsens variabler ikke blir målt på en fullgod måte. Testingen kan da føre til en trussel mot den statistiske validiteten i tillegg til det kan føre til en trussel mot begrepsvaliditet. I gjennomføringen av testingen i denne undersøkelsen ble det lagt vekt på å skape trygge rammer rundt testsituasjonen og testingen ble utført i trygge omgivelser og testene ble gjennomført etter manualen.

### **Brudd på statistiske forutsetninger**

Om en variabel ikke er normalfordelt, betyr det brudd statistiske forutsetninger for parametriske tester. Om man gjennomfører en parametrisk test vil det true den statistiske validiteten (Shadish et al, 2002). I denne undersøkelsen ble det brukt standardiserte tester. Det er med på å gjøre alle variablene; leseforståelse, breddevokabular, dybdevokabular, avkodning og kognitiv fungering, normalfordelte. Skulle det oppstå avvik fra den statistiske forutsetningen om normalfordeling, kan det innebære at sjansen for å trekke gal slutning om at det er sammenheng mellom leseforståelse og vokabular (type 1-feil) øke.

Kan vi konkludere med at det er oppnådd statistisk validitet i denne undersøkelsen? Ut fra diskusjonen over kan en konkludere med at det er signifikante resultater blant annet på bakgrunn av at det har vært en stor utvalgsstørrelse (N). Kort oppsummert viser det at Raven, ordkjedetesten og deltesten Ordforståelse i WISC III er signifikante, undersøkelsen har en stor utvalgsstørrelse og at alle variablene er normalfordelt, dette er med på å gjøre denne undersøkelsen til statistisk valid. Ved å ha en god statistisk styrke øker det sjansen for å unngå type II-feil. Siden BPVS II har et ikke signifikant resultat kan dette være med på å svekke den statistiske validiteten, men gjennom en høy reliabilitet kan man på annen side si at BPVS II er med på å øke den statistiske validiteten, det vil bli drøftet under i neste avsnitt.

### 5.5.2 Reliabilitet

Christophersen (2007) kategoriserer en reliabilitetskoeffisient på .60 som moderat reliabel, og en Alpha på .80 som relativt høy. Testene som ble reliabilitetstestet i denne undersøkelsen hadde en relativt høy reliabilitetskoeffisient. Deltesten i ordforståelse i WISC III hadde en Alpha = .80 og BPVS II hadde en Alpha på = .87. Dette kan karakteriseres som at testene er reliable i følge Gall et al. (2007), som betrakter reliabilitet over .80 som tilstrekkelig akseptabelt i de fleste sammenhenger.

En faktor som kan være med på å påvirke reliabiliteten, er usikker skåring og vurdering av testene (Gall et al., 2007; Kleven, 2002b). Dette gjelder i størst grad deltesten Ordforståelse i WISC III hvor det gis rom for bruk av skjønn. I henhold til manualen for WISC III og deltesten Ordforståelse er det gitt føringer for hva som gir ett poeng og hva som gir to poeng. I manualen foreligger det også føringer for hva som gir null poeng. Dette er med på å gjøre skåringen av testen bedre og det øker testens reliabilitet. Selv med konkrete og spesifiserte skåringskriterier er det vanskelig å skape gode skåringskriterier på en begrepstest, dette kan være med på å svekke reliabiliteten. I denne undersøkelsen har flere personer vært med på gjennomføring og skåring på deltesten Ordforståelse i WISC III, man kan anta at av den grunn vil ulike skjønnsvurderinger påvirke resultatene. Dette kan også føre til svekkelse av reliabiliteten.

BPVS II er en test som gir en liten mulighet for gjetting. Elevene skulle ved denne testen peke på et av fire bilder som de mente tilhørte ordet de fikk av testleder. Dette gir rom for å sette et spørsmål ved reliabiliteten til testen. Ved gjetting kan elevene skåre bedre enn det som er forventet. BPVS II er en kjent og standardisert test, som er brukt til måling av

breddevokabular i flere undersøkelser. Siden testen er standardisert og normert kan det antas at slike målefeil er tatt med i standardiseringen av testen. Derfor kan en med en viss grad si at testen ut fra dette er reliabel.

Tilfeldige målefeil kan være med på å svekke reliabiliteten til en undersøkelse. Testene som er med i denne undersøkelsen er blitt utført av lærerne ved skolene, mastergradsstudenter og forskningsassistenter ved ISP. I forkant av testingen fikk alle opplæring i administrering og skåring av testene. Dette er med på å øke reliabiliteten og minske tilfeldige målefeil i denne undersøkelsen. Det kan da antas at undersøkelsen kan sies å være reliabel.

### **5.5.3 Indre validitet**

Ifølge Shadish et al. (2002) sier indre validitet noe om kausal sammenheng/relasjon mellom to sett av operasjonaliseringer. I et ikke – eksperimentelt design er det spesielt viktig å belyse den indre validiteten, og det vil være nødvendig å vurdere alternative tolkninger av resultatene fra undersøkelsen (Lund, 2002b). Videre i dette kapittelet vil undersøkelsens indre validitet og da andre mulige årsaker til de tidligere presenterte resultatene drøftes.

#### **Alternative årsaksforklaringer:**

Den viktigste trusselen mot indre validitet kalles retningsproblemet og med dette menes vanskeligheten med å avgjøre hva som er årsak og hva som er virkning (Lund, 2002a).

Når det gjelder spørsmål om det finnes noen kausale relasjoner mellom variablene og tolkning av regresjonskoeffisienter ( $R^2$ ), er det umulig å si noe om hva som er årsak eller virkning i relasjonene, dette kan bare studeres gjennom en undersøkelse med et eksperimentelt design (Shadish et al. 2002; Pedhazur, 1997; Tabachnick og Fidell, 2007).

Når det gjelder spørsmålet om hva som er årsak og virkning mellom vokabular og leseforståelse, kan man anta at det er en gjensidig relasjon mellom variablene; begge spiller en kausal rolle for ferdighetene på den andre variabelen. Videre kan man spørre seg hva som er retningen i slike relasjoner. Det bør ligge til grunn en bakenforliggende teori for hvilke variabler man undersøker som henholdsvis avhengig – og uavhengigvariabel; altså blir det viktig å ta i betraktning den teoretiske bakgrunnen for hvilke analyser som gjøres. Utvikling av leseforståelse henger tett sammen med omfanget av barnas vokabular, og selv om de har knekt ordavkodningen vil den videre leseutviklingen lide hvis barnet ikke har et ordforråd som

kan støtte forståelsen. Selv om et dårlig vokabular kan påvirke ordavkodingen, vil barn som møter skolen med et dårlig utviklet ordforråd først og fremst få vansker med å utvikle en god leseforståelse. Lesing er en god kilde for vokabularutviklingen, men hvis vokabularet i utgangspunkt er dårlig kan skoletekster fort bli for vanskelige (Frost, 2003; Frost, 1999). Vokabular er et begrep som kan sees på som et sett av ord som er basis i byggingen og forståelsen av en setning (Anglin, 1993). Leseforståelse er en prosess som Bråten (2007) definerer som ”å utvinne og skape mening ved å gjennomføre og samhandle med skrevet tekst” (s.45). Noe som understøtter at relasjonen mellom vokabular og leseforståelse går begge veier er de hypotesene som Anderson og Freebody (1981) beskriver. Kort oppsummert dreier disse seg om at det er ikke vokabularkunnskap som er med på å gi en god leseforståelse, men det er kunnskaper på begge områder som er den underliggende evnen til å bli en god leser og få et godt vokabular (ibid.). I denne undersøkelsen ble leseforståelse valgt som den avhengige variabel og vokabular, delt inn i bredde og dybde, ble valgt som uavhengige variabler. Grunnen til dette var å finne ut om bredde – og dybdevokabular i ulik grad predikerer en god leseforståelse.

Ved designet som er brukt i denne undersøkelsen er den indre validiteten en svakhet, da det ikke er mulig å trekke sikre slutninger om kausale sammenhenger. Men ved bruk av hierarkisk multippel regresjonsanalyse, styrker denne analysemetoden i seg selv den indre validiteten. Grunnen til dette er at det kontrolleres for de variablene som settes inn i likningen, før den ønskede prediktorvariabelen, som en ønsker å studere i forhold til kriterievariabelen, føres inn (Kleven, 2002b). I analysen når det ble undersøkt om breddevokabular kunne predikere leseforståelse, ble effekten fra dybdevokabular holdt konstant slik at denne variabelen ikke hadde innflytelse på bidraget breddevokabular forklarte. Og når det ble undersøkt om dybdevokabular kunne predikere leseforståelse, ble effekten fra breddevokabular holdt konstant slik at denne variabelen ikke hadde innflytelse på bidraget dybdevokabular forklarte.

Testene som er brukt skal måle begrepene som ønskes og undersøkes på en tilfredsstillende måte. Den indre validiteten blir svekket dersom instrumentene som brukes ikke er standardiserte, og dermed kan det være med på å skape kunstige resultater. Denne trusselen kalles instrumentering og eksempler på slike kunstige resultater kan være ikke – konstant enhet på en antatt intervallskala, ”golv” – eller ”tak” – effekt på måleskalaen og systematiske endringer i skåringskriterier (Lund, 2002a). I denne undersøkelsen er testene som ble benyttet

standardiserte, dermed er ikke instrumentering en trussel mot den indre validiteten. I utgangspunktet er WISC ment å måle intelligens (Weschler, 1991). Men i denne undersøkelsen benyttes deltesten Ordforståelse i WISC for å teste dybdevokabular. Her må testpersonen muntlig framstille synonymer og/eller gode forklaringer på en rekke enkeltord. British Picture Vocabulary Scale (BPVS) (Dunn et al., 1997). Testen ble standardisert på norsk i 2006 ved Institutt for Spesialpedagogikk ved Universitetet i Oslo. I denne studien brukes BPVS for å måle breddevokabular. Ordkjedetesten er standardisert (Høien og Tønnesen, 2007), og denne undersøkelsen bruker testen som mål på avkodingsferdigheter. Raven betraktes som en svært god test for nonverbal intelligens, altså generell kognitiv fungering (Raven, 1986). Svakheter og styrker ved de ulike testene vil beskrives nærmere under drøfting av begrepsvaliditet.

#### **5.5.4 Begrepsvaliditet**

Begrepsvaliditet kan ofte knyttes til svakheter rundt hvordan begrepet man ønsker å måle er operasjonalisert på. Dette blir sett på som den største trusselen mot begrepsvaliditet (Gall et al., 2007; Kleven 2002b). Shadish et al. (2002) utvider begrepsvaliditet, i tillegg til undersøkelsens variabler, til å innebære alt som er spesielt med utvalget. Begrepsvaliditet kan deles inn i to sider. Den ene siden er å forstå begrepet, mens den andre er å fastsette betydningen av det (ibid). Det kan oppstå problemer i forhold til begrepsvaliditet i forbindelse med å karakterisere og måle utvalget, testene, rammene rundt og resultatene i forhold til undersøkelsen (Lund, 2002a). Tilfeldige feil i avhengig og eller uavhengige variabler kan resultere i feil i estimeringen av  $R^2$ , og dette kan også inntreffe dersom noen av variablene inneholder systematiske feil (Pedhazur, 1997). Om en uavhengig variabel i regresjonsanalysen underestimeres, vil dette føre til at også andre uavhengige variabler kan bli overestimert som prediktorer for kriterievariabelen.

De relevante begrepene som operasjonaliseres i denne undersøkelsen er leseforståelse, breddevokabular, dybdevokabular, ordavkodning og generelle kognitive evner. I denne undersøkelsen er det operasjonalisering av mer teoretiske begreper som kan være problematisk i forhold til begrepsvaliditet. Raven, som måler generelle kognitive evner, har en sterk forklaringsvarians til leseforståelse målt med Nasjonale prøver i lesing så det kan her tenkes at den måler mer elevens kognitive evner enn leseforståelse. Det er med dette mulig å si at generelle kognitive evner har en sterk sammenheng med leseforståelsen. For å måle

vokabularbredde ble det brukt testen BPVS II, som er en standardisert test for å måle vokabularbredde. Mens for å måle dybdevokabular ble det brukt deltesten Ordforståelse i WISC III. Er disse gode nok måleinstrumenter for å måle det som undersøkelsen er ute etter, selv om de også er kjente tester for å måle bredde- og dybdevokabular? I og med at de også er standardiserte tester og brukt i andre undersøkelser, er det i denne undersøkelsen gått ut i fra at de måler bredde- og dybdevokabular på er god nok måte og vi har fått de resultatene som var ventet. Ordkjedetesten er standardisert og er en screeningprøve som "fanger opp" de elevene som har vansker med å avkode ord (Høien & Tønnesen, 2007). I denne undersøkelsen anses testen som et godt nok mål for ordavkodning. Kleven (2002b) nevner at det kreves av måleinstrumentene at de dekker et representativt utvalg av det vi ønsker å måle dersom en skal oppnå god innholdsvaliditet.

Ovenfor er det vist til flere faktorer ved testoppgaven som kan være med på å svekke begrepsvaliditeten. Det kan være mange ulike årsaker til at det kan inntreffe systematiske feil. Disse kan inntreffe dersom testadministrator systematisk har administrert testen på en feilaktig måte. Dette kan være knyttet til ukorrekte prosedyrer ved gjennomføring og skåringen, noe som kan resultere i at resultatene i indikatorene ikke dekker begrepet det har vært ønskelig å måle (Pedhazur, 1997). Ved skåring av deltesten Ordforståelse i WISC III er det lagt inn føringer for skåring, men det må og legges til grunn en viss grad av skjønn og vurderinger av testadministratoren. Ved innhenting av data til denne undersøkelsen har det vært mange som har gjennomført og skåret testen, i og med at det er en del av et større forskningsprosjekt, og da vil det kunne gi mulighet for ulike vurderinger av et svar.

Som nevnt tidligere har Shadish et al. (2002) valgt å utvide begrepsvaliditet til å omfatte utvalget også. Undersøkelsens utvalg er 8. klassinger fra to skoler i Oslo. Ut i fra dette er det ikke mulighet for å si noe om utvalget representere populasjonen. Familieforhold, foreldrenes utdanning, oppdragelse og verdier er utenfor kontroll, dette på grunn av at det eneste kriteriet som er blitt gitt til denne undersøkelsen, er at elevene skulle ha norsk som morsmål. Dette viser at faktorer utenfor selve testoppgavene kan også påvirke begrepsvaliditeten.

Som denne diskusjonen viser er det flere faktorer som kan være med å true begrepsvaliditeten. Truslene i denne undersøkelsen kan være litt uklare. Det er blitt brukt standardiserte tester som man kan si er gode nok til å måle det de skal, men gjennomføringen og skåringen av testene kan påvirke resultatene.



### 5.5.5 Ytre validitet

Spørsmålet om ytre validitet knytter seg til om resultatene kan generaliseres til populasjonen av typiske utviklede barn i 8. klasse. I forbindelse med dette blir det viktig å belyse hvordan utvalget i undersøkelsen har blitt valgt ut for å sikre den ytre validiteten. Siden undersøkelsen er tilknyttet forskningsprosjektet KiSP, er skolene som er med i prosjektet forhåndsbestemt av ledelsen for prosjektet. En måte å gardere seg mot trusselen for ytre validitet, er trekking av et tilfeldig utvalg fra populasjonen (Lund, 2002a). I denne undersøkelsen er det få utvalgskriterier, den består av 8. klassinger (12 – 13 åringer) som er majoritets elever, dvs. at minoritets elever er valgt bort. Når det gjelder generaliseringer utover de som deltok i undersøkelsen er det flere forhold som kan virke inn på den ytre validiteten. Likheten mellom utvalget og populasjonen er med på å påvirke hvorvidt det man har funnet kan gjelde for flere enn elevene som er testet. I denne sammenhengen vil ikke utvalget representere et sterkt selektert utvalg, siden det er så få utvalgskriterier som legges til grunn. Dette vil ikke være en stor trussel mot den ytre validiteten at gruppen er så egenartet at resultatene ikke kan generaliseres over populasjonen.

At en undersøkelse sjelden gir høy ytre validitet er et annet sentralt forhold ved generaliseringsproblematikk (Lund, 2002b). I denne forbindelse blir det viktig å sammenligne resultatene fra denne undersøkelsen, med andre studier gjort innenfor det samme feltet. Den ytre validiteten styrkes ved at resultatene i denne undersøkelsen i sin helhet viser seg å stemme overens med den tidligere forskningen. En trussel mot den ytre validiteten i denne sammenheng er bortvelgelsen av minoritetsspråklige elever. Man kan mest sannsynlig forvente at gjennomsnittet på denne gruppen, som kun består av majoritetsspråklige elever, blir høyere enn landsgjennomsnittet som består av både majoritets – og minoritetsspråklige elever. Grunnene til at minoritetsspråklige elever er valgt bort, er fordi det kan være stor spredning i disse elevenes norskkunnskaper. Man kan da anta at noen av elevene med minst norskkunnskaper er med på å påvirke gjennomsnittet på Nasjonale prøver i lesing på landsbasis.

## 5.6 Oppsummering

Problemstillingen i denne undersøkelsen er hvilke sammenhenger det er mellom vokabular og leseforståelse blant 8. klassinger. Videre er problemstillingen delt inn i to forskningsspørsmål der vokabularbreddens og vokabulardybdens innvirkning på leseforståelsen ville undersøkes. Ettersom vokabular og leseforståelse henger tett sammen og påvirker hverandre gjensidig, kan det forventes korrelasjoner mellom variabler.

I denne studien fant vi slik som i mange andre undersøkelser på samme område (Muter, Hulme, Snowling og Stevenson, 2004; Nation og Snowling, 2004; Roth, Speece og Cooper, 2002; Sènèchal, Ouellette og Rodney, 2006; Share og Leikin, 2004; Snow, Tabors, Nicholson og Kurland, 1995) at vokabulardybde i større grad enn vokabularbredde predikerer leseforståelsen i positiv retning. Ouellette (2006) er den som har gjennomført en nyere undersøkelse angående sammenhengen mellom vokabular og leseforståelse. Han delte også inn i bredde- og dybdevokabular. De funnene som han fant i sin undersøkelse stemmer overens med funnene i denne undersøkelsen. Forskjellen i disse to studiene er at Ouellette (2006) undersøkte 4. klassinger og i denne undersøkelsen består utvalget av 8. klassinger, men likevel ser vi at det er dybdevokabular som har sterkest forklaringsvarians på leseforståelsen. Regresjonsanalysen i foreliggende undersøkelse viser at deltesten Ordforståelse i WISC III, som måler dybdevokabular, har den største unike forklaringsvariansen på 3,2 %, sammenlignet med BPVS II, som måler breddevokabular, som har en unik forklaringsvarians på 2,1 % etter at det er kontrollert for Ordkjedetesten og Raven. I korrelasjonsanalysen i denne undersøkelsen korrelerer testen for breddevokabular (BPVS II) og testen for dybdevokabular (deltesten Ordforståelse i WISC III) omtrent like mye til leseforståelsen målt med testen Nasjonale prøver i lesing. Dette viser at vokabularbredde også har en sammenheng med leseforståelsen, men at denne sammenhengen synes ikke å være like sterk når man kommer opp i høyere klassetrinn. En årsak til at den variansen BPVS II og deltesten Ordforståelse i WISC III har til leseforståelse målt med Nasjonale prøver i lesing ikke er så stor, kan være på grunn av kolinaritet. Da er det så stor korrelasjon mellom de uavhengige variablene at de ikke gir noe unikt bidrag til den avhengige variabelen (Pallant, 2007).

Et nokså uforventet funn var den sterke variansen testen Raven hadde til leseforståelsen da målt med testen Nasjonale prøver i lesing. Raven hadde en unik forklaringsvarians på 45,8 %

til leseforståelsen. En årsak til dette kan, som nevnt tidligere, være at Nasjonale prøver i lesing krever såpass mye av elevens kognitive evner og at Raven sin unike forklaringsvariabel er av den grunn reell.

Ordavkoding er en selvfølgelighet i leseopplæringen og danner grunnlaget for lesing, og er av den grunn tatt med i denne undersøkelsen. Det viser seg her at ordavkoding testet med Ordkjedetesten ikke hadde så sterk korrelasjon til leseforståelsen sett i forhold til bredde – og dybdevokabular. Dette kan mest sannsynlig forklares med at på 8. klassetrinn er det forventet at ordavkodingen er automatisert og derfor ikke er en like sterk faktor som det vokabular er (Bråten, 2007).

I denne undersøkelsen ble det brukt Nasjonale prøver i lesing som test for leseforståelsen. Det som viser seg her er at testen for leseforståelse er ganske utslagsgivende for resultatene en får i forhold til sammenhengen mellom leseforståelse og vokabular. Det kunne derfor ha vært interessant å se på ulike lesetester; hva de egentlig tester og hvordan de egner seg i undersøkelser der man vil måle nivået på en persons leseforståelse.

Denne studien er en del av et større longitudinelt forskningsprosjekt, noe som gir muligheter til videre innsikt i forholdet mellom vokabular og leseforståelse ved at barn blir fulgt over tid. Ved at mange av disse spørsmålene kan sees på ved senere tidspunkt kan det gi kunnskap om kausale sammenhenger. Mange spørsmål gjenstår ubesvart eller uklare, og det er behov for flere longitudinelle undersøkelser på dette feltet. Det kunne ha vært interessant å følge det samme utvalget elever over tid, for å se hvordan vokabularutviklingen ble etter intervensjonsprogrammer i vokabularlæring, for å se om en slik intervensjon har effekt på leseforståelsen. Selv om vi i denne undersøkelsen ikke avdekker noen nye funn, kan det likevel være interessant å undersøke dette temaet grundigere når det gjelder eldre barn. Undersøkelser som er tidligere har blitt gjort og som beskrives i empiridelen er stort sett undersøkelser gjort på yngre elever.

Hensikten med denne undersøkelsen har vært å få ytterligere forståelse for temaet slik at man kan dra nytte av det gjennom et barns leseutviklingsprosess. Lesing er en god kilde for vokabularutviklingen, men hvis vokabularet i utgangspunkt er dårlig kan skoletekster fort bli for vanskelige. Det er viktig at læreren forsikrer seg om at elevene har det ordforrådet som trengs for å forstå en fagtekst, dette er med på å gi det nødvendige grunnlaget for at videre læring skal kunne finne sted (Frost 2003; Frost, 1999). Barn med dårlig vokabular kan være i

fare for å få leseforståelsesvansker, noe som igjen kan føre til innvirkning på deres skolegang. Et dårlig vokabular kan påvirke ulike forhold ved leseutvikling, som avkoding, leseflyt og leseforståelse. En grunnleggende vokabularinnlæring, der både bredden og dybden vektlegges, vil ha en positiv innvirkning på leseforståelsen.

## Kildeliste

- Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. I J. Guthrie (Red), *Comprehension and teaching: Research reviews*. (s. 77 – 117). Newark, DE: International Reading Association.
- Anglin, J. M. (1993). Vocabulary development: A morphological analysis. I: *Monographs of the society for research in child development*, 238, vol. 58 (10).
- Baker, S. K, Simmons, D. C & Kameenui, E. J. (1998). Vocabulary Acquisition: Research Bases. I D. C Simmons, & E. J Kameenui (Red.), *What Reading Research Tells Us About Children with Diverse Learning Needs. Bases and Basics* (side 183 – 218). New Jersey: Lawrence Erlbaum Associates, Publishers.
- Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia and real-time processing. I: *Language and cognitive processes*, 12 (5/6)/1997, 507-584.
- Befring, E. (2007). *Forskningsmetode med etikk og statistikk*. Oslo: Det norske samlaget.
- Biemiller, A. (1999). *Language and reading success*. Cambridge, MA: Brookline Books.
- Bishop, D. M. V. (1997). *Uncommon understanding: development and disorders of language comprehension in children*. Hove: Psychology press.
- Brown, L., Sherbenou, R.J., & Johnsen, S.K. (1997). TONI – 3: *Test of Nonverbal Intelligence* (3<sup>rd</sup> ed.). Austin, TX: Pro – Ed.
- Brudholm, M. (2002). *Læseforståelse – hvorfor og hvordan?* København: Alinea.
- Bråten, I. (1997). Leseforståelse. *Nordisk Pedagogik*, 17, 95 – 110.
- Bråten, I. (2007). Leseforståelse – komponenter, vansker og tiltak. I I. Bråten (Red.), *Leseforståelse. Lesing i kunnskapssamfunnet - teori og praksis* (side 45 – 81). Oslo: Cappelens Forlag.
- Bø, I. & Helle, L. (2008). *Pedagogisk ordbok*. Oslo: Universitetsforlaget.

- Carlisle, J.F. & Rice, M.S. (2002). *Improving Reading Comprehension. Research – Based Principles and Practices*. Baltimore: York Press, Inc.
- Carlsten, C.T.(1993). *Leseprøve. 7. klasse bokmål*. Oslo: Universitetsforlaget.
- Carpenter, P.A., Just, M.A. & Shell, P. (1990). What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. *Psychological Review*. Vol. 97, No. 3, 404 – 431.
- Christophersen, K. (2006). *Databehandling og statistisk analyse med SPSS*. Oslo: Unipub.
- Coleman, J. (1998). Cognitive reality and the phonological lexicon: A review. *Journal of Neurolinguistics*, 11, 295 – 320.
- Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation. Design & Analysis Issues for Field Settings*, Houghton Mifflin Company, Boston.
- Craver, R.P. (1977). *Technical manual for the National Reading Standards*. Kansas City, MO: Revrac Publications.
- Craver, R.P. (1990). Intelligence and Reading Ability in Grades 2 – 12. *Intelligence*, 14, 449 – 455.
- Daneman, M. & Green, I. (1986): Individual differences in comprehending and producing words in context. I: *Journal of memory and cognition*, 25/1986, 1-18.
- Dickinson, D. K., Tabors, P. O. (2001). *Beginning Literacy with Language: Young Children Learning at Home and School*. Baltimore: Brooks Publishing Company.
- Drum, P. A. & Konopak, B. C. (1987): Learning word meanings from written context. I M. G. McKeown (Red.), *The nature of vocabulary acquisition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Dunn, L.M & Dunn L.M. (1981). *Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.
- Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale*. Second Edition. Testbook. London.

- Finbak, L. (1998). Lese – og skrivevansker – dysleksi? En undersøkelse av utvalgs-kriterier og sammenhenger mellom avkoding og staving. Doktoravhandling, Institutt for spesialpedagogikk, Universitetet i Oslo.
- Frost, J. (1999). *Lesepraksis – på teoretisk grunnlag*. Oslo: Cappelen Akademiske Forlag.
- Frost, J. (2003). *Prinsipper for god leseopplæring. Innføring i den første lese- og skriveopplæringen*. Oslo: Cappelen Akademiske Forlag.
- Gall, M. P., Gall, J. P. & Borg, W. R. (2007). *Educational Research. An Introduction*. New York: Longman Publishers.
- Gathercole, S. & Baddeley, A. (1993): *Working memory and language*. Hove: Erlbaum.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore: Paul H. Brookes.
- Høyen, T. & Tønnesen, G. (2007). Ordkjedetesten. Logometrica AS.
- Just, M.A., & Carpenter, P.A. (1987). *The Psychology of Reading and Language Comprehension*. Boston: Allyn and Bacon.
- Kleven, T.A. (2002a). Ikke – eksperimentelle design. I T. Lund (Red.), *Innføring i forskningsmetodologi* (side 265-286). Oslo: Unipub.
- Kleven, T.A. (2002b). Begrepsoperasjonalisering. I T. Lund (Red.), *Innføring i forskningsmetodologi* (side 141-183). Oslo: Unipub.
- Lahey, M. (1988). *Language disorders and language development*. Needham, MA: Macmillan.
- Levelt, W.J.M., Roelofs, A., & Meyer A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1- 75.
- Lewis-Beck, M.S. (1980). *Applied Regression. An Introduction*. London: Sage Publications.
- Lund, T. (2002a). Metodologiske prinsipper og referanserammer. I T. Lund (Red.), *Innføring i forskningsmetodologi* (side 79-123). Oslo: Unipub.

- Lund, T. (2002b). Generaliseringsproblematikk. I T. Lund (Red.), *Innføring i forskningsmetodologi* (side 125-140). Oslo: Unipub.
- Molander, B. og Skauge, I.L. (2009). *Lese Lære Lykkes*. Oslo: Gyldendal.
- Muter, V., Hulme, C., Snowling, M.J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary and grammatical skills as foundation of early reading development: Evidence from a longitudinal study. *Development Psychology*, 40, 665 – 681.
- Nagy, W. E. (1988). *Teaching Vocabulary to Improve Reading Comprehension*. ERIC. Clearinghouse on Reading and Communication Skills.
- Nagy, E. W. & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. I M. G. McKeown, & M. E. Curits (Red.), *The nature of vocabulary acquisition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Nation, K., & Snowling, M.J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27, 342 – 356.
- NESH, (2006). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Forskningsetiske komiteer.
- Ouellette, G.P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. I: *Journal of educational psychology*, 98/3/2006, 554-566.
- Pallant, J. (2007). *SPSS. Survival Manual*. New York: Open University Press.
- Pedhazur, E.J. (1997). *Multiple regression in behavioral research. Explanation and prediction* (3<sup>rd</sup> ed). Orlando, FL: Harcourt Brace.
- Petersen, D. K. (2008). *Hvad har betydning for eleveres læseforståelse?* Fjerritslev: Landsforeningen af Læsepedagoger.
- Pikulski, J. J. (2006). Fluency: A Developmental and Language Perspective. I S. J Samuels, & A. E. Farstrup (Red.) *What Research Has to Say About Fluency Instruction*. Newark: International Reading Association.



- Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 26(2), 282 – 307.
- Raven Standard Progressive Matrices test ABC* (Raven). (1986). London: Oxford Psychologists Press.
- Richard, G. J. & Hanner, M. A. (1995). *Language Processing Test-Revised*. East Moline, IL: LinguiSystems, Inc.
- Rommetveit, R. (1972): *Språk, tanke og kommunikasjon: Ei innføring i språkpsykologi og psykolingvistikk*. Oslo: Universitetsforlaget.
- Roth, F.P., Speece, D.L., & Cooper, D.H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research*, 95, 259 – 272.
- Samuelstuen, M.S. og Bråten, I. (2005). Decoding, knowledge and strategies in comprehension of expository text. *Scandinavian Journal of Psychology*, 46, 107 – 117.
- Sènèchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the early vocabulary to future reading. I S. B. Neuman, & D. Dickinson (Red.), *Handbook of early literacy research: Vol. 2* (pp. 173 – 182). New York: Guilford Press.
- Shadish, W. R, Cook, T. D. & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.
- Share, D.L. & Leikin, M. (2004). Language impairment at school entry and later reading disability. Connections at lexical versus supralexical levels of reading. *Scientific Studies of Reading*, 8, 87 – 110.
- Snow, C.E., Tabors, P.O., Nicholson, P.E. & Kurland, B.F. (1995). SHELL: Oral language and early literacy in kindergarten and first grade children. *Journal of Research in Childhood Education*, 10, 37 – 47.
- Snow, C.E., Burns, M.S. & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

- Stahl, S. A. & Nagy, W.E. (2006). *Teaching Word Meanings*. London: Lawrence Erlbaum Associates, Publishers.
- Stanovich, K.E., Cunningham, A.E., & Feeman, D.J. (1984). Intelligence, cognitive skills and early reading progress. *Reading Research Quarterly*, 19, 278 – 303.
- Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360 – 407.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. I M. G. McKeown (Red.), *The nature of vocabulary acquisition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Sternberg, R.J., & Powell, J. (1983). Comprehending verbal comprehension. *American Psychologist*, 38, 878 – 893.
- Sticht, T.G. (1979). Applications of the Audread modell to reading evaluation and instruction. I L. B. Resnick, & P. Weaver (Red.), *Theory and practice of early reading*. (Vol. 1, s. 209 – 226). Hillsdale, NJ: Erlbaum Associates.
- Strømsø, H. (2007). Høytlesing, hurtiglesing og leseforståelse – en historie om lesing og forskning om leseforståelse. I I. Bråten (Red.), *Leseforståelse. Lesing i kunnskapssamfunnet - teori og praksis* (side 20 - 44). Oslo: Cappelen Forlag.
- Sørensen, P.M. (2006). Statistikk. I K. Fuglseth, & K. Skogen (Red.), *Masteroppgaven i pedagogikk og spesialpedagogikk* (side 184 – 207). Oslo: Cappelen Akademisk Forlag.
- Tabachnick, B.G. & Fidell, L.S. (2007). *Using multivariate statistics*. Boston, MA: Pearson Education, Inc.
- Tannenbaum, K.R., Torgesen, J.K. & Wagner, R.K. (2005). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10(4), 381 – 398.
- van Dijk, T.A. & Kintsch, W. (1978). Cognitive psychology and discourse: Recalling and summarizing stories. I W.U. Dressler (Red.), *Current trends in textlinguistics* (pp. 61 – 81). New York: De Gruyter.

- Vellutino, F.R. (2003). Individual differences as sources of reading comprehension in elementary school children. I A.P. Sweet & C.E. Snow (Red.), *Rethinking reading comprehension* (pp. 51 – 81). New York: The Guilford Press.
- Vermeer, A. (2001): Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. I: *Applied psycholinguistics*, 22, 217-234.
- Wagner R.K., Muse, A.E. & Tannenbaum, K.R. (2007). Promising Avenues for Better Understanding Implications of Vocabulary Development for Reading Comprehension. I R. K. Wagner, A. F. Muse, & K. R. Tannenbaum (Red.), *Vocabulary Acquisition. Implications for Reading Comprehension* (side 276 – 291). New York – London: The Guildford Press.
- Wechsler, D. (1991). *The Wechsler Intelligence Scale for Children* (3<sup>rd</sup> ed.). New York: Psychological Corp.
- Woodcock, R.W. (1998). WRMT – R: *Woodcock Reading Mastery Tests – Revised*. Circle Pines, MN: American Guidance Services.

**Nettsteder / elektroniske dokumenter:**

<http://www.udir.no/Tema/Lareplaner/>

<http://www.utdanningsdirektoratet.no/Tema/Nasjonale-prover/>