

# Identifisering av barn med spesifikke språkvansker

*Standardiserte testers validitet*

Wenche J. Elsebutangen og Elin S. Holst-Larsen



Masteroppgave i spesialpedagogikk ved  
Det utdanningsvitenskapelige fakultet,  
Institutt for spesialpedagogikk

UNIVERSITETET I OSLO

Vår 2010



# **Identifisering av barn med spesifikke språkvansker**

© Wenche J. Elsebutangen & Elin S. Holst-Larsen

2010

Identifisering av barn med spesifikke språkvansker

Wenche J. Elsebutangen & Elin S. Holst-Larsen

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

# Sammendrag

**Bakgrunn, formål, problemstilling:** Denne oppgaven er tilknyttet forskningsprosjektet Kunnskapsgenerering i det spesialpedagogiske praksisfeltet (KiSP) ved Institutt for Spesialpedagogikk og handler om identifisering av barn med spesifikke språkvansker. De fleste barn gjennomgår en normal språkutvikling. Det finnes imidlertid noen barn som ikke tilegner seg et funksjonelt språk på en adekvat måte, uten at faktorer som vanligvis ledsager vansker med språklæring, er fremtredende. For disse barna er språket den primære vansken. Slike vansker karakteriseres som spesifikke språkvansker (Bele, 2008; Leonard, 2000). Tradisjonelt er standardiserte tester brukt for å gi et mål på hvor langt et barn er kommet i læring og utvikling på testtidspunktet. Barns språkferdigheter kan til en viss grad måles med tester av denne typen, men det er en svakhet ved slike tester at de ofte viser mer globale mål på ekspressive og reseptive ferdigheter, og ikke i så stor utstrekning kan avdekke de spesifikke vanskene hvert enkelt barn har. Ifølge Haywood (1997) er ikke problemet hva de standardiserte testene måler, men hva de ikke måler. I en undersøkelse av nyere dato stiller Bishop & McDonald (2009) spørsmål ved om tradisjonelle språktester mangler sensitivitet overfor klinisk viktige kjennetegn på språkvansker. Vi har valgt å se nærmere på to slike tradisjonelle språktester, TROG-2 og BPVS II (heretter kalt TROG og BPVS). Oppgavens problemstilling er: *I hvilken grad identifiserer TROG og BPVS barn med spesifikke språkvansker på en måte som er valid i forhold til læreres beskrivelse av vanskene?*

**Metode/design:** Problemstillingen belyses med en empirisk-kvantitativ metodisk tilnærming. Undersøkelsen har et ikke-eksperimentelt design, og er deskriptiv i sin form. Data om barns språkferdigheter er blitt samlet inn ved bruk av tester (TROG, BPVS og Raven) og lærervurderinger (spørreskjemaet 20 spørsmål om språkferdigheter). TROG måler grammatisk forståelse og BPVS måler reseptivt vokabular. Raven gir et mål på generell kognitiv fungering. Hver av språktestene i kombinasjon med Raven brukes til å angi om barn kan ha SSV, og spørreskjemaet 20 spørsmål om språkferdigheter er det kriteriet testene måles mot.

Utvalget i undersøkelsen består av 80 barn fra 2., 5., og 6. trinn ved to grunnskoler i Oslo. 47 av barna er jenter, og 33 er gutter. Barna er valgt ut basert på følgende seleksjonskriterier: barna skal være uten hørselshemming, ha nonverbal intelligens innen normalområdet, ha norsk som morsmål, og være testet eller vurdert med alle instrumentene som er nevnt over.

Måleinstrumentene som er benyttet, er valgt fra testbatteriet for hovedprosjektet. I tillegg til de overnevnte instrumentene er Språk 6-16 benyttet som supplerende informasjon i forbindelse med drøftingen av validiteten til TROG og BPVS.

**Resultater:** Resultatene i denne undersøkelsen viser at testene og lærerne (uttrykt ved 20 spørsmål) stort sett er enige om hvilke barn som ikke har spesifikke språkvansker, men at det er et misforhold mellom tester og lærervurderinger når det gjelder å identifisere barn som kan ha spesifikke språkvansker. I noen tilfeller angir lærerne gjennom spørreskjemaet 20 spørsmål om språkferdigheter at barn har en vanske, mens testene ikke fanger det opp. I andre tilfeller angir testene at barn har en vanske, mens lærerne ikke gir indikasjoner på dette. Det ble funnet en korrelasjon mellom TROG og BPVS som hadde lav styrke, men var signifikant. Korrelasjonen mellom TROG og Raven var moderat, mens korrelasjonen mellom BPVS og Raven var nesten ubetydelig. TROG og BPVS hadde svært god spesifisitet, det vil si evne til korrekt å identifisere alle individer uten vansker. Testene oppfylte ikke kravet om sensitivitet, det vil si evne til å identifisere alle individer med en vanske (Rydz, Shevell, Majnemer, & Oskoui, 2005). Det er foretatt en t-test for uavhengige utvalg for å sammenligne gjennomsnittsskåren på TROG og BPVS hos barn som kan ha SSV og barn som ikke har SSV ifølge 20 spørsmål. De seks barna som lærerne angir som barn med mulig SSV, har svakere gjennomsnittsskåre på både BPVS og TROG enn barna som lærerne angir som normalspråklige.

**Konklusjon:** SSV er en sammensatt vanske som er vanskelig å oppdage. Denne undersøkelsen viser at det kreves flere tilnærminger til identifisering av barn med denne vansken, og at reseptive språktester ikke bør brukes alene, men bør suppleres med en eller flere ekspressive språktester. Selv bruk av ekspressive og reseptive tester har begrensninger, fordi de ikke fanger opp hvordan barnet fungerer i sin naturlige kontekst. Denne undersøkelsen viser at det fortsatt er behov for forskning på spesifikke språkvansker, også med tanke på å kunne identifisere barn med SSV på en mer valid måte. Det er ikke noe entydig forhold mellom et barns skåre på en test og læreres oppfatning av barnets språkferdigheter i hverdagen. Tester bør derfor suppleres med en økologisk tilnærming til identifisering og kartlegging. For å legge til rette for at forskere og klinikere skal kunne drive en evidensbasert praksis, bør testens utgivere kunne dokumentere testens egnethet til identifisering av barn med spesifikke språkvansker.







# Forord

To lærerike og spennende år som masterstudenter i logopedi ved Universitet i Oslo er snart over. Studiet avsluttes med denne masteroppgaven. En stor takk går til vår veileder, spesialpsykolog Ernst Ottem, som har gitt oss inspirasjon, god støtte og hjelp underveis i prosessen. En takk går også til Peer Møller Sørensen som har vært behjelpelig med å gjøre statistikken mer forståelig. Det har vært en lærerik prosess, med både oppturer og nedturer. Det har vært en stor fordel å arbeide i tospann. Dette har gitt oss mange gode diskusjoner og erfaringer. En stor takk går også til våre ektemenn som har gjort dette mulig å gjennomføre.

Oslo, mai 2010

Elin Synøve Holst-Larsen og Wenche Jensen Elsebutangen

# Innholdsfortegnelse

1	INNLEDNING .....	1
1.1	Bakgrunn, tema og formål .....	1
1.2	Oppgavens oppbygging .....	4
1.2.1	Presiseringer om bruk av terminologi .....	5
1.2.2	Presiseringer om vedlegg .....	6
2	TEORETISK BAKGRUNN .....	7
2.1	Hva er spesifikke språkvansker? .....	7
2.1.1	Beskrivelse av fenomenet.....	7
2.1.2	Forekomst.....	7
2.1.3	Utvikling i synet på spesifikke språkvansker .....	8
2.1.4	Lingvistiske kjennetegn hos barn med spesifikke språkvansker.....	10
2.1.5	Undergrupper .....	12
2.1.6	Arv.....	13
2.1.7	Kognitive faktorer .....	13
2.1.8	Konsekvenser av språk- og kommunikasjonsvansker.....	15
2.2	Identifisering av barn med SSV.....	16
2.2.1	En psykometrisk tilnærming til identifisering.....	16
2.2.2	En økologisk tilnærming til identifisering .....	19
2.3	Testers validitet og reliabilitet .....	22
2.3.1	Testvaliditet.....	22
2.4	Forskning på testers sensitivitet og spesifisitet.....	26
3	METODISK TILNÆRMING .....	31
3.1	Vår undersøkelse .....	31
3.1.1	Utvalg .....	31
3.1.2	Design og metode.....	33
3.1.3	Gjennomføring .....	34
3.1.4	Måleinstrumenter .....	35
3.1.5	Supplerende måleinstrument .....	38
3.1.6	Analyse av data .....	39
3.1.7	Gruppering etter språkferdigheter .....	43

3.2	Validitet og reliabilitet i undersøkelsen.....	45
3.3	Etikk.....	47
4	RESULTATER .....	49
4.1	Innledning.....	49
4.2	Deskriptiv statistikk.....	49
4.2.1	Karakteristikk av fordelinger .....	49
4.2.2	Reliabilitetsanalyse.....	54
4.2.3	Korrelasjoner .....	54
4.2.4	Hvordan samsvarer måleinstrumentene med hverandre?.....	56
4.2.5	Prediktiv verdi .....	60
4.2.6	T-test for uavhengige utvalg .....	62
4.2.7	Enkel deskriptiv analyse.....	62
4.2.8	Hva kjennertegner barn med mulig SSV ifølge 20 spørsmål?.....	65
4.2.9	Hva kjennetegner barn med mulig SSV ifølge TROG?.....	70
4.2.10	Hva kjennetegner barn med mulig SSV ifølge BPVS?.....	71
5	DRØFTING.....	74
5.1	Drøfting av resultater i undersøkelsen.....	74
5.1.1	Måler TROG og BPVS alle aspekter ved begrepet SSV?.....	74
5.2	Gjør testene nøyaktige målinger? .....	77
5.2.1	Casestudier som viser ulike utfall av en test .....	78
5.2.2	Drøfting av forhold som kan føre til underrapportering .....	82
5.2.3	Drøfting av forhold som kan føre til overrapportering .....	87
5.3	Drøfting av validiteten og reliabiliteten til undersøkelsen .....	88
5.3.1	Drøfting av statistisk validitet .....	88
5.3.2	Drøfting av begrepsvaliditeten .....	89
5.3.3	Ytre validitet.....	92
5.3.4	Andre forhold som kan ha påvirket validiteten og reliabiliteten.....	92
5.3.5	Oppsummering av undersøkelsen .....	94
5.3.6	Konklusjon .....	96
6	VEIEN VIDERE .....	98
	Register.....	99
	Litteraturliste .....	100
	Vedlegg 1: 20 spørsmål.....	105

Vedlegg 2: enkel deskriptiv analyse.....	106
Vedlegg 3: beskrivelse av TROG-2 .....	108
Vedlegg 4: beskrivelse av BPVS II.....	109
Vedlegg 5: beskrivelse av Raven SPM .....	110





# 1 INNLEDNING

## 1.1 Bakgrunn, tema og formål

Denne masteroppgaven er tilknyttet forskningsprosjektet Kunnskapsgenerering i det spesialpedagogiske praksisfeltet (KiSP) ved Institutt for Spesialpedagogikk, som har som ambisjon å forskningsforankre erfaring og innsikt i det spesialpedagogiske praksisfeltet. Formålet med dette arbeidet er å bidra til å heve kvaliteten på den spesialpedagogiske profesjonsutdanningen (KiSP, 2009). Som masterstudenter har vi fått anledning til å bli med i første fase av forskningsprosjektet, og har bidratt til å kartlegge den språklige kompetansen til elever ved to barneskoler. Barnas språkferdigheter er blitt kartlagt både ved hjelp av tester og et spørreskjema der læreren vurderer symptomer på språkvansker for hvert barn. Data fra kartleggingen som er foretatt som en del av KiSP-prosjektet, benyttes også i denne undersøkelsen.

I vår undersøkelse har vi valgt å se nærmere på hvordan en kan identifisere barn med spesifikke språkvansker (SSV) ut fra prestasjoner på tester og observasjoner av hvordan barn fungerer språklig. Formålet med dette er å avdekke om de to språktestene TROG-2 og BPVS II (heretter kalt TROG og BPVS) gir en økologisk valid identifisering av barn med spesifikke språkvansker i utvalget vårt. Vi kontrollerer dette ved å undersøke om de standardiserte testene samsvarer med læreres vurdering av elevenes språkferdigheter gjennom spørreskjemaet 20 spørsmål om språkferdigheter.

De fleste barn gjennomgår en normal språkutvikling. Det finnes imidlertid noen barn som ikke tilegner seg et funksjonelt språk på en adekvat måte, uten at faktorer som vanligvis ledsager vansker med språklæring, slik som hørselstap, lav nonverbal IQ og nevrologisk skade, er fremtredende. For disse barna er språket den primære vansken. Slike vansker karakteriseres som spesifikke språkvansker (Bele, 2008; Leonard, 2000).

SSV betegnes som et skjult handikap, og det er flere grunner til det. For det første er det ofte vanskelig å se eller høre at et barn har et slikt handikap. For det andre er det vanskelig å forstå og diagnostisere et slikt handikap. Det fører til at noen barn må slite med problemene alene. Det blir anslått at 5-7 % av alle barn har spesifikke språkvansker (Leonard, 2000; Ottem & Lian, 2008a). Det vil si at sannsynligvis ett til to barn i hver skoleklasse har spesifikke

språkvansker, og hos mange av disse vil vansken blant annet føre til problemer med lesing og skriving. Barna risikerer også å utvikle adferdsmessige, sosiale og psykiske problemer (Hughes, Turkststra, & Wulfeck, 2009; Ottem & Lian, 2008a).

Spesifikke språkvansker kan manifestere seg på flere ulike måter, og det er liten konsensus om hva som kjennetegner barn med SSV (Conti-Ramsden, Crutchley, & Botting, 1997). Det finnes ingen universell definisjon av spesifikke språkvansker som forskerne er enige om (Aram, Morris, & Hall, 1992). Som gruppe defineres barn med spesifikke språkvansker ofte ikke ved egenskaper de har, men enten ved eksklusjonskriterier som angir at de er normale bortsett fra svekket språklig fungering (Fletcher & Ingham, 1995), eller ved diskrepanskriterier som angir at det er en forskjell mellom språklig fungering på den ene siden og andre ikke svekkede sider ved kognitiv fungering på den annen (Aram, Morris, & Hall, 1993). Bruken av eksklusjonskriterier kan tenkes å tilsløre at det er tale om en heterogen gruppe, og påvirker hvordan vi beskriver barn med SSV (Conti-Ramsden, et al., 1997).

Problemet med å finne en god definisjon på spesifikke språkvansker hos barn har gjort det vanskelig å identifisere slike vansker på en reliabel måte (Johnston, 1988; Stark & Tallal, 1981). Aram, Morris & Hall (1993) sammenlignet metoder for identifisering av barn med SSV. De undersøkte klinisk identifisering av SSV i forhold til diskrepanskriterier av ulike typer, og konkluderte med at ingen enkeltstående metode var i stand til å fange opp den kliniske gruppen av barn med SSV.

TROG ble lansert i norsk versjon i 2009, og norsk standardisering av BPVS er under utarbeiding. Testmanualene gir dokumentasjon for testenes innholds- og kriterievaliditet, men dokumenterer ikke testenes validitet med tanke på identifisering av barn med spesifikke språkvansker. Dette er bakgrunnen for at vi ønsker å undersøke hvor godt disse testene er egnet til å identifisere barn med spesifikke språkvansker i vårt utvalg. Selv om språktester kan være i stand til å avdekke forskjellige ferdighetsnivåer hos barn, er de ikke nødvendigvis i stand til å identifisere barn med spesifikke språkvansker (Gray, Plante, Vance, & Henrichsen, 1999).

Tradisjonelt er tester brukt for å gi et mål på hvor langt et barn er kommet i læring og utvikling på testtidspunktet. Tester har sitt utspring i det som kan kalles en medisinsk tradisjon, der hovedfokus er sykdom eller diagnose. I denne tradisjonen kan språk anses som et system der interaksjonen er et sekundært fenomen. Dette språksynet er formalistisk eller



monologisk (Linell, 1998). I et monologisk språksyn rettes fokuset mot språkstruktur og individuell språkkompetanse, og kommunikasjon anses som isolerte handlinger som utveksles mellom individer. Barns språkferdigheter kan til en viss grad måles med standardiserte tester. Det er imidlertid en svakhet ved standardiserte tester at de ofte viser mer globale mål på ekspressive og reseptive ferdigheter, og ikke i så stor utstrekning kan avdekke de spesifikke vanskene hvert enkelt barn har. De har kanskje ikke den ønskede kliniske verdien som trengs for å tilrettelegge en godt tilpasset intervensjon for hvert enkelt barn (Lyster, 2008). Haywood (1997) mener problemet ikke ligger i hva de standardiserte testene måler, men hva de ikke måler. Det vil alltid være nødvendig å kombinere mål på kognitiv og språklig fungering med observasjoner av hvordan barnet fungerer i dagliglivet (Lie & Tetzchner, 2010). Selv om språktester er reliable, kan de mangle validitet dersom barn med svak skåre på testene ikke vekker bekymring hos lærere eller foreldre i hverdagen, også kalt overrapportering (Bishop & McDonald, 2009). Det samme er tilfellet dersom lærere og foreldre fanger opp språkproblemer som testene ikke identifiserer. Dette kalles underrapportering.

20 spørsmål om språkferdigheter er en liste med utsagn som beskriver symptomer på språkvansker (Ottem, 2009). Skjemaet skal fylles ut av en lærer eller andre nærpå personer som kjenner barnet godt. Måleinstrumenter av denne typen kan ha sitt utspring i det som kan kalles en sosial tradisjon som legger vekt på hvordan personer fungerer i samspill med andre. I en slik tradisjon anses språk som funksjonalistisk eller dialogisk. Dialogismen understreker at interaksjon alltid skjer i en kontekst (Lind, 2002). Da vektlegges kommunikasjonsaspektet og formen anses for å være sekundær (Linell, 1998). Spørreskjemaet kan sies å ha økologisk validitet, da lærerne vurderer eleven i en naturlig kontekst uten at barnet skal prestere noe under bestemte betingelser. Ifølge Botting, McCracken, & Conti-Ramsden (2002) kan det være at lærere og lærervurdering er i stand til å identifisere vansker som ikke fanges opp av standardiserte språklige og kognitive tester. I så fall kan lærevurdering fungere som en valid klinisk vurdering av et barns språklige fungering.

Selv om både klinisk vurdering og psykometriske formler er blitt benyttet til å finne diskrepans mellom språk og andre aspekter i utviklingen, ser det ut til å være en økende tendens til å kreve at psykometriske formler brukes (McCauley & Demetras, 1990). Standardiserte psykometriske diskrepanskriterier er mer restriktive og kanskje mindre sensitive overfor språkvansker enn klinisk vurdering basert på et barns språk i en naturlig kontekst (M. Dunn, Flax, Sliwinski, & Aram, 1996). I en undersøkelse av nyere dato stiller

Bishop & McDonald (2009) spørsmål ved om tradisjonelle språktester mangler sensitivitet overfor klinisk viktige kjennetegn på språkvansker.

Denne bakgrunnen danner utgangspunkt for problemstillingen i undersøkelsen vår:

*I hvilken grad identifiserer TROG og BPVS barn med spesifikke språkvansker på en måte som er valid i forhold til læreres beskrivelse av vanskene?*

For å svare på problemstillingen stiller vi følgende forskningsspørsmål underveis:

- Gjør testene nøyaktige målinger? Vi vil belyse dette spørsmålet ved å undersøke sammenhengen mellom språktestene og spørreskjemaet 20 spørsmål om språkferdigheter. Dette spørsmålet handler om testenes kriterievaliditet.
- Måler testene TROG og BPVS alle aspekter ved begrepet spesifikke språkvansker? Vi vil belyse dette spørsmålet ved å undersøke sammenhengen mellom BPVS, TROG og Raven. Dette spørsmålet handler om språktestenes begrepsvaliditet.
- Kan et eventuelt misforhold mellom språktestene og spørreskjemaet 20 spørsmål om språkferdigheter forklares med forhold knyttet til økologisk validitet?

## 1.2 Oppgavens oppbygging

Oppgaven er delt inn i seks kapitler:

Kapittel 1: Presenterer formål, bakgrunn og problemstilling for undersøkelsen. I tillegg gis en presisering av terminologi og vedlegg som benyttes i vår undersøkelse.

Kapittel 2: Redegjør for hva spesifikke språkvansker er og hvilke tilnærminger som kan brukes til å identifisere barn med SSV. Til slutt blir viktige aspekter knyttet til testvaliditet beskrevet.

Kapittel 3: Redegjør for ulike metodiske tilnærminger til vår undersøkelse, herunder en beskrivelse av validitet og reliabilitet. I forbindelse med validitet blir de fire validitetstypene til Cook og Campbell (1979) presentert. Avslutningsvis blir etiske aspekter knyttet til prosjektet diskutert.

Kapittel 4: Presenterer resultater fra undersøkelsen.

Kapittel 5: Drøfter resultater i undersøkelsen opp mot teori og empiri, for deretter å drøfte validiteten og reliabiliteten til undersøkelsen. Til slutt presenteres en konklusjon der problemstillingen besvares.

Kapittel 6: Presenterer tanker om hvordan identifisering av SSV kan forbedres.

### 1.2.1 Presiseringer om bruk av terminologi

Begrepet *spesifikke språkvansker (SSV)* refererer til en betydelig svekkelse av språkferdighetene uten at faktorer som vanligvis ledsager vansker med språklæring, slik som hørselstap, lav nonverbal IQ og nevrologisk skade, er fremtredende (Leonard, 2000). Barn med spesifikke språkvansker utvikler seg ellers normalt. Hos disse barna blir språkvansken det primære problemet (Ottem & Lian, 2008a).

Av hensyn til språklig variasjon brukes begrepet *spesifikke språkvansker* og forkortelsen *SSV* om hverandre.

Med *identifisering* i denne oppgaven mener vi å avdekke om et barn har spesifikke språkvansker. For å kunne avdekke dette må man innhente informasjon om barnets språkferdigheter og/eller språklige fungering i hverdagen.

Med *testvaliditet* menes om en test måler det den skal måle (McCauley, 2001; Sattler, 2001).

Med *psykometrisk* menes kvantitativ kartlegging av en persons psykologiske trekk eller egenskaper (Sattler, 2001).

*Økologisk validitet* vil si å beskrive og forstå aktiviteter i en naturlig kontekst (Svartdal, 2009). Spørreskjemaet 20 spørsmål kan sies å være økologisk valid, fordi det gir indikasjoner om hvordan barn fungerer språklig i naturlige situasjoner.

Ved omtale av spørreskjemaet 20 spørsmål om språkferdigheter veksles det mellom betegnelser som *spørreskjema*, *20 spørsmål* og *20 spørsmål om språkferdigheter*.

TROG-2 , BPVS II og Raven SPM blir i denne oppgaven omtalt som TROG, BPVS og Raven.

## **1.2.2 Presiseringer om vedlegg**

Spørreskjemaet 20 spørsmål om språkferdigheter, språktestene TROG og BPVS og den nonverbale testen Raven er beskrevet i metoddelen (avsnitt 3.1.4), men vi har også valgt å ta med en kopi av spørreskjemaet og beskrivelse av alle måleinstrumentene som vedlegg. Dette har vi gjort med tanke på lesere som ikke kjenner disse måleinstrumentene så godt, fordi vi omtaler de ulike måleinstrumentene før det gis noen beskrivelse av dem. For lesere som ikke er kjent med de ovennevnte måleinstrumentene, anbefales det derfor å lese disse vedleggene først (se innholdsfortegnelsen).

## 2 TEORETISK BAKGRUNN

I dette kapitlet redegjøres det for hva spesifikke språkvansker er og hvilke tilnærminger som kan brukes til å identifisere barn med språkvansker. Viktige aspekter knyttet til testvaliditet blir også beskrevet.

### 2.1 Hva er spesifikke språkvansker?

#### 2.1.1 Beskrivelse av fenomenet

Begrepet ”spesifikke språkvansker” (SSV) refererer til en betydelig svekkelse av språkferdighetene uten at faktorer som vanligvis ledsager vansker med språklæring, slik som hørselstap, lav nonverbal IQ og nevrologisk skade, er fremtredende (Leonard, 2000). Barn med spesifikke språkvansker utvikler seg ellers normalt. Hos disse barna blir språkvansken det primære problemet (Ottem & Lian, 2008a).

#### 2.1.2 Forekomst

Det anslås at ca. 5-7 % av alle barn har SSV, og at vansken forekommer tre ganger så hyppig hos gutter som hos jenter. Det finnes mange estimater av forekomsten av SSV i litteraturen, men dokumentasjon for dataene foreligger ofte ikke (Leonard, 2000; Ottem & Lian, 2008a). Ved en studie av femåringer ble det funnet en forekomst av SSV på 7,4 % (Tomblin, 1996, ref. i Leonard, 2000). For mange barn med SSV vedvarer språkvansken. De fleste barn med SSV utvikler seg språklig med tiden, men mange har fortsatt språklige svakheter i sen barndom, ungdom og noen også i voksen alder (Leonard, 2000). Noen barn vokser tilsynelatende vanskene av seg, mens andre ikke gjør det (Ottem & Lian, 2007). Van der Lely & Battel (2003, ref. i Ottem & Lian, 2007) mener at barn som vokser språkvanskene av seg, ikke har spesifikke språkvansker, men en forsinket språkutvikling. Ifølge (Ottem & Lian, 2008a) forekommer en forsinket språkutvikling (både spesifikke språkvansker og språkvansker som er sekundære i forhold til andre vansker) i en mindre andel i løpet av førskolealderen. En studie gjort av Whithurst og Fischel (1994) viser at antallet barn med forsinket språkutvikling synker med økende alder. Ifølge disse forskerne har mange barn forsinket språkutvikling i førskolealder, men mange av disse tar igjen sine jevnaldrende innen skolestart. Man kan

derfor ikke vite med sikkerhet om det er tale om spesifikke språkvansker eller en forsinket språkutvikling før barna er i femårsalderen.

### **2.1.3 Utvikling i synet på spesifikke språkvansker**

Barn med uforklarlige vansker med å tilegne seg språk, har fått mange ulike merkelapper i årenes løp (Bishop, 1997; Kamhi, 1998). Den stadig skiftende terminologien som er blitt brukt for å beskrive disse barna, reflekterer delvis utviklingen i vår forståelse av vansken (Friel-Patti, 1999). Generelt sett er det slik at jo mindre forståelse man har av en tilstand, jo mer varierende og inkonsekvent er terminologien vi bruker når vi refererer til den (Bishop, 1997). Spørsmålet om hva man skal kalle barn med spesifikke språkvansker, har skapt stor forvirring og diskusjon (Kamhi, 1998). Problemet med termer er at man ikke vet om de gir en presis betegnelse av fenomenet (Bishop, 1997). Den stadig skiftende terminologien kan ha ført til at man lett kan ha mistet interessante detaljer i studier av disse barna (Leonard, 2000). I det følgende presenteres noen hovedtrekk i utviklingen i synet på spesifikke språkvansker sammen med eksempler på den skiftende terminologien som er blitt brukt, for å illustrere hvordan forståelsen av fenomenet har endret seg.

I grove trekk kan utviklingen i synet på SSV skisseres slik: I begynnelsen hadde forskerne hovedfokus på språkproduksjon. Allerede i 1822 publiserte Gall en beskrivelse av barn som hadde klare problemer med språket, men som ikke viste tegn til andre kjente vansker: "There are many children ... who do not speak to the same degree as other children although they understand well or are far from being idiotic" (Gall, 1822, ref. i Leonard, 2000, p. 5). Med tiden ble det gradvis gjort et skille mellom barn som hadde vansker med produksjon og de som hadde vansker med både forståelse og produksjon, selv om begge typer vansker ble innlemmet i kategorien SSV. Barn med en medfødt hjerneskade ble etter hvert utelukket fra gruppen. Etter hvert ble for øvrig kriteriene for SSV snevret inn for å sikre at barn med tydelig svekket nonverbal intelligens, hørsel og oralmotorikk ikke ble innlemmet i gruppen (Leonard, 2000). Liebmann (1898) referert i Leonard (2000) kan ha vært den første til å beskrive undergrupper av barn med spesifikke språkvansker. Han beskrev en gruppe barn med motoriske vansker, mens andre grupper besto av barn som bare oppfattet enkeltord og barn som manglet evne til å oppfatte språk. Siden har forskerne gjort mange forsøk på å finne undergrupper, uten at man har kunnet enes om dem.

Når det gjelder vanskens natur ble det tidlig fremsatt en teori om at det ikke dreide seg om en nevrologisk vanske, men om et funksjonelt problem (Cøen, 1886, ref. i Leonard, 2000).

Vansker knyttet til oppmerksomhet og minne ble for øvrig allerede på 1800-tallet ansett som viktige årsaker til språkvansker (Treitel, 1893, ref. i Leonard, 2000). Eisenson (1968) referert i Bishop (1997) mente at barn med SSV hadde problemer med å identifisere språklyder, og anså fenomenet som et grunnleggende problem med å skille mellom lyder. Han har senere fått støtte for sitt syn gjennom eksperimentelle bevis som tyder på at SSV kan skyldes en auditorisk prosesseringsvanske (for eksempel Tallal, 1976). Vi kommer tilbake til hvilke årsaksteorier som er mest aktuelle i dag.

Tidlig på 1900-tallet ble termer som "congenital aphasia" brukt i stor grad i engelsk og fransk litteratur for å beskrive barn med språkvansker. Vansker med både å forstå og produsere språk var inkludert i denne kategorien. På begynnelsen av 1920-tallet begynte man å bruke dysfasi i tillegg til termen afasi. En grunn til denne gradvise forandringen av begrep var at dysfasi var mer nøyaktig, fordi forstavelsen *dys* betyr at det er problemer med språket, mens forstavelsen *a* impliserer fravær av språk. Etter hvert gikk man bort fra å bruke begge disse termene. En grunn til det var at de var knyttet til nevrologisk betingede vansker i forbindelse med hjerneskade. En annen grunn var at man ble opptatt av å beskrive lingvistiske trekk ved barns språk. Termene som ble brukt, inneholdt ofte ordet "language" sammen med en beskrivelse av en ledsagende vanske (Leonard, 2000). På 1960-tallet økte den faglige interessen for barn med spesifikke språkvansker med The Stanford Conference on Childhood Aphasia (Kamhi, 1998). Fra dette tidsrommet ble det også tatt i bruk et vell av ulike termer for samme fenomen i den engelske litteraturen. Noen eksempler er "delayed speech", "language disorder" og "language impairment" (Leonard, 2000). I dag brukes ofte termen "specific language impairment" (forkortet SLI) om fenomenet på engelsk og spesifikke språkvansker (forkortet SSV) på norsk. Ordet spesifikk betyr at språket er den primære vansken, og at ingen annen vanske foreligger (Bishop, 1997). Den varierende begrepsbruken fortsetter også i dag. Et eksempel på det finnes i 10. revisjon av den norske utgaven av diagnosemanualen ICD-10, der vansken kalles spesifikke utviklingsforstyrrelser av tale og språk (WHO, 2000).

Spørsmålet om hvorvidt barn med spesifikke språkvansker tilegner seg språket senere enn andre barn eller har selektive vansker, har vært mye debattert. I over 30 år har forskere forsøkt å avgjøre om de lingvistiske kjennetegnene til barn med SSV er de samme som hos yngre

normalspråklige barn, eller om de er av en helt annen karakter. Disse spørsmålene diskuteres fortsatt i nyere forskning (Leonard, 2000).

#### **2.1.4 Lingvistiske kjennetegn hos barn med spesifikke språkvansker**

Språk er et komplekst fenomen, og barn med SSV kan ha vansker innen flere lingvistiske dimensjoner (Leonard, 2000; Rice, 2004). Det er ofte disse klinikerer ønsker å avdekke. I det følgende vil en kort beskrivelse av semantiske, fonologiske, morfologiske og pragmatiske vansker hos barn med SSV bli presentert.

##### **Semantiske vansker**

Språkets semantikk viser til språkets innhold og betydning (Lyster, 1994). Semantikk knyttes ofte til ordlæring, begrepsoppbygging og ordforråd. Det ser ut til at barn med SSV utvikler ordforrådet saktere enn sine jevnaldrende. De trenger flere gjentakelser for å lagre ord i langtidsminnet, noe som knyttes til kapasiteten for bearbeiding og minne. Videre kan de ha vansker med å knytte mening til ordets lydklang på en nøyaktig og effektiv måte, samt det å utnytte syntaktisk kunnskap for å tolke mening i det de hører. Svakere og færre semantiske assosiasjoner og mindre effektiv og nøyaktig gjenkalling av ord fra langtidsminnet er også noe som kan observeres hos disse barna. (Bishop, 1997; Leonard, 2000). Innholdsord som substantiv og verb ser ut til å være mer komplisert å lære enn funksjonsord som adjektiv og preposisjoner. Barn med SSV skårer lavere enn sine jevnaldrende på ordforrådstester (Rice, 2004). De kan også streve med ordletingsvansker, noe som er lettest å registrere i skolealder. Symptomer på dette kan være at de har uvanlig lange pauser og utstrakt bruk av ord som ”dings” og ”sånn” (Leonard, 2000). Rygvold (2008) nevner at de kan omskrive ord, slik at for eksempel kjole blir jentegenser. Språkforståelsen til disse barna kan bli dårligere fordi de tilegner seg nye ord sent, og de kan få vansker med å uttrykke tanker og ideer, hendelser og relasjoner.

##### **Fonologiske vansker**

Barn med fonologiske vansker har problemer med å tilegne seg lydsystemet i språket. Dette er et kognitivt fenomen og ikke det samme som artikulasjonsvansker, der vansken er å artikulere språklyder. Forveksling og forenkling av lyder i ord som er vanlig hos normalspråklige barn



fram til treårsalderen, kan hos barn med SSV vedvare opp i skolealder. Barn som er lite fonologisk bevisste når de nærmer seg skolealder, kan ha fonologiske vansker (Rygvold, 2008).

### **Grammatiske vansker (morfologi og syntaks)**

Morfologi kan betegnes som læren om formverket i et språk (Wangensteen, 2005), og har med språkets grammatikk å gjøre. Morfologiske vansker ser ut til å være karakteristisk for barn med SSV uavhengig av språk. Dette er vist ved studier i engelskspråklige land, men også norske studier har vist det samme (Bjerkan, 2000). Barn med SSV bruker bøyninger tilfeldig og på en lite konsekvent måte. De gjør mange morfologiske feil og bruker færre ordbøyninger, noe som gjør at språket ligner på yngre barns språk. De har større problemer med morfologi enn normalspråklige barn, og studier peker på at barn med SSV som gruppe ikke alltid tar igjen sine jevnaldrende på dette området (Rice, 2004).

Den syntaktiske utviklingen, det vil si å sette sammen ord til lengre ytringer, ser også ut til å skje seinere hos barn med SSV enn hos barn på samme alder. Dersom ytringen inneholder mye informasjon, kan det også være vanskelig for dem å forstå setninger der syntaksen er enkel (Bishop, 1997). Barn med SSV bruker ofte grunnmønsteret subjekt-vebal-objekt mer enn det som er vanlig. Ordrekkefølgen utgjør ikke det største problemet for disse barna i hverdagen, men i en testsituasjon vil dette problemet bli tydeligere (Rygvold, 2008).

### **Pragmatiske vansker**

Pragmatikk er språket i bruk og tolkning av språklige ytringer i forskjellige situasjoner (Sveen, 2005). Barn som har ordforråd, språkforståelse og grammatisk kunnskap som ikke er aldersadekvat, kan streve med å bruke språket i sosialt samspill med andre. De kan ha vansker med å ta andres perspektiv og vite hva mottakeren trenger av opplysninger for å forstå det som blir sagt. Barn med SSV kan oppleves som vanskelige å snakke med, da de kan ha en tendens til å ta lite språklig initiativ, ikke gi adekvate svar og ikke reparere dialogen ved å stille oppfølgingsspørsmål. De pragmatiske vanskene blir ofte sekundære i forhold til vanskene disse barna har med semantikk og morfologi (Rice, Warren, & Betz, 2005).

Selv om spesifikke språkvansker manifesterer seg ulikt, finnes det noen trekk som er typiske: Barn med disse vanskene begynner ofte sent å snakke, bruker enkle grammatiske strukturer

(utelater fortid eller hjelpeverb langt utover den alderen da dette vanligvis beherskes), de har et begrenset vokabular, både i produksjon og forståelse, et svakt verbalt korttidsminne og vansker med å forstå komplekst språk, spesielt ved rask tale (Bishop, 2006).

Som vi ser over kan barn med SSV ha vansker innen ulike lingvistiske dimensjoner. Språk er et komplekst fenomen, og vanskene kan ytre seg forskjellig fra individ til individ. For å kunne identifisere barna på en tilfredsstillende måte og for å kunne gi barnet best mulig tilpasset oppfølging har man forsøkt å identifisere undergrupper av SSV.

### **2.1.5 Undergrupper**

De fleste forskere og klinikere er enige om at barn med SSV er en heterogen gruppe, men det har likevel ikke vært mulig å enes om en inndeling i undergrupper (Bishop, 1998). En inndeling som var vanlig tidligere og som også delvis består i dag, er et skille mellom ekspressive og reseptive former for språkvansker. Et eksempel på dette finner vi i ICD-10, der overbegrepet spesifikke utviklingsforstyrrelser av tale og språk blant annet deles inn i spesifikk artikulasjonsforstyrrelse, ekspressiv språkforstyrrelse og impressiv (reseptiv) språkforstyrrelse. Spesifikk artikulasjonsforstyrrelse vil si at barnets bruk av språklyder er under det som er aldersadekvat, mens språkferdighetene ellers er normalt utviklet. Med ekspressiv språkforstyrrelse menes at barnets evne til å uttrykke seg verbalt er betydelig under det som er ventet på grunnlag av barnets mentale alder, og at språkforståelsen er normal. Denne vansken kan opptre samtidig som artikulasjonsforstyrrelsen. Med impressiv språkforstyrrelse menes at språkforståelsen er under det som forventes ut fra barnets mentale alder. Ifølge diagnosemanualen forekommer en av eller begge de ekspressive vanskene nesten alltid samtidig med de reseptive vanskene (WHO, 2000). Inndelingen over er blitt kritisert av blant annet Bishop (1997), da mye tyder på at barn med ekspressive vansker ofte også har reseptive vansker, selv om de ekspressive vanskene ofte er mest fremtredende. Det er for eksempel lettere å oppdage avvik i barns ekspressive språk enn å vite hva barn egentlig forstår, men det betyr ikke at de har rene ekspressive vansker (Bishop, 1997). Når det gjelder barn med språklydvansker har de ofte svakheter på andre språkområder i tillegg, og det er derfor blitt stilt spørsmål ved om dette kan betraktes som en egen undergruppe (Leonard, 2000).

Det er gjort flere forsøk på å finne undergrupper av barn med SSV ved hjelp av ulike statistiske metoder (Aram & Nation, 1975) og kliniske beskrivelser av barns språkvansker

(Rapin & Allen, 1983). Ifølge Bishop (1997) er det svært få klinikere som bruker de kategoriene som ble funnet i de ovennevnte studiene. Det er også gjort forsøk på å finne valide undergrupper ved å kombinere statistiske og kliniske metoder (Conti-Ramsden, et al., 1997). En longitudinell studie av den samme gruppen barn identifiserte tre undergrupper: ekspressiv SSV, ekspressiv/reseptiv SSV og kompleks SSV. Den sistnevnte gruppen besto av barn som hadde leksikalske, syntakstiske, semantiske og pragmatiske vansker uten å ha fonologiske vansker (Conti-Ramsden & Botting, 1999a). Lingvistisk baserte klassifikasjoner virker mer lovende enn et skille mellom reseptivt og ekspressivt språk, men også disse trenger å forbedres og valideres (Bishop, 1997). Problemet med å identifisere undergrupper kan skyldes at språkvanskene kan forandre seg markert med alderen, slik at den samme vansken kan opptre noe ulikt i utviklingsforløpet. Barn med SSV kan derfor gå fra å tilhøre en undergruppe til en annen (Bishop, 1997; Conti-Ramsden & Botting, 1999a).

Vi har nå sett at forskerne har gjort mange forsøk på å finne undergrupper av SSV for å få en bedre forståelse av fenomenet. Et annet spørsmål som er mye undersøkt i forskningslitteraturen er spørsmålet om hva som kan være årsaken til SSV.

### **2.1.6 Arv**

Det er mange teorier om hva som kan være årsaken til SSV, og mye tyder på at vansken kan ha en betydelig arvelig komponent. SSV opptre tre ganger så ofte hos gutter som hos jenter (Leonard, 2000), og ulike studier har vist at mellom en tredel og halvparten av barna med SSV har foreldre eller søsken som har eller har hatt språklæringsvansker (Leonard, 2000; Rice, 2004). I tvillingstudier ble det funnet en konkordansrate (det vil si at tvillingene er like når det gjelder vansken) hos eneggede tvillinger på 80-86 %, mens toeggede tvillinger hadde en konkordansrate på 38-48 % (Bishop, North, & Donlan, 1996). Problemet er ikke alltid genetisk, og i stedet for å lete etter én enkelt underliggende faktor, bør SSV betraktes som et resultat av to eller flere underliggende vansker (Bishop, 2006). Slike underliggende vansker kan ha en tilknytning til ulike kognitive faktorer, og noen av disse beskrives i neste avsnitt.

### **2.1.7 Kognitive faktorer**

Ulike kognitive mekanismer kan lede til spesifikke språkvansker hos barn. Kunnskap om slike mekanismer kan være viktig for å spesifisere gode markører for SSV. Gode markører kan gjøre det mulig å predikere språkvansker og gi en bedre spesifisering av

inkludjonskriteriene for SSV (Ottem & Lian, 2008b). I det følgende presenteres først fire av de mest kjente teoriene om kognitive faktorer som kan være underliggende årsaker til spesifikke språkvansker, og dernest ulike markører for identifisering av SSV.

Begrepet spesifikke språkvansker innebærer at barn har svake språkferdigheter mens andre aspekter ved kognitiv fungering er normale. Det er imidlertid belegg for å si at vanskene omfatter mer enn språket (Ottem, 2004).

Tallal (1976) har utviklet en teori om at en grunnleggende perseptuell svikt kan ligge til grunn for språkvansker. Hun hevder at det for både en normal og avvikende språkutvikling finnes såkalte kritiske hjernemekanismer som muliggjør all persepsjon av temporale sekvenser. Hørselen, som er den viktigste tidssansen, hjelper oss til å skille mellom lyder og huske rekkefølgen av lyder som kommer sekvensielt og i rask rekkefølge. Det auditive systemet må ha en god temporal oppløsningsevne for at hjernen skal oppfatte strukturelle forskjeller. En svikt her vil kunne føre til en svikt i barnets språkutvikling. Barnet vil kunne få problemer med å oppfatte tale, og dette kan igjen føre til vansker med å lære nye ord (Tallal, 1976).

I motsetning til Tallal har Baddeley, Gathercole & Papagno (1998) en teori om at den ”fonologiske løkken”, som er en ”mekanisme” for læring av nye ord og ytringer, har en betydning for språktilegnelsen. En svikt her kan føre til en svak utvikling av barnets vokabular og især læring av nye ord, og det kan vanskeliggjøre læring av nytt språk gjennom hele livet.

Ikke alle forskere er enige om at en svikt i verbalt korttidsminne er årsak til SSV hos barn. Hulme og Roodenry (1995) mente at slike vansker er en konsekvens av mer generelle språkvansker. Brown og Hulme (1996) videreutviklet teorien og mente at vansker med produksjon av språk og persepsjon vil kunne føre til problemer med verbalt korttidsminne og språktilegnelse.

Noen forskere mener også at spesifikke språkvansker kan skyldes generelle begrensninger i kapasiteten til informasjonsbearbeidelse (Johnston, 1994; Leonard, 2000). Man antar at språkvanskene oppstår som en følge av en begrenset kapasitet til å bearbeide informasjon eller ineffektiv bearbeidelse av informasjon. På grunn av langsom informasjonsbearbeidelse kan det tenkes at barn med spesifikke språkvansker trenger mer tid enn sine jevnaldrende til å oppfatte informasjon (Montgomery, 2005). Et spørsmål som er blitt debattert innen forskning

på SSV, er i hvilken grad grammatiske vansker skyldes manglende grammatiske kunnskaper, og i hvilken grad de har sammenheng med bearbeidingsproblemer (Bishop, 1997).

Teoriene ovenfor har ulikt syn på hva som kan ligge til grunn for en språkvanske, og benytter derfor ulike markører til identifisering av SSV. Ifølge Tallals teori er ART (Auditiv Repetisjons Test) som går ut på informasjonsbehandling av lydsekvenser, en god markør (Tallal, 1976). Ifølge Baddely (2003) er bruk av nonordrepetisjonstester en god markør for SSV fordi denne typen tester belaster minnet for lydstrukturer i ord. Ifølge Brown og Roodenry (1995) kan ikke en spesifikk test pekes ut som en god markør, men tester på grunnleggende språklige funksjoner som semantikk og syntaks kan brukes. TROG og BPVS er eksempler på den sistnevnte typen markør.

Flere av teoriene enes om at SSV kan skyldes forsinket modning av områder i hjernen som har med språktilegnelsen å gjøre. ERP eller "event related brain potentials" er en relativt ny metode som måler hjernens reaksjoner på auditive stimuli. I en artikkel refererer McArthur & Bishop (2004) til nyere forskning som viser at modning av det auditive systemet fortsetter oppover i ungdomsalderen. En kan anta at prosessen er forsinket med rundt fire år hos barn med SSV, slik at det vil bli et gap mellom barnets kronologiske alder og språkalders som strekker seg langt oppover i tid, på tross av at barn med SSV utvikler seg hele tiden.

### **2.1.8 Konsekvenser av språk- og kommunikasjonsvansker**

En språk- og kommunikasjonsvanske kan gi en dyptgående og varig virkning i et barns liv (Ottem & Lian, 2008a). En språkvanske vedvarer ofte over lang tid, og barnet står i fare for å utvikle adferdsmessige, sosiale, skolerelaterte og psykiske problemer (Hughes, et al., 2009; Ottem & Lian, 2008a). I en studie av Tallal et al. (1996) ble det anslått at mer enn 80 % av barna med SSV kan få vansker med lesing. Vi har valgt å ikke utdype dette nærmere i vår undersøkelse, da det ikke er relevant for vår problemstilling.

Spesifikke språkvansker er en skjult vanske, og den kan få store negative ringvirkninger. Det er derfor viktig å identifisere barn med spesifikke språkvansker, slik at de kan få den oppfølgingen de trenger.

## 2.2 Identifisering av barn med SSV

Spesifikke språkvansker er et skjult problem som ofte oppdages når krav til kognitive funksjoner som oppmerksomhet, hukommelse, resonnering og språk øker – når læringstrykket øker (Bele, 2008). Det at spesifikke språkvansker er et skjult problem, gjør at en del tilfeller forblir uoppdaget. Nettopp fordi det er vanskelig å oppdage en språkvanske, er det viktig å ha gode måleinstrumenter for å identifisere problemet (Bishop, 1997). Identifisering av barns språkvansker er viktig av to årsaker: Identifisering er ofte første skritt på veien mot at barnet kan få hjelp. For det andre kan korrekt identifisering bidra til å forebygge de negative konsekvensene som kan være forbundet med en språkvanske (McCauley, 2001).

Det bør brukes ulike tilnærminger til identifisering og kartlegging av spesifikke språkvansker (Bishop, 1997; Sattler, 2001). Ifølge Sattler (2001) bør det være fire grunnpilarer i en undersøkelsesprosedyre: normerte standardiserte tester, uformelle kartleggingsprosedyrer, intervju og observasjoner. Det er viktig å identifisere språkvansken tidlig for at den ikke skal hemme barnets utvikling. I det følgende vil vi se nærmere på de to ulike tilnærmingene som benyttes til å identifisere barn med spesifikke språkvansker i denne undersøkelsen: psykometrisk tilnærming ved bruk av tester og økologisk tilnærming ved bruk av spørreskjemaet 20 spørsmål til lærerne.

### 2.2.1 En psykometrisk tilnærming til identifisering

I forskning er det vanlig å bruke standardiserte tester når det skal avdekkes om et barn har spesifikke språkvansker (Bishop & McDonald, 2009). Til dette formålet benyttes både språktester og nonverbale tester. Skårene på testene brukes som eksklusjons- og inklusjonskriterier, og disse vil bli beskrevet nedenfor. Fordelen ved å bruke tester er at de gir muligheten til å vurdere språkvansken i en standardisert sammenheng og se den i forhold til normative data. Det er ikke enkelt å avgjøre hvilke språkferdigheter som skal inkluderes når man skal definere en vanske, og det er liten konsensus om dette. Det er imidlertid enighet om at det er viktig å bruke tester med gode psykometriske egenskaper, det vil si standardiserte tester med god validitet og reliabilitet (Bishop, 1997). Disse egenskapene beskrives i avsnitt 3.2 Validitet og reliabilitet i undersøkelsen.

Det anbefales vanligvis å bruke instrumenter som tester flere ekspressive og reseptive språkfunksjoner (Bishop, 1997). I diagnosemanualer som ICD-10 gis det imidlertid ingen

retningslinjer om hvilke språktester som skal benyttes, selv om språktester kan ha ulik sensitivitet (Spaulding, Plante, & Farinella, 2006).

### **Inklusjons- og eksklusjonskriterier**

For å avgjøre om et barn har spesifikke språkvansker, er det vanlig å bygge på et sett av inklusjonskriterier (som beskriver språkvansken) og eksklusjonskriterier (andre utviklingsforstyrrelser) (Bishop, 1997; Leonard, 2000; Ottem & Lian, 2008a).

Inklusjonskriteriene fastsettes på grunnlag av skårer på standardiserte språktester, men det er store forskjeller mellom hvilke kriterier som brukes. Kriterier utviklet av Stark og Tallal har satt en standard for kriteriene som brukes i empirisk forskning. Kriteriene Stark og Tallal har anbefalt for å identifisere denne gruppen av barn, er at de skal ha en skåre på standardiserte språktester som er 6-12 måneder under kronologisk eller mental alder. De utelukket barn som hadde problemer med hørsel, barn der foreldre eller lærere rapporterte om betydelig emosjonelle vansker eller adferdsproblemer, barn med en nonverbal IQ på -1 SD eller lavere, og barn med nevrologiske vansker eller store artikulasjonsvansker/fonologiske vansker. Stark og Tallal brukte spesielt utvalgte tester til dette formålet, og anbefalte bruk av flere språktester (Plante, 1998). Selv om Stark og Tallal har satt standarden for bruk av inklusjons- og eksklusjonskriterier i forbindelse med identifisering av SSV, er det stor variasjon i kriteriene som brukes.

SSV beskrives i de internasjonalt anerkjente diagnosemanualene ICD-10 (International Classification of Diseases) som er utarbeidet av Verdens helseorganisasjon (WHO, 2000) og DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) som er utarbeidet i USA av American Psychiatric Association (APA, 2000). I det følgende konsentrerer vi oss først og fremst om ICD-10. I tilknytning til ICD-10 er det utarbeidet diagnostiske kriterier for forskning (WHO, 1993). Disse kriteriene skiller mellom en ekspressiv og en reseptiv språkvanske. For å få en SSV-diagnose etter disse kriteriene, kreves en skåre på en ekspressiv eller reseptiv språktest som er 2 standardavvik eller mer under populasjonens gjennomsnitt, og språkferdighetene skal være minst 1 SD under nonverbal IQ. Ved identifisering av en ekspressiv vanske skal reseptive ferdigheter ikke være mer enn 2 SD under barnets kronologiske alder. Denne cut-off-grensen er tilfeldig og uvalidert (Spaulding, et al., 2006).

Ifølge Leonard (2000) er et vanlig kriterium for identifisering av spesifikke språkvansker at barna skal ha en språktestskåre under gjennomsnittet (-1,25 SD eller lavere) ledsaget av en

nonverbal IQ innenfor normalområdet. For å sikre at vansker med språket det primære, ekskluderes barn med tilleggsvansker (som hørselshemming og nylige mellomørebetennelser) og barn som ikke fungerer godt sosialt.

Ifølge Spaulding, Plante & Farinella (2006) viste en gjennomgang av forskningsartikler utgitt i American Speech-Language-Hearing Association (ASHA) fra august 2003 til april 2004 at de fleste forskere satte cut-off-grensen på en språkskåre mellom 1 SD og 1,5 SD under Mean på en eller flere språktester. Denne operasjonaliseringen stemmer med en antakelse som er utbredt hos mange forskere, for eksempel Leonard (1991), om at barn med spesifikke språkvansker befinner seg i forlengelsen av normalområdet. Et problem med cut-off-grenser er imidlertid at de brukes på en hvilken som helst test, selv om det ikke foreligger evidens for at barn med språkvansker sannsynligvis vil få en tilsvarende eller lavere skåre (Spaulding, et al., 2006). Et misforhold mellom den fastsatte cut-off-skåren og skårer som vanligvis er forbundet med vansker på den valgte testen, kan føre til systematisk under- og overidentifisering av vansker (Plante & Vance, 1994).

Det er en stor diskusjon om Stark og Tallals metode med inklusjons- og eksklusjonskriterier fortsatt skal representere en standard for identifisering av barn med SSV. Fortsatt leter man etter metoder for å identifisere barn med denne vansken på en stabil, konsistent og nøyaktig måte (Plante, 1998). Flere forskere anbefaler å slutte å bruke vilkårlige cut-off-skårer, for i stedet å utvikle evidensbaserte cut-off-skårer for hver enkelt test (Spaulding, et al., 2006). En empirisk utviklet cut-off-skåre på en bestemt test ville oppnå høyest nivå av sensitivitet (dvs. prosent av barn med en språkvanske som diagnostiseres som barn med en vanske) og sensitivitet (dvs. prosent av barn med normalutviklede språkferdigheter som diagnostiseres som barn med normalutviklet språk) (Spaulding, et al., 2006).

### **Spesifikke språkvansker og nonverbal IQ**

Som eksemplene over viser, er det et vanlig kriterium for diagnosen SSV at nonverbal IQ skal være innenfor det normale variasjonsområdet, og grensen settes ofte til over 85. IQ-kriteriet er imidlertid omstridt, fordi vi foreløpig ikke vet nok om sammenhengen mellom IQ og språkvansker. Hvis en slik sammenheng er svak eller ikke finnes, bør man være forsiktig med å bruke nonverbal IQ som eksklusjonskriterium (Ottem & Lian, 2007). Fey, Long & Cleve (1994) fant at barn med nonverbal IQ over gjennomsnittet og barn med nonverbal IQ under gjennomsnittet hadde sammenlignbare språkvansker. Begge gruppene hadde også nytte av



behandling for vanskene (Fey, et al., 1994). Bishop (1997) beskriver lignende funn i sine tvillingstudier. Plante (1998) anbefaler at grensen settes ved en nonverbal IQ på 70 eller 75, selv om dette medfører at barn som ikke har diskrepans mellom språklige og nonverbale ferdigheter inkluderes i SSV-gruppen.

Bruk av et IQ-kriterium er mest omstridt i klinisk sammenheng, fordi IQ trolig ikke har noe å si for hvilken nytte barn har av behandling. Når formålet med diagnostisering er å avdekke hvilke barn som har behov for klinisk oppfølging, fraråder Bishop derfor å benytte strenge eksklusjonskriterier, fordi man da kan utelukke barn som kan dra nytte av behandling. I forskningssammenheng anses kriteriet som mer hensiktsmessig, fordi det bidrar til at man forsker på "rene" grupper (Bishop, 1997; Leonard, 2000).

### **2.2.2 En økologisk tilnærming til identifisering**

Det ligger noen begrensninger i å definere en vanske på grunnlag av statistiske kriterier, uten å vurdere i hvilken grad vansken hemmer barnets fungering i det daglige (Bishop, 1997). En test gir bare et bilde av barnets prestasjon på et bestemt tidspunkt, og gir ikke nødvendigvis et komplett bilde av barnets språkferdigheter. Det finnes mange faktorer som kan påvirke testresultatet, og en testsituasjon er kunstig. Det er derfor viktig at en test suppleres med andre former for utredning. Tester er et hjelpemiddel til kartlegging av språkets innhold og form, men favner ikke barnets daglige språk i samhandling med andre (Rygvold, 2008).

Urie Bronfenbrenner (1974, ref. i McCauley, 2001, s. 79) beskrev i sin tid forskning på barns utvikling som "the study of the strange behavior of children in strange situations for the briefest possible period of time". Denne uttalelsen rettet søkelyset mot en bekymring om at forskere ikke fanget opp essensielle faktorer som påvirker barn fordi de ikke studerte dem i naturlige situasjoner sammen med mennesker som sto dem nær. Bronfenbrenner forsøkte å forstå barns utvikling gjennom observasjoner av barn og deres nærpå personer i naturlige kontekster. Bronfenbrenner utviklet siden en bioøkologisk modell som la vekt på at både egenskaper ved barnet og konteksten barnet befinner seg i, påvirker utviklingsprosessen (McCauley, 2001).

I nyere tid har Verdens helseorganisasjon innført et klassifiseringssystem med et utgangspunkt som trolig er inspirert av Bronfenbrenners bioøkologiske modell.

Klassifiseringssystemet ICF (International Classification of Functioning, Disability and

Health) er et rammeverk for måling av helse og funksjonshemming som ble undertegnet av alle medlemslandene i mai 2001. I motsetning til diagnosesystemene ICD-10 og DSM-IV, som har sitt utspring i en medisinsk tradisjon der hovedfokus er sykdom eller diagnose, har ICF et økologisk perspektiv på kartlegging. Hovedfokus for ICF er ikke sykdom eller diagnose, men hvilken funksjonsevne personer har i samspill med omgivelsene. ICF legger vekt på å klassifisere menneskers funksjonsevne relatert til mulighet for deltakelse i sosialt liv. Evnen til å delta vil kunne avhenge av at omgivelsene tilrettelegges. ICF er ikke ment å erstatte den internasjonale sykdomsklassifikasjonen ICD-10, men å utfylle den (WHO, KITH, & helsedirektoratet, 2004).

Innen et økologisk perspektiv på kartlegging står en helhetlig vurdering sentralt, og økologisk validitet blir her viktig. Det vil si å beskrive og forstå aktiviteter i en naturlig kontekst (Svartdal, 2009). Økologisk validitet dreier seg om hvorvidt forskningssituasjonen ”ligner” på situasjonen man ønsker å si noe om. Fordelen med å vurdere barns adferd i en naturlig situasjon, er at barnet ikke er klar over at de blir observert, noe som kan eliminere muligheten for at adferden til personen som studeres påvirkes (Svartdal, 2009).

I denne undersøkelsen brukes spørreskjemaet 20 spørsmål om språkferdigheter som en økologisk tilnærming til kartlegging av språkvansker. Det skal fylles ut av noen som kjenner barnet godt og er i stand til å vurdere eleven i forhold til barn på samme alder. I vår undersøkelse er dette lærere. Som vi har sett kan lærervurdering fungere som en valid klinisk vurdering (Botting, McCracken, & Conti-Ramsden, 2002). Spørreskjema kan hjelpe lærere til å bli bevisst sin tause kunnskap slik at den kommer barnet til gode. Begrepet taus kunnskap vil si kunnskap som kommer til syne gjennom handlinger og kan være vanskelig å formulere i fullstendig språklig form (Florin & Johannessen, 1988). Ifølge Polanyi (1967) er slik kunnskap ikke verbalisert, men kunnskap som kan settes ord på. Hiim & Hippe (1998) sier at det ofte dreier seg om en individuell kunnskap og kunnskapstilegnelse, men det kan også være lite eksplisitt kunnskap som er allmenn innenfor yrkesgruppen. Spørreskjemaet 20 spørsmål hjelper lærerne til å vurdere barnets språkligferdigheter, men er ikke tilstrekkelig i seg selv til å gi en samlet vurdering av disse (Ottem, 2009). Spørreskjemaet 20 spørsmål kan sies å være økologisk valid, da lærerne vurderer eleven i en naturlig kontekst uten at barnet skal prestere noe under bestemte betingelser.

Bishop har utviklet Children's Communication Checklist (CCC) med tanke på å vurdere sider ved kommunikasjonsvansker som ikke så godt evalueres ved en standardisert test (Bishop,

1998). Dette er en sjekklister med utsagn om barns kommunikasjonsferdigheter som i likhet med 20 spørsmål skal fylles ut av barnets nærpersoner (Bishop & McDonald, 2009). En fordel ved instrumenter av denne typen er for det første at de gir muligheten til å innhente informasjon om daglig kommunikasjon fra noen som kjenner barnet godt, og for det andre at de kan omfatte en bredere del av barnets språk og kommunikasjon enn det tester gjør (Bishop & McDonald, 2009). Et problem med sjekklister er at informantene som fyller dem ut, kan ha varierende forutsetninger for å forstå utsagnene, og at svarene preges av informantenes subjektive vurderinger (Bishop & McDonald, 2009).

I 1997 skrev Bishop at hun ikke kjente til at noen har forsøkt å sammenligne testbaserte diagnoser av SSV til vurderinger ved foreldre, lærere, jevnaldrende eller barnet selv av hvordan kommunikasjonsvansker påvirker hverdagen (Bishop, 1997). Siden den gang er det gjort flere studier som har undersøkt denne sammenhengen, som for øvrig også er utgangspunktet for vår undersøkelse, og funn fra noen av dem presenteres i det følgende.

### **Forskning på lærervurderingers validitet**

I en studie fant Botting, McCracken & Conti-Ramsden (2002) at det var et misforhold mellom indikasjoner fra testene TROG, BPVS, WISC III og lærervurdering. Ifølge artikkelforfatterne kan det være at lærere er i stand til å identifisere vansker som ikke fanges opp av standardiserte språklige og kognitive tester. En sannsynlig mekanisme bak dette er at en kumulativ virkning av noen års vanske, eller en rekke ulike ferdigheter kan påvirke skoleferdighetene, selv om det ikke fanges opp av tester som er utviklet for å vurdere enkelttyper av språkvanske (slik som vokabular). En akkumulering av problemer på tvers av flere ferdighetsområder kan også føre til en svakere prestasjon på skoleprøver som krever flere av disse. Ifølge artikkelforfatterne kan lærervurdering i så fall fungere som en valid klinisk vurdering.

I forbindelse med standardiseringen av testen Språk 6-16, ble barnets lærere bedt om å rangere barnas språklige kompetanse, i tillegg til at de ble testet med Språk 6-16 av en annen person enn læreren. Det ble funnet en sterk sammenheng mellom lærernes vurdering av barnas språkkompetanse og testresultatene. Det var imidlertid ikke alltid samsvar mellom de to kartleggingsmetodene. I noen tilfeller mente læreren at barnet hadde store språkvansker selv om testresultatene var gode, og i andre tilfeller var testresultatene svake, mens læreren mente at eleven fungerte godt (Ottem, 2009).

I en studie fant Bishop, Laws, Adams og Norbury (2006) at Children's Communication Checklist var svært godt egnet til å identifisere barn med språkvansker, men at samsvaret mellom foreldres vurdering og læreres vurdering av barns vansker bare var moderat.

Vi har i det forgående sett på hvordan vi kan identifisere barn med SSV ved bruk av ulike tilnærminger som spørreskjemaer og tester. Tester står fortsatt sentralt i identifiseringen av barn med denne vansken. Det er derfor viktig å vite i hvilken grad tester er valide og reliable. Vi vil se nærmere på aspekter som har betydning for testers validitet og reliabilitet i det følgende.

## 2.3 Testers validitet og reliabilitet

Denne studien har som formål å undersøke validiteten til språktestene TROG og BPVS i forhold til kriteriet 20 spørsmål om språkferdigheter. I det følgende presenteres teori om testvaliditet, i tillegg til informasjon om de to testenes validitet og reliabilitet som er oppgitt i testmanualene samt forskning på testers sensitivitet og spesifisitet.

### 2.3.1 Testvaliditet

Testvaliditet dreier seg om hvorvidt en test måler det den skal måle, og dermed om den setter oss i stand til å gjøre tolkninger på grunnlag av testresultater. Testens utgiver er ansvarlig for å gi evidens for at testen er valid for bestemte formål, og den som bruker testen, er ansvarlig for valid bruk av testresultatene. I motsatt fall kan vi ikke stole på testresultatene. Nettopp dette gjør at det ikke er én enkelt type validitet som egner seg for alle testformål. Selv om en tests utvikler hevder at testen er valid, betyr det ikke at den er valid for alle formål. Validitet må sees i sammenheng med hva måleresultatet skal brukes til. Validitet er heller ikke noe absolutt, men noe en test har i ulik grad (Kleven, 2002a; McCauley, 2001; Sattler, 2001).

Testvaliditet dreier seg om to sentrale spørsmål: hva måler testen og hvor godt måler den det?

Standards for Educational and Psychological Testing, som er et retningsgivende referansearbeid om testvaliditet og andre sider ved testing, understreker at testskårer ikke har eller mangler validitet i seg selv, men at det er *tolkningen* av testene som er eller ikke er valid. Tolkningen er en slags påstand som vi gjør om testskårene (AERA, 1999).

Metodene som brukes for å vise at en test sannsynligvis er valid for et generelt formål, er blitt flere og mer avanserte i årenes løp (McCauley, 2001). AERA (1999) opererer med fem ulike kilder til validitet som kan brukes til å tolke testskårer til forskjellige formål. Vi har valgt å holde oss til en inndeling som brukes i mye av den øvrige litteraturen om validitet: innholdsvaliditet, kriterievaliditet og begrepsvaliditet (Befring, 2007; Kleven, 2007a; Sattler, 2001). Disse validitetstypene behandles ofte som sideordnede, men begrepsvaliditet kan anses som en overordnet form for validitet i forbindelse med måling. Innholdsvaliditet og kriterievaliditet kan anses som aspekter ved begrepsvaliditet, selv om det kan være disse underaspektene som er av størst interesse (Kleven, 2002a). Det er de to sistnevnte som er mest sentrale i vår undersøkelse, da vi anser spørsmål om innholdsvaliditet for å være viktigere å utrede ved utvikling av tester. Vi vil allikevel nevne alle tre, da de er tett sammenvevde.

a) *Innholdsvaliditet* gjelder om de interne komponentene i en test representerer området testen er ment å måle. Når vi evaluerer innholdsvaliditeten, må vi vurdere egnetheten til testens bestanddeler (Sattler, 2001). Spørsmål som er relevante i den forbindelse, er om testen måler det aktuelle området, om testspørsmålene er adekvate, om testen inneholder nok informasjon til å dekke det den skal måle på en god måte og hvilket mestringsnivå innholdet vurderes på (Sattler, 2001).

b) *Kriterievaliditet* gjelder forholdet mellom testskårer og en type kriterium eller resultat, slik som rangeringer, klassifiseringer eller andre testskårer (Sattler, 2001). Kriteriet vi bruker for å finne ut om TROG og BPVS gir en valid identifisering av spesifikke språkvansker, er 20 spørsmål om språkferdigheter. Som mål på kriterievaliditet i denne undersøkelsen brukes beregning av testenenes sensitivitet og spesifisitet. Siden dette er en studie som undersøker sammenhenger mellom variabler, er kriterievaliditet helt sentralt. Kriteriet må i likhet med testen ha adekvate psykometriske egenskaper: Det bør la seg måle, være upartisk/nøytralt og relevant for testens formål (Sattler, 2001). Det finnes to typer kriterierelatert validitet: prediktiv og samtidig (Kleven, 2002a). Vi konsentrerer oss om samtidig validitet fordi prediktiv validitet ikke er relevant i vår sammenheng.

*Samtidig validitet* gjelder om testskårene er relatert til et kriteriemål som er tilgjengelig på det aktuelle tidspunktet (Sattler, 2001). I forbindelse med vår undersøkelse antar vi at korrelasjonen mellom lærernes oppfatning av barnas språkferdigheter og barnas prestasjon på testene er et uttrykk for samtidig validitet.

Kriterier har ikke nødvendigvis fullkommen reliabilitet og validitet. En lærervurdering kan ha dårligere reliabilitet enn testen, og innholds- og begrepsvaliditeten til kriteriet kan være usikkert (Kleven, 2002a). Da vårt kriterium nettopp er lærervurdering, er dette noe vi må være oppmerksom på i forbindelse med tolkning av data i undersøkelsen vår.

*c) Begrepsvaliditet* er definert som ”grad av samsvar mellom begrepet slik det er definert teoretisk og begrepet slik vi lykkes med å operasjonalisere det” (Kleven, 2002a, p. 176). Vi kan skille mellom teoretisk og empirisk validitet. Vurdering av teoretisk validitet går ut på å evaluere validiteten til begrepene som blir benyttet (faglig konsensus i oppfatningen av begrepet). Vi vet at det ikke er en fullstendig faglig konsensus om begrepet spesifikke språkvansker, og det operasjonaliseres noe ulikt fra en undersøkelse til en annen. Vurdering av teoretisk validitet går også ut på å evaluere samsvaret mellom formell og operasjonalisert begrepsdefinisjon (Befring, 2007). Operasjonaliseringen av begrepet SSV er et viktig spørsmål i vår undersøkelse. Definisjonene er ofte for uklare til å gi klare svar på spørsmål om teoretisk validitet. Det må derfor foretas en empirisk vurdering i tillegg (Kleven, 2002a). Den empiriske vurderingen av begrepsvaliditet går ut på å vurdere om de relasjonene vi kan observere, stemmer med relasjoner vi teoretisk ville forvente (Kleven, 2002a). Dette kan vi til en viss grad få informasjon om via måleinstrumentet. Alpha-koeffisienten kan si noe om hvorvidt begrepsvaliditeten er akseptabel. I tilfeller der begrepet er sammensatt, kan måleinstrumentet inneholde flere delområder, og da kan vi forvente en korrelasjon mellom disse. De viktigste empiriske indikasjonene på begrepsvaliditet får vi ved å gjøre en ytre vurdering. Det vil si at måleresultatet sammenlignes med målinger i andre situasjoner og målinger av andre begreper. Konvergerende validitet vil si at man undersøker om resultatet korrelerer med andre målinger som det ut fra teorien kan tenkes å korrelere med. Diskriminerende validitet undersøkes ved å se om vårt måleresultat måler andre begrep som det ikke forventes å ha sammenheng med (Kleven, 2002a).

Verken teoretisk eller empirisk vurdering kan gi tilstrekkelig sikkerhet for konklusjonene. Vi baserer vurderingen av begrepsvaliditeten på antatt teoretisk kunnskap, og dette brukes så til å se om våre målinger bekrefter våre teorier. Vi må ta hensyn til denne usikkerheten og drøfte den (Kleven, 2002a).

Vi har sett at testens utgiver er ansvarlig for å dokumentere at en test er valid for et bestemt formål. I de neste avsnittene vil vi se på hvilke indikasjoner testmanualene gir om validiteten til TROG og BPVS.

## **Validiteten og reliabiliteten til TROG-2**

**Kriterievaliditet:** Det er ikke foretatt korrelasjonsstudier med andre reseptive grammatiske tester i forbindelse med den norske standardiseringen, men det ble beregnet korrelasjoner mellom dataene fra TROG og BPVS. Korrelasjonene varierer litt med alder, men er signifikante i alle aldersgrupper. De sterkeste korrelasjonene finner vi i 4-5-årsgruppen, og styrken avtar gradvis med stigende alder. Dette forklares med at den grammatiske kompetansen som måles med TROG, flater ut i de øvre aldersgruppene, mens ordforrådet derimot er i stadig utvikling (Bishop, 2009).

For den engelske versjonen av TROG ble diskriminerende validitet vurdert ved å sammenligne TROG-skårer for barn som fikk logopedisk opplæring med skårene til kontrollbarn på samme alder. Foreldre fylte også ut Childrens's Communication Checklist - Second Edition for en del av disse barna. Konklusjonen ble at TROG er sensitiv overfor barns kommunikasjonsvansker (Bishop, 2009).

**Reliabilitet:** I testmanualen gis det dokumentasjon for at denne testen har svært høy reliabilitet (Bishop, 2009). Sannsynligheten for at en person svarer riktig på alle de fire oppgavene i en blokk ved en tilfeldighet, er .004 (Bishop, 2003). En blokk består av fire oppgaver med samme grammatiske konstruksjon.

## **Validiteten og reliabiliteten til engelsk versjon av BPVS-II**

**Innholdsvaliditet:** Dersom BPVS anses som et mål på reseptivt vokabular (for standard engelsk), kan validiteten anses å komme fra innholdet. Stimulusordene kan betraktes som et utvalg fra det samlede engelske ordforrådet, og har en bredde på 18 kategorier. I manualen står det at BPVS følger en tradisjon med måling av reseptivt vokabular som har solid evidens for validitet.

**Kriterievaliditet:** Det er dokumentert god kriterievaliditet for PPVT, men det finnes få studier som gir direkte statistisk evidens for samtidig eller prediktiv validitet for BPVS. Korrelasjonen mellom den originale BPVS og TROG er 0.59 (Bishop, 1989, ref. i Dunn, et al., 1997). McCormick et al. (1994, ref. i Dunn et al., 1997) fant i en studie at PPVT-R var den beste prediktoren for kognitive ferdigheter hos små barn. Det å måle reseptivt vokabular slik det gjøres i BPVS, kan ikke likestilles med å forklare ord ekspressivt, men begge deler tapper barnets forståelse av det talte ord. Det er bare uttrykksmåten som er ulik.

**Reliabilitet:** Reliabiliteten for engelsk versjon av BPVS er svært høy. Test-retest-reliabiliteten er litt lavere enn for indre konsistens (Dunn, Dunn, Whetton, & Burley, 1997).

## 2.4 Forskning på testers sensitivitet og spesifisitet

Standardiserte språktester brukes i stor grad til å identifisere barn med spesifikke språkvansker, både klinisk og i forskning. Det er sjelden blitt stilt spørsmål ved i hvilken grad disse er valide og effektive verktøy for identifisering av klinisk signifikante språkvansker. Flere studier har vist at det er et misforhold mellom tester og klinikernes vurderinger (Aram, et al., 1992; Tomblin et al., 1997). Ifølge Bishop & McDonald (2009) er det allment kjent at det kan være et ganske stort misforhold mellom hvilke barn med SSV som identifiseres ved hjelp av språktester, og hvilke barn som identifiseres klinisk. De to forskerne mener at dette hovedsakelig kan skyldes to forhold: For det første har språktester varierende sensitivitet overfor klinisk viktige språkproblemer, for det andre fant de at barn av lav sosial status ikke like lett ble fanget opp av hjelpeapparatet, men at språktestene fanget opp at de hadde vansker.

En studie som ble foretatt av Gray, Plante, Vance og Henrichsen (1999) av den diagnostiske nøyaktigheten til fire vokabulartester, konkluderte med at det er viktig at statistisk evidens for validitet stemmer overens med det som er formålet med kartleggingen. Dokumentasjon for begrepsvaliditet er sentralt hvis hensikten med en test er å beskrive språkferdigheter, men er ikke tilstrekkelig i seg selv til å støtte identifiseringsformål. I studien hadde vokabulartestene evidens for samtidig validitet og diskriminerende validitet, men ingen av testene var godt egnet til å identifisere barn med SSV i utvalget. Selv om barna med SSV som gruppe hadde lavere skårer enn normalspråklige barn på fire ulike vokabulartester, falt skårene innenfor det som vanligvis regnes for å være normalområdet.

### Hvor nøyaktige er testene?

Dette spørsmålet er knyttet til prediktiv verdi, som igjen er forbundet med prediktiv validitet (Sattler, 2001). Prediktiv verdi er et begrep som brukes ved beskrivelse av diagnostiske tester eller screeningmetoder om sannsynligheten for at en person med utslag på testen (positiv test) virkelig har den undersøkte risikofaktoren (positiv prediktiv verdi), eller om sannsynligheten for at en person uten utslag på testen virkelig er uten den undersøkte risikofaktoren (negativ



prediktiv verdi) (Braut, 2009). For å finne den prediktive verdien måles testene mot et kriteriemål som vanligvis baseres på personens prestasjon, på diagnoser stilt av helsepersonell eller lærervurderinger (Sattler, 2001).

Figuren under er en modell for vurdering av den prediktive verdien til en test.

**Figur 1: Mulige utfall av en test, målt mot et kriteriemål**

		Dårlig utfall på testen	Godt utfall på testen
Kriteriemål (20 spørsmål)	I risikozonen	<b>a) sann positiv (treff)</b>	<b>b) falsk positiv (falsk alarm)</b>
	Ikke i risikozonen	<b>c) falsk negativ (bom)</b>	<b>d) sann negativ (korrekt)</b>

Basert på Sattler (2001)

- a) Sann positiv vil si at testen plasserer barnet i risikogruppen, og at dette samsvarer med kriteriemålet. (Positiv betyr her at barnet er i risikozonen for vansker eller dårlig resultat.)
- b) Falsk positiv vil si at testen plasserer barnet i risikogruppen, men at dette ikke samsvarer med kriteriemålet. Dette kalles overrapportering.
- c) Falsk negativ vil si at testen plasserer barnet utenfor risikogruppen, men at dette ikke bekreftes av kriteriemålet. Dette kalles underrapportering. (Negativ betyr her at barnet klassifiseres utenfor risikozonen for å ha problemer eller dårlig resultat.)
- d) Sann negativ vil si at testen plasserer barnet utenfor risikozonen for vansker, og at dette bekreftes av kriteriemålet (Rydz, et al., 2005).

Hvor nøyaktig en test måler det den skal måle, angis vanligvis som testenes sensitivitet og spesifisitet. Med sensitivitet menes evnen til å identifisere alle individer med en vanske.

Spesifisitet er evnen til korrekt å identifisere alle individer uten vansker (Rytz et al, 2005). I vårt tilfelle er vansken spesifikke språkvansker.

Følgende studier er eksempler på forskning som dokumenterer at tester ikke alltid har fullkommen sensitivitet og spesifisitet:

**Eksempler på mulig overrapportering (falskt positivt utfall):** I en studie fant Tomblin et al. (1997) at bare 29 % av barna som tilfredsstilte SSV-kriteriene på et testbatteri tidligere hadde vært identifisert som barn med spesifikke språkvansker. I en annen studie ble over 45 % av et tilfeldig utvalg av åtteåringer i en kontrollgruppe identifisert som barn med SSV da det ble benyttet psykometrisk testing (Aram, et al., 1992). I en tredje studie ble det rapportert at språktesting avdekker en betydelig andel av hittil ukjente språkvansker (Cohen, Davine, & Kelly-Meloche, 1989). De nevnte studiene kan være eksempler på overrapportering. Årsaken til denne overrapporteringen kan være at språkvansken forblir uidentifisert på grunn av manglende kunnskap om denne typen vansker, som igjen kan være en følge av at det ikke er tilgjengelig ekspertise på området. En annen mulighet er at disse tilfellene er falske positive tilfeller, det vil si eksempler på underrapportering (se kvadrant b), som oppstår fordi språktestene som ble benyttet, ikke hadde tilstrekkelig reliabilitet. Selv om språktestene skulle være tilstrekkelig reliable, kan de mangle validitet hvis barn gjør det dårlig på språktestene, uten at lærere eller foreldre oppfatter at barnet har en vanske (Bishop & McDonald, 2009).

**Eksempler på mulig underrapportering (falskt negativt utfall):** Andre studier har funnet tilfeller av barn som er fanget opp av hjelpeapparatet på grunn av at de har språkvansker, men ikke innfrir psykometriske kriterier for språkvansker. Klassifiseringen testene gjør av disse barna, kan være falsk negativ (kvadrant c). Et eksempel som illustrerer dette er en studie som viste at 35 % av førskolebarna som allerede var henvist til logoped hadde adekvat språkutvikling ifølge tester (Keegstra, Knijff, Post, & Goorhuis-Brouwer, 2007). Mange av disse barna hadde høyt utdannede foreldre, og forskerne konkluderte med at foreldre med høy utdanning var mer tilbøyelige til å være overbekymret for språkutviklingen. Dersom foreldrene ikke hadde rett i at barna hadde en språkvanske, kan testens angivelse være sann negativ, det vil si at testene underrapporterer vansken (se kvadrant d). En annen mulighet er at barna kan ha hatt språkvansker som hadde løst seg på det tidspunktet da testen ble tatt, eller at de hadde vansker (f.eks. stemmevansker) som ikke involverte språkvansken. Conti-Ramsden og Botting (1999b) gav enda en forklaring. Det kan være at standardiserte tester ikke fanger opp signifikante språkproblemer.

Målet bør være å fastsette diagnostiske kriterier som reduserer antall barn som hører til i kvadrant b og c (figur 1), og at man øker antall barn som hører til i kvadrant a og d. Det vil si at man øker sensitiviteten og spesifisiteten til den diagnostiske prosessen (Bishop & McDonald, 2009).

## **Oppsummering av kapittel 2**

Begrepet spesifikke språkvansker refererer til en betydelig svekkelse av språkferdigheter uten at faktorer som vanligvis ledsager vansker med språklæring, slik som hørselstap, lav nonverbal IQ og nevrologisk skade er fremtredende. For disse barna er språkvansken det primære problemet. Det anslås at ca. 5-7 % av alle barn har spesifikke språkvansker (Leonard, 2000; Ottem & Lian, 2008a).

Barn med spesifikke språkvansker er en heterogen gruppe og har ulike lingvistiske kjennetegn. Selv om det finnes noen typiske trekk, kan vansken arte seg forskjellig fra individ til individ. Forståelsen av vansken har endret seg gjennom tidene og barna har fått mange merkelapper i årenes løp. Fortsatt er forskerne uenige om hvordan spesifikke språkvansker skal defineres og hva som er årsakene til SSV. De fleste enes om at SSV har en arvelig komponent, men dette er ikke hele forklaringen. Ulike kognitive mekanismer kan lede til spesifikke språkvansker hos barn.

Man har gått bort fra jakten på den ene årsaken, og innsett at det trolig er flere underliggende årsaker til vansken. Det er gjort flere forsøk på å finne valide undergrupper av SSV, men dette har vist seg å være vanskelig. En forklaring på det kan være at vansken endrer karakter med alderen, slik at et barn kan gå fra å tilhøre en gruppe til en annen (Bishop, 1997).

En språk- og kommunikasjonsvanske kan få store konsekvenser i et barns liv, fordi språkvansken ofte vedvarer over lang tid og barnet står i fare for å utvikle adferdsmessige, sosiale, skolerelaterte og psykiske problemer (Hughes, et al., 2009; Ottem & Lian, 2008a).

Spesifikke språkvansker er et skjult problem, og det har vist seg å være vanskelig å identifisere barn med spesifikke språkvansker på en valid måte ved hjelp av standardiserte tester alene (Bishop, 1997). Flere studier har vist at tester mangler sensitivitet overfor klinisk viktige tegn på språkvansker. Problemet er ikke hva testene måler, men hva de ikke måler (Haywood, 1997). Tester favner ikke barnets daglige språk i samhandling med andre (Rygvoid, 2008). Etter hvert har man også begynt å benytte en mer økologisk tilnærming til

kartlegging der barnet vurderes i en naturlig kontekst og ikke behøver å prestere noe. Spørreskjemaer som fylles ut av noen som kjenner barnet godt, kan brukes til et slikt formål.

Testvaliditet dreier seg om i hvilken grad en test måler det den skal måle. Det finnes mange aspekter ved testvaliditet, men to spørsmål er sentrale: Det første gjelder om det kan trekkes slutninger på grunnlag av testresultatene, dvs. hvor nyttige eller meningsfulle de er (innholds- og begrepsvaliditet). Det andre gjelder hvilket forhold det er mellom testskårer eller andre kriterier.

Formålet med denne undersøkelsen er å undersøke om de to språktestene TROG og BPVS identifiserer barn med SSV på en måte som er valid i forhold til læreres beskrivelse av vanskene. Vi har derfor presentert generell teori om testvaliditet, testers sensitivitet og spesifisitet og hva forskning sier om testers validitet. Vi har også sett på hva testmanualene sier om validiteten til TROG og BPVS.

## 3 METODISK TILNÆRMING

Vår undersøkelse gjennomføres som et delprosjekt i en større studie som forskergruppen Kunnskapsgenerering i det spesialpedagogiske praksisfeltet (KiSP) gjennomfører ved to skoler i Oslo. Hovedprosjektet har et eksperimentelt design der det blant annet blir foretatt pre- og posttestmålinger av elevers vokabular for å se effekten av en intervensjon. I tillegg til en rekke språk-, stave- og lesetester som er gitt individuelt og i gruppe, har elevene gjennomført en test som måler nonverbale kognitive ferdigheter og fylt ut et skjema der de vurderer sine egne læringsprosesser. Lærerne har for øvrig fylt ut skjemaet 20 spørsmål om språkferdigheter for hver elev.

### 3.1 Vår undersøkelse

Vi har vært knyttet til hovedprosjektet i pretestfasen. Vår oppgave har vært å delta i testing av elever og registrering av data. Vi har til gjengjeld fått tilgang til alle data som er blitt samlet inn, men har kun benyttet oss av data som er relevante for vår problemstilling. Ut i fra problemstillingen vår er det data fra måleinstrumentene TROG, BPVS og 20 spørsmål i tillegg til den nonverbale evnetesten Raven vi skal analysere, men vi vil også benytte tilleggsdata fra Språk 6-16 for å belyse elevenes vansker.

Vi vil i det følgende beskrive ulike aspekter ved vår undersøkelse. Vårt prosjekt har et kvantitativt design. I en undersøkelse av denne typen er det viktig å beskrive utvalg, design, metode, gjennomføring, instrumenter og analyse på en grundig måte, slik at den er etterprøvbart, og det vil vi gjøre i det følgende. Alle undersøkelser har sine begrensninger, derfor vil vi også presentere aspekter ved validitet og reliabilitet som er relevante for vår undersøkelse. Avslutningsvis beskriver vi etiske aspekter ved prosjektet for å vise at disse er ivare tatt, og oppsummerer kapittelet.

#### 3.1.1 Utvalg

Utvalget i undersøkelsen vår er hentet fra utvalget i KiSP-prosjektet. KiSP-prosjektet har et utvalg på 398 barn fra 1., 2., 5., 6. og 8. trinn ved to skoler i Oslo. Utvalget i vår undersøkelse består av 80 barn på 2., 5. og 6. trinn. 47 av disse barna er jenter og 33 er gutter. Vi har

inkludert barn med ulik alder i utvalget både for å få et større antall og fordi alder ikke har betydning for den analysen vi skal gjøre.

Vi har brukt følgende seleksjonskriterier:

- norsk morsmål
- skåre på minimum -1 SD på den nonverbale testen Raven (som tilsvarer en IQ på 85 eller høyere)
- ingen kjent hørselsvanske

Utvalget består for øvrig bare av elever som er testet eller vurdert med samtlige av måleinstrumentene som benyttes i undersøkelsen.

De ovennevnte seleksjonskriteriene er valgt med utgangspunkt i et sett med kriterier for identifisering av barn med SSV som er anbefalt av Leonard (2000) (Se avsnitt 2.2.1 om inklusjons- og eksklusjonskriterier). Vi benytter IQ-kriteriet som er anbefalt av Leonard, og ekskluderer barn med kjente hørselsvansker. Da vi av hensyn til personvernet ikke har kunnet innhente helseopplysninger, kan det ikke utelukkes at det er barn med andre tilleggsvansker i utvalget. Vi har ikke brukt språkskåre som seleksjonskriterium for utvalget, men tar utgangspunkt i svake skårer på måleinstrumentene når vi skal finne ut hvilke barn som kan ha spesifikke språkvansker. I utvalget inkluderer vi alle barn uansett språkskåre, fordi vi ønsker en fullstendig oversikt over hvordan barnas prestasjoner fordeler seg på de ulike måleinstrumentene, og fordi vi i tillegg til språktester også benytter lærervurdering i identifiseringen av barna med SSV.

Som det fremgår innledningsvis, har vi få informanter sammenlignet med hovedprosjektet. Årsaken til dette er for det første at TROG og Raven ikke er administrert til alle barna som er tilknyttet hovedprosjektet, fordi testene ikke har egnet seg for barn på alle alderstrinn. Konsekvensen av dette er at vi har måttet ekskludere alle barn på 1. og 8. trinn. For det andre har de ovennevnte seleksjonskriteriene ført til at vi har måttet utelate minoritetsspråklige barn, og dette utgjør en stor gruppe (110 barn).

### 3.1.2 Design og metode

Vi har samlet inn data til undersøkelsen ved bruk av tester og lærevurderinger. Dette gir oss kvantitativ informasjon om hvor mange barn som kan ha SSV, men gir oss også tilgang til kvalitative opplysninger om elevenes vansker. Ut i fra problemstillingen vår skal vi finne sammenhengen mellom tre variabler (BPVS, TROG og 20 spørsmål om språkferdigheter). Problemstillingen skal belyses med en empirisk-kvantitativ metodisk tilnærming.

Undersøkelsen har et ikke-eksperimentelt design, og vil være deskriptiv i sin form. I en slik undersøkelse er ikke formålet å gi noen påvirkning som skal endre tingenes tilstand, men heller å studere tingenes tilstand slik de er (Kleven, 2002b).

I en deskriptiv studie tar man ofte sikte på å studere korrelasjoner mellom variabler og variabelgrupper (Gall, Gall, & Borg, 2007) i motsetning til eksperimentelle eller kvasi-eksperimentelle design hvor forskeren vil manipulere en eller flere uavhengige variabler (Kleven, 2002b). Oppgavens design og utvalgets fordeling får konsekvenser for hvilke analyser vi kan bruke. Den mest brukte korrelasjonskoeffisienten er Pearsons  $r$  som blant annet krever at variablene er kontinuerlige og normalfordelte, og at det finnes et lineært forhold mellom disse variablene (Sattler, 2001). Med unntak av 20 spørsmål innfrir alle måleinstrumentene vi benytter disse forutsetningene. 20 spørsmål er ikke konstruert for å gi normalfordelinger, det vil si at man ikke er interessert i å differensiere mellom godt fungerende barn, men heller å få en vurdering av de svaktfungerende. Vi foretar korrelasjoner mellom BPVS og TROG for å kontrollere i hvilken grad disse testene måler det samme underliggende begrepet. Vi foretar også korrelasjoner mellom de ovennevnte testene og Raven, som brukes til å undersøke om skårene på språktestene påvirkes av IQ.

Hovedformålet til prosjektet vårt er imidlertid å undersøke sammenhengen mellom BPVS og TROG på den ene siden og 20 spørsmål om språkferdigheter på den andre. Dette handler om kriterievaliditet. Da 20 spørsmål om språkferdigheter ikke er normalfordelt, kan vi ikke foreta en korrelasjon mellom skjemaet og de to språktestene. I stedet gir vi oversikter over sammenhengen mellom skjemaet og testene ved hjelp av krysstabeller ved å gruppere skårene, og beregner testenes nøyaktighet i forhold til kriteriet ved å regne ut testenes spesifisitet og sensitivitet. Testenes evne til å skille mellom barn med og uten vansker, undersøkes også ved hjelp av en t-test for uavhengige utvalg. Sammenhengen mellom barnas skårer på samtlige måleinstrumenter i undersøkelsen belyses for øvrig ved hjelp av en enkel

deskriptiv analyse som siden brukes som grunnlag for en kvalitativ drøfting av utvalgte case. Beskrivelsen av case danner også grunnlag for drøfting av forskjeller og likhetstrekk mellom de ulike måleinstrumentene.

### **3.1.3 Gjennomføring**

Før testingen av barn ved de to skolene tok til, fikk alle masterstudentene som var tilknyttet KiSP-prosjektet, en innføring i administrering av måleinstrumentene som skulle benyttes. Dette ble gjort for å sikre at testingen skulle gjennomføres på en mest mulig konsekvent og nøyaktig måte. Testperioden strakte seg fra august 2009 til januar 2010, og testene ble i hovedsak administrert av 10 masterstudenter fra ISP og forskningsassistenter fra KiSP-prosjektet. Formalitetene rundt innhenting av samtykke og avklaring av egnet tid og sted med skolene ble ordnet av ledelsen ved KiSP-prosjektet.

Testingen med BPVS, TROG og Språk 6-16 ble gjennomført på skolen i skoletiden. Språk 6-16 ble bare administrert til barn som lærerne vurderte som språklig svake gjennom spørreskjemaet 20 spørsmål. Studentene testet ca. 30 elever hver. Det var opprinnelig beregnet at BPVS og TROG skulle administreres samtidig, og at hver elev skulle bruke 20 minutter på de to testene. Det viste seg imidlertid at særlig de yngste elevene trengte mer tid, og testingen tok rundt én time per barn. Årsaken til tidsbruken var at vi var opptatt av at elevene slappet av i testsituasjonen, og brukte derfor tid på å bli litt kjent med dem før vi begynte. Det ble også gitt pauser etter behov, slik at elevene kunne yte maksimalt.

Barna ble også testet med Raven, som er en nonverbal evnetest. Testen ble administrert som en gruppetest i klassene ved elevenes lærere. Utover de ovennevnte måleinstrumentene ble det gitt en rekke andre språkrelaterte tester. I tillegg til testene ble det innhentet informasjon om elevenes språkferdigheter ved hjelp av spørreskjemaet 20 spørsmål, som hovedlærerne fylte ut for hver elev. Dette skjedde i testperioden, og lærerne hadde ingen kjennskap til testresultater.

Det er viktig at informantene i undersøkelsen ikke kan identifiseres. Vi har derfor gitt elevene case-nr. som ikke kan spores tilbake til enkeltpersoner. Av hensyn til anonymiseringen angis ikke klassetrinn i forbindelse med casebeskrivelser. I stedet brukes benevnelsene ”småskoletrinnet” og ”mellomtrinnet”.



I det følgende gis det en nærmere beskrivelse av måleinstrumentene som benyttes i undersøkelsen.

### **3.1.4 Måleinstrumenter**

Nedenfor gis en beskrivelse av hvert av måleinstrumentene vi har brukt i undersøkelsen: TROG, BPVS, 20 spørsmål om språkferdigheter og Raven. Vi beskriver også Språk 6-16, som benyttes som supplerende måleinstrument i undersøkelsen.

#### **Test for Reception of Grammar (TROG 2)**

TROG 2 (Bishop, 2009) er norsk versjon av en britisk test som opprinnelig ble utviklet av Dorothy Bishop på slutten av 1970-tallet. Testen er oversatt og standardisert til norske forhold, og er beregnet for barn fra 4-16 år. TROG er en reseptiv språktest for kartlegging av grammatisk forståelse som kan gi et godt grunnlag både for å finne en persons funksjonsnivå og for å kunne diagnostisere spesifikke grammatiske vansker. I den britiske versjonen av manualen vises det til studier der den opprinnelige versjonen av TROG har vært brukt. Ifølge disse studiene har testen vist seg å være et godt redskap ved utredning av blant annet spesifikke språkvansker (Bishop, 2009). I den norske manualen oppgis det at testen ikke bør brukes alene i diagnostiseringen av språkvansker, men kan inngå som en del av et testbatteri.

Testen er individuelt administrert og består av 80 flervalgsoppgaver. Oppgavene er organisert i blokker med fire oppgaver for hvert grammatiske forhold som undersøkes. (En blokk er en betegnelse for en gruppe med oppgaver som tester barnets forståelse av en bestemt grammatisk konstruksjon.) Hver oppgave har fire svaralternativ som presenteres i form av bilder. Testleder skal vise en oppgave til barnet og si: ”Her har vi en bok med fire bilder på hver side. Nå skal jeg si noe, og så skal du peke på det bildet som passer best til det jeg sier.” Barnet skal så peke på det som passer med setningen testleder leser fra en stimulusbok, og behøver altså ikke å bruke språk for å avgi svar. For hver oppgave er det ett riktig svaralternativ. De tre andre svaralternativene betegnes som distraktorer, og skiller seg fra korrekt svar kun ved at man har endret et grammatisk eller leksikalt element. Eksempel: I oppgave K1 ”Kua blir jaget av jenta” representerer svaralternativ 1) ”Jenta blir jaget av kua” en grammatisk distraktor. Ordrekkefølgen er svært viktig for forståelsen, og distraktorene er konstruert på en slik måte at man kan oppdage feilmønstre hos testpersonen. Vanskegraden

er økende, og testingen avsluttes når 5 blokker etter hverandre er galt besvart. Dersom barnet svarer galt på en oppgave innen en blokk, regnes hele blokken som feil.

## **British Picture Vocabulary Scale II (BPVS II)**

British Picture Vocabulary Scale er en britisk test som er utviklet av Dunn, Dunn, Whetton og Burley (1997). Det er en individuelt administrert test som kartlegger reseptivt vokabular for aldersgruppen 3-15 år. Ifølge Dale & Reichert (1957) referert i Dunn & Dunn, Whetton & Burley (1997) kan testen predikere skolefaglig fungering. Testen er ikke oversatt og standardisert til norske forhold, så vi har benyttet den britiske versjonen av testen.

Testen består av 14 oppgavesett med 12 ord i hvert sett. Hver oppgave har fire svaralternativ som presenteres i form av bilder. Barnet skal velge det bildet som passer til stimulusordet som testleder leser opp. Testlederen sier: ”Jeg skal si et ord, og så vil jeg at du skal peke på det bildet som passer best til det jeg sier”. Man starter med det oppgavesettet som tilsvarer barnets alder (evt. forventet alder). Har barnet mer enn én feil i første oppgavesett, må tidligere oppgaver administreres for å finne barnets nedre grense. Dersom barnet har åtte eller flere feil i et oppgavesett, avsluttes testingen. Dette er barnets øvre grense. Testen går fra høy- til lavfrekvente ord, og fra konkrete til abstrakte ord og begreper. Vanskegraden er økende, og verbal respons er ikke nødvendig.

## **20 spørsmål om språkferdigheter**

20 spørsmål om språkferdigheter er et måleinstrument som brukes som det kriteriet testene måles opp mot, i undersøkelsen vår. Måleinstrumentet er et skjema for kartlegging av språkvansker hos barn og ungdom. Selv om skjemaet kalles ”20 spørsmål om språkferdigheter”, er det formulert som utsagn om barns språkferdigheter og språkrelaterte vansker. I tillegg til de 20 spørsmålene inneholder skjemaet også en del med supplerende spørsmål om tilegnelse av symbolsystemer og relasjoner til andre, men denne delen benyttes ikke i vår undersøkelse. Skjemaet er ment som et supplement til bruk av tester ved utredning av elever med spesifikke språkvansker i aldersgruppen 5 til 16 år (Ottem, 2009).

I en artikkel om sammenhengen mellom skjemaet og Språk 6-16, kalles skjemaet et observasjonsskjema (Ottem, 2009). Slik det er utformet, er det ikke så egnet til bruk i en observasjonssituasjon, men snarere til å registrere et inntrykk læreren har dannet seg over tid

av hva som kjennetegner et barns språkrelaterte vansker i skolesituasjonen. Vi synes det har mye til felles med en sjekklister eller et spørreskjema, og bruker ofte den sistnevnte betegnelsen når vi omtaler skjemaet i denne oppgaven. Hensikten med skjemaet er å gi lærere og andre fagfolk et klarere bilde av elever med språk- og kommunikasjonsvansker. Skjemaet bygger på 20 utsagn om språkvansker, og disse vurderes i forhold til andre barn på samme alder. De samme spørsmålene kan stilles på mange alderstrinn, fordi svarene er relativt aldersuavhengige. Utsagnene er gruppert i tre delskalaer etter hva de måler (se vedlegg 1): semantisk evne (utsagn 1-8), vansker med å motta et budskap (utsagn 9-14), og vansker med å formidle et budskap (utsagn 15-20). Resultatene for delskalaene summeres til en total sumskåre. Hvert utsagn har en skala fra 1 til 5 der en krysser av for 1 dersom utsagnet ikke passer for eleven. Hvis utsagnet passer, krysses det av for 5. Passer utsagnene delvis, krysses det av for tall mellom 1 og 5. Maksimal sumskåre er  $5 \cdot 20 = 100$ . Hvis antall anmerkninger øker, øker også sannsynligheten for at barnet kan ha språkvansker.

I en undersøkelse ble det rapportert at en sumskåre over 31 forekommer relativt sjelden hos typiske barn. Sumskårer over 41 er svært sjeldne hos typiske barn. I undersøkelsen regnet Ottem 31 % av maksimal skåre som en kritisk grense for å undersøke om barnet kan ha språkvansker (Ottem, 2009). Vi har satt grensen på 51 % i vårt utvalg for å være rimelig sikre på at vi har med barn med alvorlige språkvansker å gjøre. I tabell 1 i avsnitt 3.1.7 i dette kapitlet gis det en oversikt over hvordan vi har delt inn barna etter skårer på 20 spørsmål og de andre måleinstrumentene. Inndelingen begrunnes i drøftingsdelen.

### **Raven Standard Progressive Matrices (SPM)**

Raven SPM (Raven, Court, & Raven, 1986) er en nonverbal test som er ment for bedømmelse av generelle evner. Testen er beregnet for aldersgruppen 6 til 17 år, men kan også administreres til voksne. Administreringen gjøres individuelt eller i gruppe. I vår undersøkelse er testen gitt som en gruppetest. Raven SPM består av 60 testledd med varierende vanskegrad og ulike krav til logisk tenkning. Hvert ledd består av en figur eller en matrise hvor et element mangler, og oppgaven går ut på å identifisere det manglende element blant fem alternativer. Barnet skal sammenlikne mønstre og mønsterkombinasjoner og foreta analogiske slutninger. Testen er en såkalt matrisetest som måler evnen til å kunne foreta en logisk utledning av nye relasjoner og mønstre på grunnlag av kjent informasjon. Den gjør det mulig å vurdere problemløsningskapasitet uten at det benyttes verbalt språk.

### 3.1.5 Supplerende måleinstrument

For å få større bredde i dokumentasjonen av språkferdighetene tar vi med data fra deltestene Begreper, Ordspenn og Setningsminne fra screeningtesten Språk 6-16. Det er bare elever som læreren regner som språklig svake, som har tatt deltestene til Språk 6-16. Nedenfor presenteres testen.

#### Språk 6-16

Språk 6-16 er en normert screeningtest for språkvansker hos barn og ungdom i alderen 6 til 16 år. Den er utarbeidet av Ottem og Frost, og ble publisert i 2005. Testen skal avdekke om et barn har språklige ferdigheter som er adekvate for alderen, og om det er nødvendig med videre utredning av språkvanskene. Testen består av tre obligatoriske deltester som tester ekspressivt språk: Ordspenn, Setningsminne og Begreper. Ordspenn går ut på å gjenta ord. Barnet skal først gjenta to ord, deretter tre ord, fire ord og fem ord (eksempel: Rusk-Hegg-Fikk). Deltesten Setningsminne går ut på at barnet skal gjenta setninger helt korrekt. Setningslengden øker gradvis (eksempel: Pærene i hagen min er bedre enn dem i butikken). Deltesten Begreper består av to delprøver. Den første kalles motsetninger og går ut på at barnet skal finne antonymer til det testlederen sier (Er kapteinen liten eller er han \_\_\_\_\_(stor?). Den andre delen tester barnets ordkunnskap. Barnet blir bedt om å forklare hva ulike ord betyr (for eksempel: hva er en paraply?). Ordene har økende vanskegrad.

Deltestene Ordspenn og Setningsminne undersøker ulike aspekter ved korttidsminnet. Deltesten Ordspenn regnes som en god test på fonologisk korttidsminne (det vil si å kunne fastholde lydstrukturene i ord). Deltesten Setningsminne regnes som en god test på evnen til å organisere og fastholde informasjon i setninger. Den tredje deltesten, Begreper, undersøker det semantiske aspektet ved språket, det vil si barns kunnskap om ordenes betydning. Den totale sumskåren beregnes på grunnlag av resultatene i de tre deltestene.

Testen har også supplerende deltester innen områdene grammatikk, fonologisk bevissthet og lesing. Disse deltestene er ikke administrert til vårt utvalg, og vil derfor ikke bli beskrevet nærmere.

For hver deltest måles prestasjonene med en skåre som har et gjennomsnitt på 10 og et standardavvik på 3. Skårer fra 7 til 13 regnes altså for å ligge innenfor normalvariasjonen.

### 3.1.6 Analyse av data

Etter gjennomføring av testene ble de skåret og resultatene registrert i SPSS (Statistical Package for the Social Sciences). Det er viktig at måleinstrumentene som benyttes i undersøkelsen har tilstrekkelig reliabilitet. Reliabilitetsanalyse gir et mål på om måleinstrumentene er operasjonalisert på en tilfredsstillende måte, og dermed om de oppfyller krav til indre konsistens (Christophersen, 2009; Kleven, 2007a). Som vi har vist i teoridelen, angir testmanualene at TROG og BPVS er reliable måleinstrumenter (Bishop, 2003, 2009; Dunn, et al., 1997). Da vi har dokumentasjon for testenes reliabilitet, foretar vi ingen reliabilitetsanalyser av disse. Da 20 spørsmål er et forholdsvis nytt måleinstrument, kontrollerer vi reliabiliteten til dette. Videre benyttes deskriptiv statistikk til å analysere sammenhengen mellom de ulike måleinstrumentene. Sammenhengen mellom BPVS, TROG og Raven analyseres med korrelasjonsanalyse, mens sammenhengen mellom testene og 20 spørsmål undersøkes ved hjelp av en krystabell, en enkel deskriptiv analyse og t-test for uavhengige utvalg samt ved beregning av testenes nøyaktighet.

#### Reliabilitetsanalyse

Reliabilitetsanalyse gir et mål på om måleinstrumentene er operasjonalisert på en tilfredsstillende måte, og dermed om måleinstrumentene oppfyller krav til indre konsistens (Christophersen, 2009; Kleven, 2007a).

Reliabilitetskoeffisienten som er brukt, er Cronbachs alpha, som varierer mellom 0 og 1. Verdier nær 1 indikerer høy indre konsistens (Christophersen, 2009). Ifølge Svartdal (2009) betyr en alpha-verdi større enn 0.70 at man kan anta at items har en tilstrekkelig høy indre konsistens.

#### Korrelasjonsanalyse

Korrelasjon betyr ”samgang eller samvariasjon mellom forskjellige variabler” (Befring, 2007, p. 147). Det vil si at verdien til en av variablene forteller noe om verdien til den andre variabelen (Befring, 2007). Samvariasjon behøver ikke innebære kausalitet eller årsak/virkning (Svartdal, 2004). En korrelasjonsanalyse bidrar til å gi konkrete svar på fremsatte hypoteser, og gjør det mulig å estimere ut i fra et utvalg ved å vurdere om det er stor sannsynlighet for at parameteren i populasjonen har en bestemt verdi (dvs. vurdere om de faktiske resultatene

samsvarer med de forventede resultatene). I en analyse er det viktig at man inkluderer en kontrollvariabel, slik at man ikke tillegger interessante forklaringsvariabler en betydning de egentlig ikke har (Midtbø, 2007). I vår oppgave er Raven brukt som en kontrollvariabel og seleksjonsvariabel. Når Raven brukes som en kontrollvariabel, er det for å avdekke om nonverbal IQ har noen sammenheng med skårer på språktester. Når Raven brukes som en seleksjonsvariabel, er det for å eliminere bort barn med skåre  $< -1$  SD, fordi vi vil studere barn med intelligens i normalområdet.

Bivariat korrelasjonsanalyse brukes til å korrelere to og to variabler for å kunne si noe om retningen og styrken på disse (Midtbø, 2007). I vår analyse velger vi å benytte Pearsons  $r$  som korrelasjonsmål fordi variablene i undersøkelsen er på et forholdstallsnivå, og fordi sammenhengen mellom variablene er tilstrekkelig lineær. Korrelasjonsmålet angir hvor sterk en observert sammenheng er og hvordan to variabler samvarierer. Styrken på samvariasjonen mellom to variablene indikeres med et tall mellom 0 og 1 eller 0 og -1. Verdiene +1 og -1 indikerer en perfekt sammenheng, mens 0 betyr at det ikke finnes noen sammenheng. En korrelasjon kan som vi ser være negativ eller positiv. En negativ korrelasjon vil si at høye verdier på den ene variabelen med stor sannsynlighet tilsier lave verdier på den andre variabelen. Omvendt vil en positiv korrelasjon tilsa at en høy verdi på en variabel ledsages av høye verdier på den andre variabelen (Befring, 2007).

Det vurderes ut i fra hver enkelt undersøkelse hva som er en sterk eller svak korrelasjon, fordi tolkning av korrelasjonskoeffisienter avhenger av hvilke variabler som korrelerer med hverandre (Svartdal, 2004). Ifølge McCauley (2001) er korrelasjoner  $< .20$  nesten ubetydelige, korrelasjoner mellom  $.20$ -. $40$  lave, korrelasjoner mellom  $.40$ -. $70$  moderate, korrelasjoner mellom  $.70$ -. $90$  høye og korrelasjoner  $> .90$  svært høye .

Korrelasjonskoeffisienter rapporteres vanligvis sammen med et utsagn om statistisk signifikans, som beskriver sannsynligheten for at korrelasjonskoeffisienten er ulik 0 ved en tilfældighet på grunnlag av utvalgets størrelse. Statistisk signifikans angir at det er lite sannsynlig at den oppnådde verdien er tilfeldig. En korrelasjons statistiske signifikans forstås noen ganger som den viktigste indikasjonen på dens betydning, men en svært lav korrelasjonskoeffisient er sannsynligvis ikke av betydning selv om den er statistisk signifikant, fordi den ikke forklarer mye av variansen til de korrelerte målene. Jo større utvalget er, jo enklere er det å oppnå statistisk signifikans. Derfor er det slik at selv om en

statistisk signifikant korrelasjonskoeffisient alltid er ønskelig, må styrken i tillegg til signifikansen til korrelasjonen vurderes (McCauley, 2001).

## **Krysstabell**

Bivariate krysstabeller benyttes for å sammenligne to og to måleinstrumenter med hverandre. Krysstabellene gir flere detaljer om sammenhengen enn det korrelasjonene gjør, fordi de gir en oversikt over hvor mange barn måleinstrumentene kategoriserer i de ulike gruppene.

Vi kan observere tendenser i krysstabellene, men for en nøyaktig angivelse av hvor godt testene skiller mellom barn med og uten mulig SSV, regner vi ut testenenes prediktive verdi. Disse beregningene beskrives i neste avsnitt.

## **Prediktiv verdi – testenenes sensitivitet og spesifisitet**

Hvor nøyaktig en test måler det den skal måle, angis vanligvis som testenenes sensitivitet og spesifisitet (se avsnitt 2.4). Formlene for sensitivitet og spesifisitet er vist nedenfor:

$$\text{Sensitivitet} = \frac{\text{Sann positiv}}{\text{Sann positiv} + \text{falsk negativ}} \quad \text{Spesifisitet} = \frac{\text{Sann negativ}}{\text{Sann negativ} + \text{falsk positiv}}$$

(Rydz, et al., 2005, p. 7)

Vanligvis er et sensitivitetsnivå på 70 % eller mer akseptabelt, men noen anbefaler at det skal være 80 % eller mer. For spesifisitet er verdier på 70-80 % akseptable, men noen eksperter anbefaler en spesifisitet på 90 % (Rydz, et al., 2005).

## **T-test for uavhengige utvalg**

T-test for uavhengige utvalg er en test som ofte brukes til å undersøke forskjellen mellom to grupper (Kleven, 2007b). De to gruppene består her av barn som er blitt identifisert som barn

med mulig SSV ifølge 20 spørsmål, og barn som ikke har SSV ifølge 20 spørsmål. Metoden sammenlikner to gjennomsnitt. Dersom utvalgene er meget små, forutsetter t-testen at populasjonen er normalfordelt (Kleven, 2007b). I denne undersøkelsen vil t-test bli brukt til å finne ut om barn som kan ha SSV ifølge 20 spørsmål, har svakere skårer på TROG og BPVS enn barn som ifølge 20 spørsmål har mulig SSV.

## **Enkel deskriptiv analyse**

Vi har valgt å benytte enkel deskriptiv analyse i vår undersøkelse fordi den gir en komplett og oversiktlig fremstilling av alle variablene som er benyttet. Dette gir oss et grunnlag for å sammenligne skårer på ulike tester for hvert enkelt barn (Vedlegg 2). Tabellen viser gruppeinndeling i 20 spørsmål, casenummer, kjønn og trinn samt skåre på TROG, BPVS, 20 spørsmål, deltestene Begreper, Ordspenn og Setningsminne (fra Språk 6-16) og Raven.

Det er mange muligheter for utforming av en enkel deskriptiv analyse. Vi har valgt et oppsett der vi deler data i grupper som representerer ulike språkferdighetsnivåer i samsvar med lærernes vurdering av barnas språklige fungering uttrykt ved skjemaet 20 spørsmål om språkferdigheter. Vi tar utgangspunkt i lærernes vurdering, fordi 20 spørsmål om språkferdigheter er det kriteriet TROG og BPVS måles opp mot. Dette gir oss en tredeling av resultatene: Gruppe 1 er barn som ser ut til å ha normale språkferdigheter ifølge lærerne. Gruppe 2 er barn som ser ut til å ha milde/moderate språkvansker og Gruppe 3 er barn som ser ut til å ha alvorlige språkvansker/spesifikke språkvansker. Nedenfor vil vi gi en forklaring på hva vi legger i normale språkferdigheter, milde/moderate språkvansker og alvorlige språkvansker/spesifikke språkvansker. Dette blir presentert i tabell 1 i avsnitt 3.1.7. Alle råskårene, med unntak av 20 spørsmål og Språk 6-16, er gjort om til standardskårer. Vi gir derfor først en forklaring på hva standardskårer er, og presenterer deretter tabellen.

## **Standardskårer**

I denne undersøkelsen har vi gjort om råskårer for flere måleinstrumenter til z-skårer. En z-skåre er en standardskåre. Overføringen av råskårer til z-skårer ble gjennomført for hvert alderstrinn, slik at det er mulig å vurdere skåringsnivå på tvers av alderstrinn. Z-skårene er altså aldersuavhengige. En av årsakene til bruken av z-skårer er at det for RAVEN og BPVS ikke foreligger oppdaterte normtabeller for norske barn.



Z-skåre er en direkte overføring av skåreverdier  $X$  til en ny skala, der standardavviket er måleenhet og Mean er nullpunkt. Standard-skåren  $z$  er et uttrykk for hvor mange standardavvik et skåreresultat  $X$  er plassert fra Mean i fordelingen. Det er en lineær overføring fra en skala til en annen, og det betyr at relasjonene mellom de ulike verdiene er de samme, og at  $z$ -skårene kan føres tilbake til den opprinnelige skalaen. Overføring til standard-skåre er en fordel både ved statistisk analyse og i forhold til funksjonsvurdering (Befring, 2007). Med funksjonsvurdering mener vi i denne sammenhengen om barna har en språkvanske eller er normalspråklige.

## **Operasjonalisering**

Begrepet SSV er et abstrakt begrep som ikke kan observeres direkte. De fleste psykologiske og pedagogiske variabler er abstraksjoner som bare kan observeres indirekte. Når vi skal måle en slik variabel, operasjonaliserer vi den ved å uttrykke den i form av variabelens observerbare trekk. Variabelen kan derfor betraktes som en hypotetisk egenskap. Når vi skal måle variabler av denne typen, tar vi utgangspunkt i observerbar adferd som vi regner med er forårsaket av eller har forbindelse med variabelen. Dette er et slags utsnitt av et individs adferd som kan kalles adferdsindikasjoner, fordi adferden indikerer noe om variabelen (Befring, 2007).

Vi operasjonaliserer begrepet SSV på følgende måte på måleinstrumentene i vår undersøkelse: En skåre på  $-1.25$  SD eller lavere på BPVS eller TROG kombinert med en skåre på  $-1$  SD eller høyere på den nonverbale testen Raven angir mulig SSV ifølge språktestene. Skåren på Raven tilsvarende en nonverbal IQ på 85 eller høyere. En skåre på 51 % av maksimal skåre eller høyere angir mulig SSV ifølge 20 spørsmål. Utover dette stilles det krav til at personene har norsk morsmål og ikke har kjente hørselsvansker.

### **3.1.7 Gruppering etter språkferdigheter**

Vi vil finne ut hvor mange av barna i utvalget vårt som har så svake skårer på måleinstrumentene at de kan ha SSV ifølge testene og/eller 20 spørsmål. For å gjøre fremstillingen oversiktlig har vi valgt å dele barna inn i ulike grupper etter deres språkferdigheter. Hver gruppe er beskrevet og gitt en språklig betegnelse (se tabell 1).

**Tabell 1: Oversikt over gruppering av barna på TROG, BPVS og 20 spørsmål**

Måleinstrumenter	Skåre	Betegnelse
TROG og BPVS	<b>Gruppe 1:</b> språkskårer >1 SD	Svært gode språkferdigheter
	<b>Gruppe 2:</b> språkskårer > -1.25 SD til 1 SD	Gode språkferdigheter/ normalspråklige
	<b>Gruppe 3:</b> språkskårer fra -1.25 SD eller lavere	Alvorlige språkvansker /spesifikke språkvansker (SSV)
20 spørsmål	<b>Gruppe 1:</b> Skårer fra 20 til 30 prosent av maksimal skåre	Gode språkferdigheter/ normalspråklige
	<b>Gruppe 2:</b> Skårer fra 31 til 50 prosent av maksimal skåre	Milde/moderate språkvansker
	<b>Gruppe 3:</b> Skårer fra 51 prosent av maksimal skåre	Alvorlige språkvansker/spesifikke språkvansker (SSV)

Tabellen viser at barna på grunnlag av skårer på TROG og BPVS og 20 spørsmål er fordelt i tre grupper. På TROG og BPVS består gruppe 1 av barn som skårer over 1 SD over gjennomsnittet. Disse barna har språklige ferdigheter over gjennomsnittet. Gruppe 2 er barn som skårer over -1,25 SD til 1 SD på en reseptiv språkttest: Dette er barn som befinner seg innenfor normalområdet for språklige ferdigheter. Gruppe 3 består av barn som skårer -1.25 SD eller lavere på testene. Dette er barn som kan mistenkes for å ha spesifikke språkvansker.

Bruk av z-skårer i forbindelse med 20 spørsmål og Språk 6-16 gir ingen mening, da utvalget er samlet fra den nederste delen av fordelingen. På grunnlag av skårene i 20 spørsmål har vi gjort følgende inndeling: Gruppe 1 består av barn som skårer fra 20 til 30 prosent av maksimal skåre. Disse barna regner vi som barn med normale språkferdigheter. Gruppe 2 består av barn med skåre fra 31 til 50 prosent av maksimal skåre. Dette regner vi som barn som tilsynelatende har problemer med språket, men som ikke har så alvorlige vansker som barna i gruppe 3. Gruppe 3 er barn med en skåre fra 51 prosent av maksimal skåre og høyere, og det er denne gruppen vi regner med at kan ha SSV.

Skårene som benyttes for Språk 6-16 er de skalerte skårene som er oppgitt i normtabellen for testen.

## 3.2 Validitet og reliabilitet i undersøkelsen

Vi vil nå se på noen kvalitetskrav som bør vurderes i undersøkelsen for å oppnå best mulig validitet. Validitet kommer av det latinske ordet *validitas*, som betyr sannhet, troverdighet, gyldighet og styrke (Kruuse, 2005). Spørsmålet om validitet reiser problemstillingen om vi virkelig har fått et måleresultat for variabelen vi ønsker å måle, og i hvilken grad resultatet også innbefatter irrelevante faktorer.

Alle undersøkelser har begrensninger når det gjelder utvalg, mål, intervensjon og setting. Validitet kan derfor bare bli gradvis oppfylt (Befring, 2007). I forbindelse med empiriske undersøkelser spiller validiteten en sentral rolle, fordi formålet med slike undersøkelser er å undersøke om en hypotese er rett eller gal. Cook og Campbell har utarbeidet et generelt validitetssystem for kausale undersøkelser, som også kan brukes for beskrivende undersøkelser av den typen vi har. Systemet omfatter fire typer av validitet: statistisk validitet, indre validitet, begrepsvaliditet og ytre validitet. Hver validitetstype har mulige feilfaktorer eller trusler som gjør det vanskelig å trekke valide slutninger (Lund, 2002b).

Skillet mellom de ulike validitetstypene er kunstig, fordi oppnåelse av en validitetstype har påvirkning på en annen. Et unntak fra dette er statistisk validitet som kan være tilfredstilt uten at de andre er det. I det følgende vil vi ta for oss de fire validitetstypene og mulige trusler som er knyttet til disse.

*God statistisk validitet* er en forutsetning for de andre validitetstypene. For å oppnå dette forutsettes det at sammenhengen eller tendensen er statistisk signifikant og rimelig sterk. Utvalgsstørrelse og design påvirker hvor god statistisk validitet som oppnås. God statistisk validitet er en nødvendig betingelse for indre validitet, ytre validitet og begrepsvaliditet. Lav statistisk styrke er en trussel mot god statistisk validitet som øker sannsynligheten for å gjøre Type I-feil eller Type II-feil, dvs. henholdsvis forkasting av en sann nullhypotese eller akseptering av en gal nullhypotese (Lund, 2002b).

*Indre validitet* angår slutninger om kausale relasjoner (Lund, 2002b). Indre validitet er i hovedsak ikke aktuelt som kvalitetskrav i beskrivende forskning, fordi man ikke kan si hva som er årsak og hva som er virkning. Dette kommer av at man i denne typen forskning ikke manipulerer variabler, og heller ikke kontrollerer for en tredje variabel (Lund, 2002b, 2002c).

Med *begrepsvaliditet* mener vi ”grad av samsvar mellom begrepet slik det er definert teoretisk og begrepet slik vi lykkes med å operasjonalisere det” (Kleven, 2002a, p. 176). Anvendt på denne konkrete undersøkelsen kan vi si at i hvilken grad de testene som er valgt måler de begrepene som ønskes å måles, er et uttrykk for begrepsvaliditet (Lund, 2002b).

*Ytre validitet* angår generalisering rettet mot en spesiell individpopulasjon, situasjon eller tid. Ytre validitet styrkes dersom utvalget består av relevante individer, situasjoner og tider (Lund, 2002b). Det er tre mulige trusler mot ytre validitet. Den første gjelder interaksjonen mellom uavhengig variabel og individer, situasjoner og tider. Den andre vedrører individhomogenitet, det vil si at individene i undersøkelsen er en forholdsvis ensartet gruppe, og at resultatene ikke lar seg generalisere til andre persongrupper. Den siste trusselen er ikke-representative individutvalg, det vil si at generalisering til en bestemt populasjon er usikker fordi forsøkspersonene ikke er representative for populasjonen (Lund, 2002b). En måte å styrke ytre validitet på, er å sikre at utvalget er representativt ved å bruke et tilfeldig utvalg i populasjonen (Lund, 2002b).

*Reliabilitet* betyr pålitelighet. Høy reliabilitet i forskning er et uttrykk for i hvilken grad data er fri for tilfeldige målingsfeil. Tilfeldige målingsfeil betyr feil som oppfører seg tilfeldig, og som vil jevne seg ut i det lange løp. Det er ofte et spørsmål om hvor nøyaktig eller konsistent en test måler det den skal (Kleven, 2002a). Ingen tester er fri for målefeil, og målefeil kan

oppstå hvis en test ikke skåres eller gjennomføres nøyaktig (Gall, et al., 2007). Det finnes to prinsipielt forskjellige måter å bedre reliabiliteten på. Den ene er å bruke standardisering av målingsprosedyrer, den andre er å nøytralisere effekten av målingsfeilene ved å utnytte at tilfeldige feil jevner seg ut i det lange løp (Kleven, 2002a). Bruk av standardiserte tester eller flere observatører kan bidra til å nøytralisere effekten av målingsfeilene.

### **3.3 Etikk**

Ved gjennomføringen av forskningsprosjekter er det viktig å ivareta etiske hensyn til elevene som testes. NESH (2009) legger vekt på at barn og unge som deltar i forskning, har særlig krav på beskyttelse i tråd med deres alder og behov. Barn må gis tilstrekkelig informasjon på sitt aldersnivå om prosjektet de skal delta i. Det er viktig at forskeren er seg sitt ansvar bevisst og har tilstrekkelig kunnskap om barn, slik at både metode og innhold av forskningen kan tilpasses det aktuelle alderstrinnet.

Når barn under 15 år skal delta i forskning, kreves det samtykke fra foreldrene. Selv om man har foreldrenes samtykke, er det viktig at man ser på barnet som et individuelt subjekt, og barnets egen aksept er nødvendig fra det er gammelt nok til å gi den. Barna som ble testet fikk en kort innføring om formålet med prosjektet før vi startet testingen, og fikk mulighet til å ta pauser mellom testene ved behov. Vi forsøkte å gi barna en mestringsfølelse uansett hvordan de presterte på testene, fordi dette ikke var en behandlingssituasjon der vi kunne følge opp eventuelle svake sider. Det ble derfor viktig å gi positiv tilbakemelding under hele testsituasjonen, og det ble også understreket overfor elevene at det ikke var meningen at de skulle kunne svare på alt. Vårt prosjekt er knyttet opp mot KiSP-prosjektet, og de formelle etiske hensyn som frivillighet, anonymitet og informert samtykke er ivaretatt av prosjektledelsen.

#### **Oppsummering av kapittel 3**

Utvalget i undersøkelsen vår er hentet fra utvalget i KiSP-prosjektet. Utvalget i vår undersøkelse består av 80 barn på 2., 5. og 6. trinn ved to skoler i Oslo. 47 av barna er jenter og 33 er gutter. Vi har benyttet seleksjonskriterier som gjør at utvalget vårt består av elever uten kjente hørselsvansker som har norsk som morsmål og nonverbal intelligens innenfor

normalområdet. Samtlige elever er testet med BPVS, TROG og Raven i tillegg til at læreren har fylt ut skjemaet 20 spørsmål om språkferdigheter for dem.

Data til undersøkelsen er samlet inn ved hjelp av testene TROG, BPVS og Raven i tillegg til spørreskjemaet 20 spørsmål om språkferdigheter. Som tilleggsinformasjon brukes data fra screeningtesten Språk 6-16.

Problemstillingen belyses med en empirisk-kvantitativ metodisk tilnærming, har et ikke-eksperimentelt design, og er deskriptiv i sin form. Skårer fra alle måleinstrumenter i undersøkelsen, med unntak av 20 spørsmål og Språk 6-16, er gjort om til standardskårer. Analysene som er benyttet er: reliabilitetsanalyse av 20 spørsmål, korrelasjonsanalyse, krystabeller, beregning av prediktiv verdi, t-tester og en enkel deskriptiv analyse.

Det er gjort kort rede for ulike kvalitetskrav som bør vurderes i undersøkelsen for å oppnå best mulig validitet. Til slutt ble etiske hensyn som er viktige å ivareta i forbindelse med undersøkelsen vår presentert.

I neste kapittel redegjøres det for resultater fra undersøkelsen.

# 4 RESULTATER

## 4.1 Innledning

Målet for denne undersøkelsen er å avdekke om TROG og BPVS gir en valid identifisering av barn med spesifikke språkvansker sett i forhold til lærernes vurdering av barna ved bruk av spørreskjemaet 20 spørsmål om språkferdigheter.

Når vi skal avdekke dette, må vi for det første finne ut om det er samvariasjon mellom TROG, som er et mål på grammatisk forståelse, og BPVS, som er et mål på reseptivt vokabular. Det forventes en viss samvariasjon dersom testene måler det samme underliggende begrepet (SSV), men da testene måler ulike aspekter ved begrepet forventes ikke samvariasjonen å være veldig sterk. Derneft undersøker vi hvordan hver av språktestene korrelerer med Raven, for å kontrollere om språktestene er rene språktester eller tester andre faktorer. I forbindelse med de ovennevnte undersøkelsene benyttes korrelasjonsanalyser. Analysene skal gi indikasjoner om språktestenes begrepsvaliditet.

For det andre vil vi finne svar på om det er en sammenheng mellom TROG og BPVS på den ene siden og 20 spørsmål på den andre. Dette er hovedspørsmålet i undersøkelsen, og handler om kriterievaliditet. Til dette formålet benyttes krystabeller, beregning av testenes nøyaktighet i forhold til 20 spørsmål og t-test for uavhengige utvalg. For å få informasjon om barns skårer på de ulike måleinstrumentene vil det bli presentert case ved hjelp av en enkel deskriptiv analyse. Til slutt undersøker vi hva som kjennetegner barn med mulig SSV ifølge de ulike instrumentene.

## 4.2 Deskriptiv statistikk

### 4.2.1 Karakteristikk av fordelinger

Når vi sammenligner fordelinger i tabeller og histogrammer, vil vi se at de varierer på ulike måter. Vi vil finne ut om resultatene i vår undersøkelse er normalfordelte, fordi dette har betydning for valg av analyseverktøy. I det følgende presenteres fordelingen av skårer på TROG, BPVS, Raven og 20 spørsmål for utvalget vårt, som består av 80 barn (47 jenter og 33 gutter) på 2., 5. og 6. trinn. Fordelingene presenteres i histogrammer og en tabell over

statistiske mål. Skårene for TROG, BPVS og Raven er oppgitt som z-skårer, mens 20 spørsmål om språkferdigheter har en egen skala.

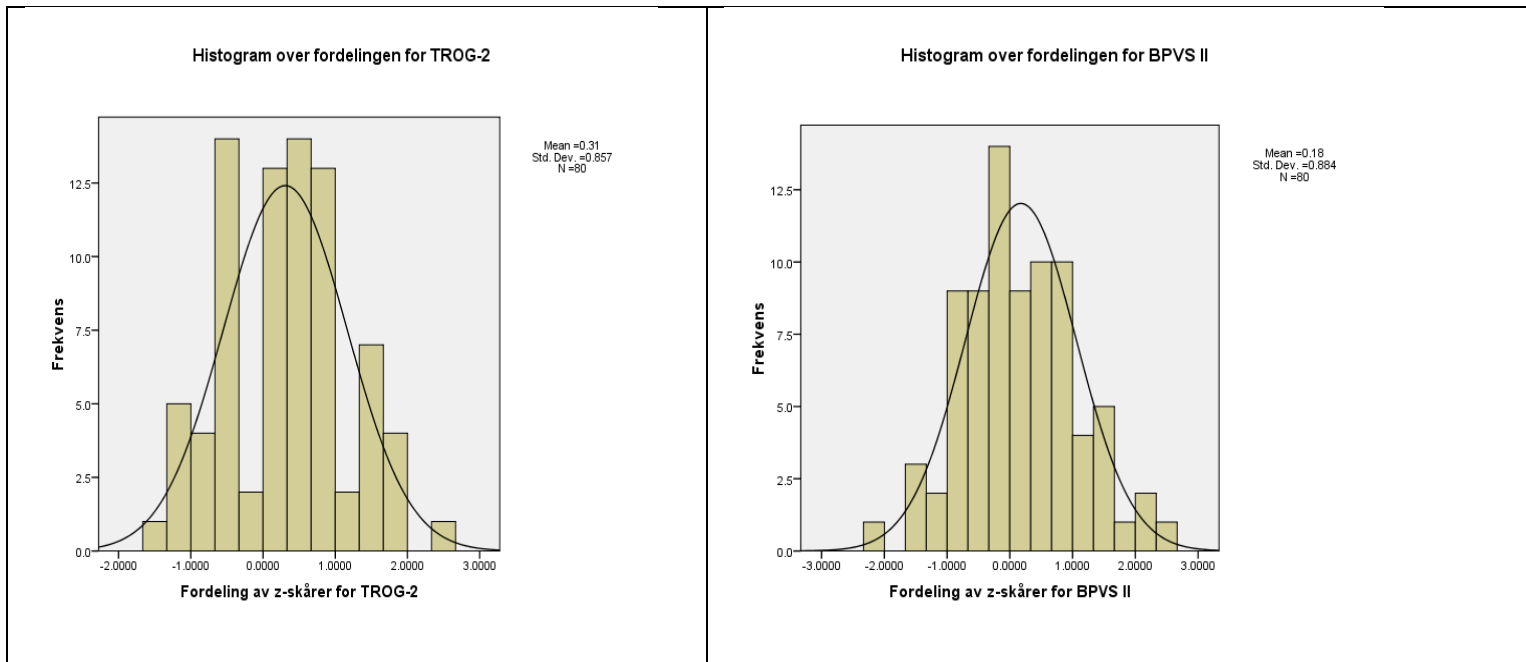
Fordelingen av skårer blir sett i forhold til en normalfordelingskurve, også kalt Gausskurven. Gausskurven er en symmetrisk, klokkeformet grafisk kurve der Mean, median og modus har samme verdi. Når vi deler variabelaksen inn med standardavvik som måleenhet for en slik fordeling, vil 68 % av fordelingen befinne seg +/- 1 standardavvik fra Mean, og praktisk talt alle observasjoner vil befinne seg i området +/-3 standardavvik fra Mean (Befring, 2007). Selv om en kurve kan se normalfordelt ut, kan den ved nærmere ettersyn ha en liten skjevhet som kan vise større eller mindre avvik fra normalfordelingen (Cristophersen, 2009). En skjevhetsfordeling kan være positivt skjev, og er da høyreskjev. Det vil si at den har en lang hale til høyre. Alternativt kan den være negativt skjev, det vil si venstreskjev. Den har da en lang hale til venstre (Field, 2009). Kurtosis viser om fordelingen er spissere eller flatere enn normalfordelingen. Kurtosis kan være positiv. Det vil si at den er spiss. Den kan også være negativ, og da er den flat (Cristophersen, 2009). I en normalfordeling er verdien på kurtosis og skjevhet tilnærmet 0. Jo nærmere man er denne verdien, jo nærmere er man en normalfordeling. Verdier for skjevhet og kurtosis som er fra -1 til 1, indikerer at fordelingen er lite skjev og at det er lite avvik fra normalfordelingen. Tilsvarende verdier rundt -2 og 2, indikerer en moderat skjev fordeling, og verdier inntil -3 og 3 en sterkt skjev fordeling. Verdier som ligger utenfor -3 og 3 karakteriseres som en svært sterkt skjev fordeling (Cristophersen, 2006). Når vi skal vurdere normalfordelingen for måleinstrumentene som er benyttet i undersøkelsen vår, blir det presentert et histogram for hvert av dem i tillegg til en tabell som viser sentrale statistiske mål for fordelingene.

## **Histogrammer over fordelinger**

Nedenfor vises histogrammer over fordelingen til alle måleinstrumentene i undersøkelsen. Den horisontale aksene viser fordelingen av z-skårer, og den vertikale aksene viser frekvensen av disse skårene.

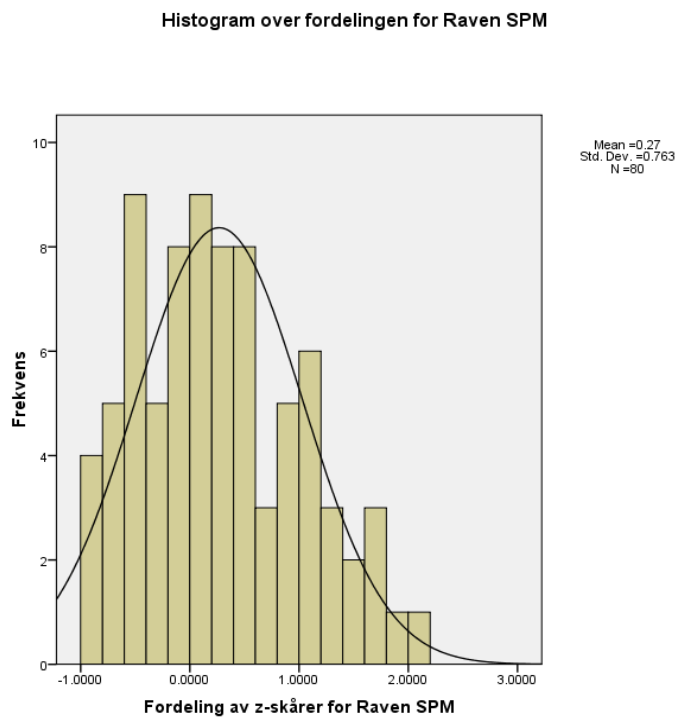


**Figur 1 viser fordelingen til TROG og BPVS.**



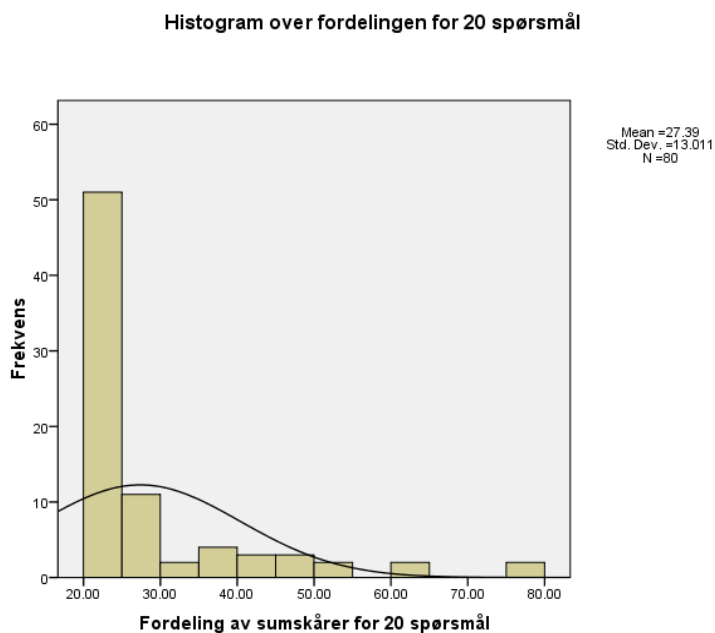
Vi ser av histogrammene over at TROG 2 og BPVS II har en tilnærmet normalfordeling.

**Figur 2 viser fordelingen til Raven**



Raven har en noe avvikende normalfordeling med en kurve som er en anelse høyreskjev. Grunnen til dette er at utvalget vårt består av barn som har IQ fra -1 SD fra Mean og høyere, og at vi dermed ikke får med barn med svake skårer.

**Figur 3 viser fordelingen til 20 spørsmål**



Figur 3 viser at 20 spørsmål om språkferdigheter ikke er normalfordelt. Vi ser at det er en opphopning av skårer til venstre for kurven. Den lange halen mot høyre viser at kurven er høyreskjev (positivt skjev). For 20 spørsmål om språkferdigheter innebærer en lav skåre at barnet har en liten vanske, mens en høyere skåre indikerer en større vanske. Utvalget vårt består av barn med en normal språkutvikling, og vi ser at mange har oppnådd gode resultater - klart de fleste har skårer rundt 20, som er det beste resultatet. 20 spørsmål er ikke laget for å gi normalfordeling, men for å gi en nyansert vurdering av barn med språkproblemer. Vi vil nå se nærmere på relevante statistiske mål for fordelingen til TROG, BPVS, 20 spørsmål og Raven.

### **Sentrale statistiske mål for fordelingene**

For å gi en mer nøyaktig beskrivelse av utvalget, presenterer vi en tabell med sentrale statistiske mål for fordelingene. Tabell 1 inneholder informasjon om antall barn i

undersøkelsen (N), Mean (M), median (Md), standardavvik (SD), minimumsskåre (Min), maksimumsskåre (Max), fordelings skjevhet, og om fordelingen er spiss eller flat (kurtosis).

**Tabell 1: Statistiske mål for TROG, BPVS, Raven og 20 spørsmål.**

	N	M	Md	SD	Min.	Max.	Skjevhet	Kurtosis
TROG	80	.30	.35	.85	-1.65	2.40	.096	-.353
BPVS	80	.17	.10	.88	-2.06	2.40	.131	-.038
20 spørsmål	80	27.38	20.00	13.01	20.00	77.00	2.17	4.49
Raven	80	.26	.16	.76	-.94	2.05	.379	-.670

De statistiske målene for TROG og BPVS i Tabell 1 viser at begge testene er tilnærmet normalfordelte. Dette ser vi fordi Mean og Median er omtrent like. For øvrig er standardavviket under 1, og det vil si at hovedvekten av fordelingene befinner seg innenfor normalområdet. Skjevhet og kurtosis er tilnærmet lik 0. Kurtosis har negativ verdi for alle de tre testene, noe som betyr at kurvene har en noe flat fordeling. For Raven ser vi at Mean (M=0.27) er litt større enn Medianen (Md=0.16). Det betyr at kurven er noe høyreskjev. Standardavviket på 0.76 befinner seg innenfor +/- 1 standardavvik fra Mean, og er altså innenfor normalområdet. Verdiene for skjevhet og kurtosis er under 1, og er derfor ikke store. Vi ser at kurtosis også her har en negativ verdi, noe som betyr at kurven har en noe flat form.

Når det gjelder 20 spørsmål, ser vi at Mean er 27.39 og Median er 20. Mean er større enn Medianen, noe som gjør at vi får en positivt skjev kurve. Skjevheten er over 2 og dermed moderat skjev. Kurtosis er over 4, og kurven er dermed veldig spiss.

Vi har sett på hvordan skårene fordeler seg, og har funnet at BPVS, TROG og Raven er tilnærmet normalfordelte, mens 20 spørsmål ikke er det. Denne informasjonen ligger til grunn

for hvilke statistiske analyser vi velger til å presentere resultater som besvarer vår problemstilling. Vi er interessert i om det er samvariasjon mellom disse måleinstrumentene for å finne ut om testene kan betraktes som mål på det samme underliggende begrepet og om testene også påvirkes av kognitiv fungering. For å avdekke dette vil vi som et første steg finne ut om TROG, BPVS og Raven korrelerer med hverandre. Til dette formålet benytter vi en korrelasjonsanalyse. Vi vil også finne ut om 20 spørsmål om språkferdigheter samvarierer med TROG og BPVS. Da 20 spørsmål ikke er normalfordelt, kan vi ikke bruke en korrelasjonsanalyse mellom dette og andre måleinstrumenter, men undersøker i stedet sammenhengene ved hjelp av krystabeller, t-test, beregning av sensitivitet og spesifisitet og en enkel deskriptiv analyse.

Før vi går videre med analyser, vil vi i neste avsnitt foreta en reliabilitetsanalyse for 20 spørsmål.

#### **4.2.2 Reliabilitetsanalyse**

Ved gjennomføring av reliabilitetsanalyse for 20 spørsmål fant vi at måleinstrumentet hadde en Cronbachs alpha-verdi på .96. Det var ingen items som skilte seg ut med lavere eller høyere verdier. Når alpha-verdien ligger like oppunder 1, betyr det at måleinstrumentet har en god indre konsistens.

#### **4.2.3 Korrelasjoner**

Det antas at både TROG og BPVS hver for seg kan identifisere SSV, og at testene derfor kan betraktes som mål på det samme underliggende begrepet. I denne sammenhengen er det interessant å finne ut i hvilken grad testene måler det samme, og om testene påvirkes av kognitiv fungering målt ved den nonverbale testen Raven. I det følgende benyttes korrelasjonsanalyse for å undersøke dette.

Korrelasjon viser som nevnt i Kapittel 3 (avsnitt 3.1.6) samgang mellom forskjellige variabler (Befring, 2007). Nedenfor foretar vi en korrelasjon mellom TROG, BPVS og Raven. Det vil bli benyttet bivariat korrelasjonsanalyse for å si noe om signifikansen og styrken på variablene. I vår analyse velger vi å benytte Pearsons  $r$  som korrelasjonsmål fordi variablene i vår undersøkelse er på et forholdstallsnivå, og fordi sammenhengen mellom variablene er tilstrekkelig lineær (Svartdal, 2004).

## Korrelerer TROG, BPVS og Raven med hverandre?

Tabell 2: korrelasjoner mellom TROG, BPVS og Raven

	TROG	BPVS	Raven
TROG	1		
BPVS	.363**	1	
Raven	.405**	.225*	1

\* Korrelasjonen er signifikant på .05-nivået

\*\*Korrelasjonen er signifikant på .01-nivået

Korrelasjonen mellom BPVS og TROG er  $r = .363$  ( $p < .001$ ) og mellom TROG og Raven er  $r = .405$  ( $p = .000$ ). Det betyr at korrelasjonene er signifikante på .01-nivå, og at man med 99 % sikkerhet kan si at resultatet ikke skyldes tilfeldigheter (Fraenkel & Wallen, 2008). BPVS og Raven korrelerer på  $r = .225$  ( $p < .045$ ) og er signifikant på .05-nivå. Det betyr at vi med 95% sikkerhet kan si at resultatet ikke skyldes tilfeldigheter. For å finne ut hvor stor del av variansen som er felles for de to variablene, kan korrelasjonskoeffisienten kvadreres (Connolly, 2007; Kleven, 2007b). Når vi gjør det, ser vi at BPVS og TROG har 13 % av variansen felles, TROG og Raven har 16 % av variansen felles og at BPVS og Raven har 5 % av variansen felles. Det meste av variansen er derfor uforklart. Det betyr at det er andre forhold som virker inn på sammenhengen, som ikke fanges opp av testene i undersøkelsen.

**Konklusjon:** Bivariat korrelasjonsanalyse viser at det er signifikante sammenhenger mellom måleinstrumentene i undersøkelsen vår. Vi fant en korrelasjon på .01-nivå mellom BPVS og TROG og mellom TROG og Raven. Det betyr at man med 99 % sikkerhet kan si at resultatet ikke skyldes tilfeldigheter. Signifikansen er noe svakere for BPVS og Raven, men innenfor .05-nivået. Det betyr at man med 95 % sikkerhet kan si at resultatet ikke skyldes tilfeldigheter (Fraenkel & Wallen, 2008). Det meste av variansen er uforklart for alle korrelasjonene.

Vi har nå funnet korrelasjonene mellom TROG, BPVS og Raven. I det følgende vil vi finne ut hvordan de ulike måleinstrumentene samsvarer med hverandre når det gjelder å identifisere barn med SSV.

#### **4.2.4 Hvordan samsvarer måleinstrumentene med hverandre?**

Som vi har sett av korrelasjonsanalysene, samsvarer måleinstrumentene bare delvis med hverandre. For å avdekke hvordan de ulike måleinstrumentene grupperer de ulike barna, brukes krysstabeller. Krysstabellene gjør det også mulig å få mer detaljert informasjon om forskjeller og likheter mellom måleinstrumentene enn det korrelasjonene gjør. Nedenfor presenteres tre krysstabeller som viser sammenhengen mellom måleinstrumentene. For å gjøre tabellene mer oversiktlige, har vi som nevnt tidligere inndelt variablene i grupper. Utvalget består av barn med nonverbal IQ  $>-1$  SD, det vil si i den delen av materialet der en skal forvente å finne de spesifikke språkvanskene.

Vi vil i det følgende beskrive sammenhengen mellom de ulike måleinstrumentene slik de fremkommer i krysstabellene. For hver krysstabell kommenteres først hovedtendensen. Deretter tar vi for oss hvilke barn som kan ha SSV ifølge de ulike måleinstrumentene.

## Hvilken sammenheng er det mellom TROG og BPVS?

**Tabell 3: Krysstabell som viser sammenhengen mellom TROG og BPVS**

		TROG				
		Gruppe 1	Gruppe 2	Gruppe 3	Totalt	
BPVS	Gruppe 1	Antall	5	8	0	13
		% av total	6.2%	10.0%	.0%	16.2%
	Gruppe 2	Antall	9	53	1	63
		% av total	11.2%	66.2%	1.2%	78.8%
	Gruppe 3	Antall	0	3	1	4
		% av total	.0%	3.8%	1.2%	5.0%
	Totalt	Antall	14	64	2	80
		% av total	17.5%	80.0%	2.5%	100.0%

Khi-kvadrat: 13.671<sup>a</sup>; DF: 4; p-verdi: .008.

a – 6 celler (66,7%) har forventet antall under 5. Minste forventede antall er 10.

Tabell 3 viser sammenhengen mellom skårer på TROG og BPVS (Khi-kvadrat: 13.671<sup>a</sup>; DF: 4; p-verdi: .008). P-verdien er signifikant, men det er for små verdier i hver celle til at vi kan tolke khi-kvadrat-verdiene.

Det er likevel interessant å se hvordan barnas skårer fordeler seg i tabellen. Tabellen viser at TROG og BPVS enes om at hovedtyngden av de 80 barna befinner seg i normalområdet, det vil si i gruppe 2. Testene er imidlertid ikke helt samstemte om rangeringen av alle barna. TROG plasserer totalt 64 barn i denne kategorien, mens BPVS plasserer 63 barn her. Blant disse enes de om 53 barn. BPVS rangerer fire barn i gruppe 3, og TROG bare to. Testene enes bare om ett av barna i denne gruppen. Det andre barnet som TROG rangerer i denne gruppen, rangeres i gruppe 2 av BPVS.

**Konklusjon:** TROG og BPVS enes om at storparten av barna er normalspråklige, men BPVS identifiserer flere barn med SSV enn TROG. Måleinstrumentene enes bare om ett barn i gruppe 3, og det er i denne gruppen vi finner barn som kan ha SSV.

**Tabell 4: Krysstabell som viser sammenhengen mellom TROG og 20 spørsmål**

		20 spørsmål				
		Gruppe 1	Gruppe 2	Gruppe 3	Totalt	
<b>TROG</b>	Gruppe 1	Antall	12	2	0	14
		% av total	15.0%	2.5%	.0%	17.5%
	Gruppe 2	Antall	49	10	5	64
		% av total	61.2%	12.5%	6.2%	80.0%
	Gruppe 3	Antall	1	0	1	2
		% av total	1.2%	.0%	1.2%	2.5%
	Totalt	Antall	62	12	6	80
		% av total	77.5%	15.0%	7.5%	100.0%

Khi-kvadrat: 6.521<sup>a</sup>; DF: 4; p-verdi: .163.

a – 6 celler (66,7%) har forventet antall under 5. Minste forventede antall er 15.

Tabell 4 viser sammenhengen mellom TROG og 20 spørsmål. Også her er p-verdien signifikant, men det er for små verdier i hver celle til at vi kan tolke khi-kvadrat-verdiene. TROG og 20 spørsmål er enige om at hovedtyngden av barna befinner seg innenfor normalområdet av språklige ferdigheter, det vil si i gruppe 1 for 20 spørsmål og i gruppe 2 for TROG. Ifølge TROG gjelder dette 64 barn, og ifølge 20 spørsmål hører 62 barn hjemme her. Blant disse barna er måleinstrumentene enige om plassering av 49 av dem i normalgruppen. Når det gjelder hvilke barn som kan ha SSV, regner lærerne uttrykt ved 20 spørsmål seks barn i denne gruppen, mens det ifølge TROG bare er to. Måleinstrumentene er bare enige om ett av barna i denne gruppen. Det andre barnet som ifølge TROG kan ha språkvansker, hører ifølge 20 spørsmål hjemme i gruppe 1, som er de normalspråklige barna. Fem av barna som plasseres i gruppe 3 på 20 spørsmål, hører hjemme i gruppe 2 ifølge TROG.

**Konklusjon:** TROG og 20 spørsmål enes om at hovedtyngden av barna er normalspråklige, men 20 spørsmål identifiserer betydelig flere barn som kan ha SSV, enn TROG. De enes bare om ett barn i gruppe 3, som er gruppen der vi finner barn som kan ha SSV.



## Hvilken sammenheng er det mellom BPVS og 20 spørsmål?

Tabell 5: Krysstabell som viser sammenhengen mellom BPVS og 20 spørsmål

		20 spørsmål				
		Gruppe 1	Gruppe 2	Gruppe 3	Totalt	
BPVS	Gruppe 1	Antall	10	2	1	13
		% av total	12.5%	2.5%	1.2%	16.2%
	Gruppe 2	Antall	50	10	3	63
		% av total	62.5%	12.5%	3.8%	78.8%
	Gruppe 3	Antall	2	0	2	4
		% av total	2.5%	.0%	2.5%	5.0%
	Totalt	Antall	62	12	6	80
		% av total	77.5%	15.0%	7.5%	100.0%

Khi-kvadrat: 11.316<sup>a</sup>; DF: 4; p-verdi: .023.

a – 6 celler (66,7%) har forventet antall under 5. Minste forventet antall er 30.

Tabell 5 viser sammenhengen mellom BPVS og 20 spørsmål. P-verdien er signifikant, men det er for små verdier i hver celle til at vi kan tolke khi-kvadrat-verdiene. Tabellen viser at BPVS og 20 spørsmål er enige om at hovedtyngden av barna befinner seg i normalområdet for språklige ferdigheter som tilsvarer gruppe 1 for 20 spørsmål og gruppe 2 for BPVS. BPVS regner 63 barn for å være i denne gruppen, mens 20 spørsmål regner 62 barn for å høre hjemme her. Måleinstrumentene enes om 50 av barna som befinner seg i normalområdet. Når det gjelder hvilke barn som har SSV, er dette seks barn ifølge 20 spørsmål og fire ifølge BPVS. Når det gjelder barna i denne gruppen, er BPVS og 20 spørsmål bare enige om to. Ett barn som ifølge 20 spørsmål har store språkvansker, rangeres blant barna med svært gode språkferdigheter av BPVS. Tre barn som 20 spørsmål regner for å ha en alvorlig språkvanske, regnes for å ha normale språkferdigheter ifølge BPVS.

**Konklusjon:** BPVS og 20 spørsmål enes om at hovedtyngden av barna er normalspråklige. 20 spørsmål identifiserer flere barn som kan ha SSV enn BPVS. De enes bare om to barn i gruppe 3, som er barn som kan ha SSV.

#### 4.2.5 Prediktiv verdi

Prediktiv verdi evalueres ved å fastsette i prosent hvor mange korrekte klassifiseringer testen eller testbatteriet gjør, det vil si hvor nøyaktig testen plasserer personer i bestemte grupper på grunnlag av cut-off-skårer. Dette gjøres ved å fastsette i hvilken grad testen og det uavhengige kriteriet gjør de samme klassifiseringene (Sattler, 2001). I denne undersøkelsen er 20 spørsmål om språkferdigheter det uavhengige kriteriet som språktestene måles mot.

Hvor nøyaktig en test måler det den skal måle, angis vanligvis som testene sensitivitet og spesifisitet (se avsnitt 2.4). Når spesifisitet og sensitivitet skal beregnes, tar vi utgangspunkt i utfall på de ulike måleinstrumentene som fremgår av krysstabellene mellom TROG og 20 spørsmål, og BPVS og 20 spørsmål. Som nevnt tidligere er det fire mulige utfall av en test:

- a) Sann positiv vil si at barnet har en vanske, og testen fanger dette opp. Som krysstabellene viser, fanger TROG opp ett barn, mens BPVS fanger opp to barn.
- b) Falsk positiv vil si at barnet ikke har noen vanske, men testen viser at barnet har en vanske. (Dette fører til overrapportering.) Som krysstabellene viser, fanger TROG opp ett barn, mens BPVS fanger opp to barn.
- c) Falsk negativ vil si at barnet har en vanske, men testen fanger ikke dette opp. (Dette fører til underrapportering.) Som krysstabellene viser, er det fem barn som TROG ikke fanger opp, mens det er fire barn som BPVS ikke fanger opp.
- d) Sann negativ vil si at barnet ikke har noen vanske, og testen angir at barnet ikke har noen vanske. Som krysstabellene viser, fanger TROG opp 73 barn, mens BPVS fanger opp 72 barn.

Nedenfor følger beregninger av sensitivitet og spesifisitet for hver av testene.

## Sensitiviteten til testene TROG og BPVS

$$\text{Sensitivitet} = \frac{\text{Sann positiv}}{\text{Sann positiv} + \text{falsk negativ}}$$

### TROG

$$\text{Sensitivitet} = \frac{1}{1+5} \approx 0.17 \approx 17 \%$$

Sensitiviteten til TROG er på ca. 17%.

### BPVS

$$\frac{2}{2+4} \approx 0.33 \approx 33 \%$$

Sensitiviteten til BPVS er på ca. 33%.

## Spesifisiteten til testene TROG og BPVS

$$\text{Spesifisitet} = \frac{\text{Sann negativ}}{\text{Sann negativ} + \text{falsk positiv}}$$

### TROG

$$\text{Spesifisitet} = \frac{73}{73+1} \approx 0.99 \approx 99 \%$$

Spesifisiteten til TROG er på ca. 99%.

## BPVS

$$\frac{72}{72+2} \approx 0.97 \approx 97\%$$

Spesifisiteten til BPVS er på ca. 97%.

**Konklusjon:** Resultatene over viser at både TROG og BPVS har svært høy spesifisitet på henholdsvis 99 % og 97 %. TROG har en sensitivitet på 17 % og BPVS på 33 %.

### 4.2.6 T-test for uavhengige utvalg

Det er foretatt en t-test for uavhengige utvalg for å sammenligne gjennomsnittsskåren på TROG og BPVS hos barn som kan ha SSV og barn som ikke har SSV ifølge 20 spørsmål. T-testen, som er signifikant på .01 nivå, viser at de seks barna som lærerne angir har mulig SSV, har svakere gjennomsnittsskåre på både BPVS ( $M = -.69$ ) og TROG ( $M = -.64$ ) enn de som ifølge lærerne er normalspråklige ( $M = 0.25$  på BPVS og  $0.38$  på TROG). Dette viser at elever som har mulig SSV ifølge lærerne, har signifikant lavere skårer på TROG og BPVS enn elever som ifølge lærerne er normalspråklige.

### 4.2.7 Enkel deskriptiv analyse

Vi har gjort en enkel deskriptiv analyse gruppert etter skåringsnivået på 20 spørsmål som gir oss mye informasjon om testresultatene til hvert enkelt barn (se vedlegg 2). Denne analysen gir en detaljert og komplett oversikt med informasjon fra alle måleinstrumentene for hvert barn. En fordel med en slik analyse, er at den i motsetning til de bivariate krysstabellene presenterer data på en måte som gjør det lettere å sammenligne barnas skårer på de ulike måleinstrumentene med hverandre. I tillegg til barnas skårer på 20 spørsmål om språkferdigheter, BPVS og TROG, tar vi med tilleggsdata fra Språk 6-16 der det foreligger.

For å gjøre det lettere å sammenligne tester og lærevurderinger kommer vi først til å ta for oss de barna som ifølge lærerne har normalspråklig fungering, deretter de barna som har milde/moderate vansker og til slutt de barna som har alvorlige vansker ifølge lærerne. For hver av disse gruppene kommenterer vi hvilke indikasjoner TROG og BPVS gir om de

samme barnas språkferdigheter. Det er de barna som kan ha SSV ifølge ett eller flere av måleinstrumentene, vi synes det er interessant å se nærmere på. Disse presenteres nedenfor.

### **Gruppe 1: Lærerne angir at elevene har normal språklig fungering, hva sier testene?**

Vedlegg 2 viser at 62 barn (77,5 %) befinner seg i denne gruppen. I de fleste tilfellene støtter TROG og BPVS lærernes vurderinger. Testene samsvarer også stort sett med hverandre.

For tre av barna i denne gruppen er det et misforhold mellom tester og lærervurderinger (skårene oppgis i parentes):

Case 12 er en jente på småskoletrinnet som skårer innenfor normalområdet på TROG (-0.08 SD), mens BPVS (-1.40 SD) angir at dette barnet kan ha SSV. Ifølge lærerne uttrykt ved 20 spørsmål (skåre: 20) har barnet ingen vansker.

Case 34 er en jente på småskoletrinnet som skårer innenfor normalområdet på TROG (-.41 SD), mens BPVS (-1.34 SD) indikerer at hun kan ha SSV. Ifølge lærerne uttrykt ved 20 spørsmål (skåre: 20) har barnet ingen vansker.

Case 61 er en jente på mellomtrinnet der TROG (-1.28 SD) indikerer at barnet kan ha SSV, mens BPVS (.94 SD) er innenfor normalområdet i likhet med 20 spørsmål (skåre: 21). Det er gitt bare én anmerkning av læreren i forbindelse med utsagnet Har vansker med å huske ting, som har fått skåren 2.

**Konklusjon:** Det er ikke samsvar mellom testene i noen av de nevnte tilfellene. I ett tilfelle indikerer TROG at barn i denne gruppen kan ha en spesifikk språkvanske, mens BPVS gir indikasjon om dette i to tilfeller. Siden barna ikke er vurdert som språklig svake av lærerne, har de ikke tatt Språk 6-16. Vi får derfor ingen tilleggsinformasjon om disse elevene. Skåringsnivåene ligger svært nær -1.25 SD, og det er derfor lite sannsynlig at disse barna har store språkvansker i formell forstand.

### **Gruppe 2: Lærerne angir at elevene har milde/moderate vansker, hva sier testene?**

Vi vil minne om at denne gruppen består av barn der lærerne har observert vansker med språket, uten at de ifølge våre kriterier regnes som barn med spesifikke språkvansker. Gruppe

2 består av 12 barn (15 %). TROG og BPVS støtter lærernes vurdering av at det ikke er barn med SSV i denne gruppen. Vi kommenterer ikke case i denne gruppen, fordi det interessante i vår undersøkelse er å se nærmere på barn som kan ha SSV ifølge måleinstrumentene.

### **Gruppe 3 . Lærerne angir at elevene har alvorlige språkvansker, hva sier testene?**

I denne gruppen er det seks barn (7,5 %). Vi finner her de barna som ifølge lærernes indikasjoner kan ha SSV.

I følgende tilfeller er det misforhold mellom tester og lærervurderinger:

Case 75 er en gutt på mellomtrinnet. Lærerne angir at han kan ha alvorlige språkproblemer (skåre: 62), men verken TROG (-.39 SD) eller BPVS (-.67 SD) fanger opp dette. Språk 6-16 angir derimot at barnet har problemer med både Begreper (6), Ordspenn (3) og Setningsminne (3).

Case 76 er en gutt på mellomtrinnet. Lærerne angir ved hjelp av 20 spørsmål at det kan være en alvorlig språkvanske (skåre: 77). Både TROG (-1.66 SD) og BPVS (-2.07 SD) samsvarer med lærerne om at barnet har store vansker. På TROG skårer barnet over 1 SD under Mean, og på BPVS skårer barnet over 2 SD under Mean. Barnet har svake skårer på alle tre deltestene til Språk 6-16. Begreper (2) er den svakeste deltesten, men skårene på Ordspenn (5), og Setningsminne (6) er også lave. Vi ser at dette er et barn alle måleinstrumentene er enige om at har en alvorlig språkvanske, og dette kan være et barn med SSV.

Case 77 er en jente på mellomtrinnet. TROG (-.39 SD) og BPVS (-.53 SD) fanger ikke opp det læreren vurderer som en alvorlig språkvanske (skåre: 54). Språk 6-16 viser at hun har vansker med Begreper (5) og Setningsminne (4), men at Ordspenn er innen normalen (7).

Case 78 er en jente på mellomtrinnet. Læreren angir at barnet kan ha alvorlige språkvansker (skåre: 76), mens TROG (-1.01 SD) og BPVS (-.76 SD) ikke fanger opp dette. Ifølge testene har barnet ingen språkvansker. Dette barnet har ikke tatt testen Språk 6-16.

Case 79 er en gutt på mellomtrinnet. Læreren angir at han kan ha alvorlige språkvansker (skåre: 53), mens TROG (.69 SD) og BPVS (1.51 SD) ikke fanger opp dette. Dette barnet har ikke tatt testen Språk 6-16.

Case 80 er en jente på mellomtrinnet. Lærerne angir at barnet har alvorlige språkvansker (skåre: 60). BPVS (-1.64 SD) støtter lærerne, mens TROG (-1.08) angir at barnet ikke har problemer. Språk 6-16 viser at barnet har problemer med Begreper (5), men hun har gode skårer på Ordspenn (9) og Setningsminne (10).

**Konklusjon:** Vi ser at i gruppe 1 vurderer lærerne elevene som språklig sterke, og TROG gir støtte for dette i 61 av 62 tilfeller, mens BPVS gir støtte for dette i 60 tilfeller.

I gruppe 2 finner vi elever som ifølge lærerne kan ha en mild/moderat vanske, men ikke spesifikke språkvansker slik vi har definert det. TROG og BPVS er enige om at det ikke er barn med spesifikke språkvansker i denne gruppen. De to testene gir heller ingen indikasjoner på at barna i denne gruppen strever med språket.

I gruppe 3 er det seks elever som ifølge lærernes vurderinger har en alvorlig språkvanske. TROG støtter denne vurderingen i ett av seks tilfeller, og BPVS i to av seks tilfeller. Det vil si at i de fleste tilfeller samsvarer ikke testene med lærernes vurderinger når det gjelder å identifisere barn med SSV. I alle fire tilfellene der Språk 6-16 er administrert til barn i gruppe 3, er det minst én av deltestene som er under normalområdet.

Tre av barna med mulig SSV ifølge 20 spørsmål er gutter, og tre er jenter. Ett av barna med mulig SSV ifølge TROG er jente, og ett barn er gutt. Tre av barna med mulig SSV ifølge BPVS er jenter, én er gutt.

Barnas skårer på de ulike måleinstrumentene sier lite om hva disse barna eventuelt strever med i det daglige. 20 spørsmål gir en funksjonell beskrivelse av barnas vansker. Kan utsagnene i 20 spørsmål og feiltyper på TROG og BPVS forklare hvorfor de ulike måleinstrumentene vurderer barnets vanske forskjellig? Har disse barna noe til felles? I neste avsnitt vil vi se nærmere på kvalitative sider ved barnas vansker for å forsøke å finne svar på dette.

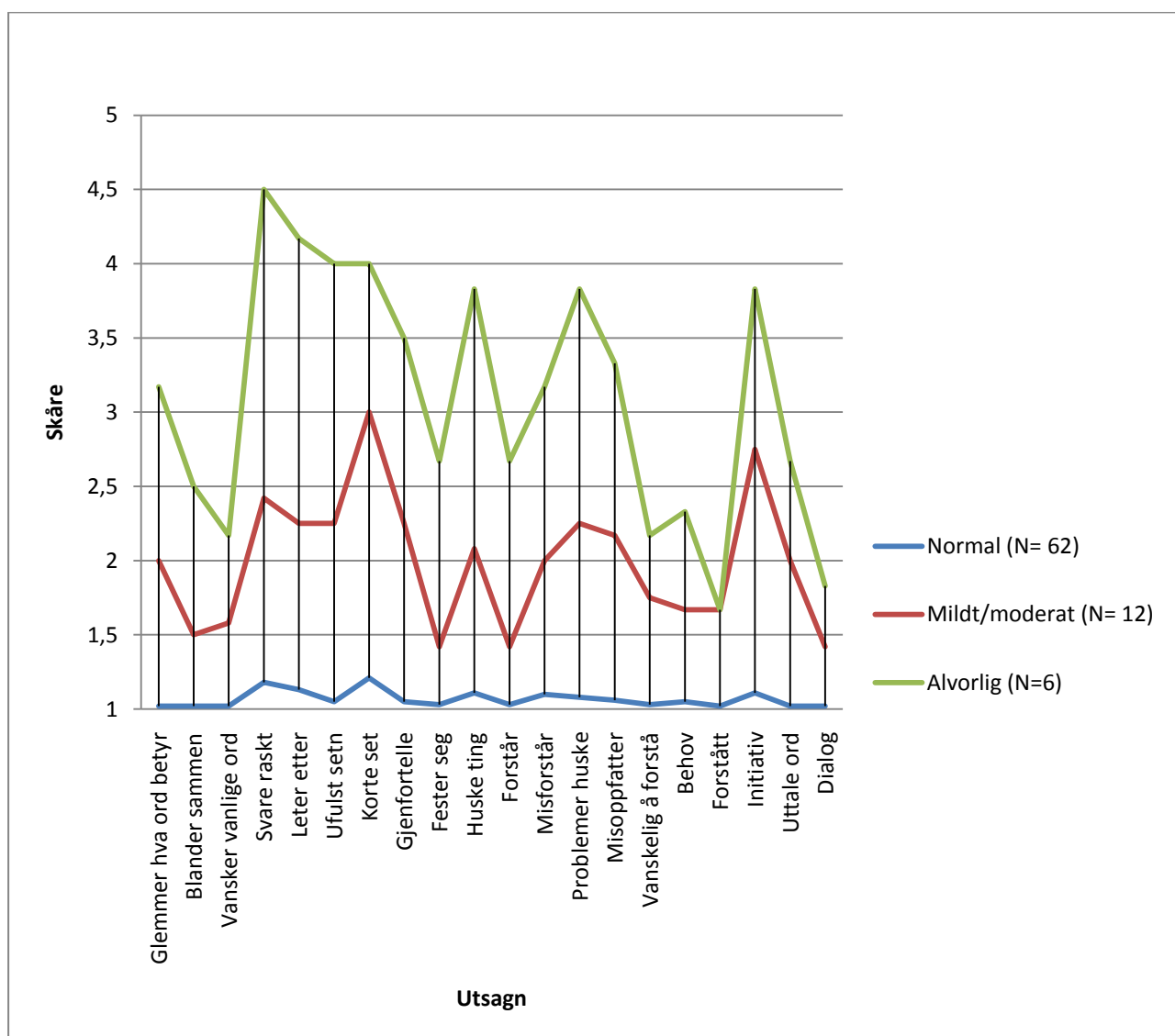
#### **4.2.8 Hva kjennertegner barn med mulig SSV ifølge 20 spørsmål?**

De 20 spørsmålene representerer en oversikt over viktige sider ved barns språklige kompetanse. Utsagnene beskriver vansker som er vanlige hos barn med spesifikke språkvansker. Ottem (2009) har gruppert disse utsagnene i tre kategorier: semantisk evne,

vansker med å motta et budskap og vansker med å formidle et budskap. I Vedlegg 1 gis en oversikt over de 20 utsagnene slik de er gruppert i disse kategoriene.

Det er ikke sikkert at barna har like store vansker innen alle de tre kategoriene. Figur 4 gir en grafisk fremstilling av hvordan normalspråklige barn og barn med milde/moderate eller alvorlige språkvansker skårer på hvert utsagn i 20 spørsmål.

**Figur 4: Fordelingen av råskårer på 20 spørsmål for normalspråklige barn, barn med milde/moderate språkvansker og barn med alvorlige språkvansker (SSV)**



Hovedtendensen i figuren over er at de normalspråklige barna i utvalget har en svært jevn, flat profil, det vil si at de har lave skårer på alle utsagnene og dermed ikke har språklige vansker



av betydning på noe område. Barna med milde/moderate vansker har en mer ujevn profil, og skårer vesentlig høyere enn de normalspråklige barna på alle utsagnene. Barna som kan ha alvorlige språkvansker, skårer vesentlig høyere enn disse igjen på nesten alle utsagnene. Vi vil se nærmere på den sistnevnte gruppen.

Figuren over viser at barna med alvorlige språkvansker i utvalget vårt har størst vansker med å svare på spørsmål like raskt som andre. En annet typisk trekk er at barna ifølge lærerne ofte leter etter de riktige ordene, bruker ufullstendige setninger, bruker korte setninger når de skal svare på spørsmål og har vansker med å gjenfortelle en historie de har hørt. Felles for alle disse utsagnene er at de hører til under kategorien Semantisk evne, som ser ut til å utgjøre den største vansken for disse barna. Barna med alvorlige språkvansker har også vansker med å huske ting og beskjeder, og disse vanskene hører innunder kategorien Vansker med å motta et budskap. Det siste utsagnet som skiller seg ut som spesielt typisk for disse barna, er at de sjelden tar språklige initiativ. Dette utsagnet hører innunder kategorien Vansker med å formidle et budskap. De tre sistnevnte utsagnene ser ut til å ha like høye skårer hos barna i denne gruppen. Ved første øyekast ser det ut til at barna har flere vansker med å motta informasjon enn å formidle informasjon, men ved nærmere ettersyn ser vi at de to utsagnene i kategorien Vansker med å motta et budskap er ganske like.

Nedenfor presenteres case for de seks barna som har mulig SSV ifølge 20 spørsmål. Det blir gitt eksempler på vansker som læreren har gitt høyest skåre (skårene 4 og 5). For hvert case oppgis for øvrig sumskåre og delskårer for hver av de tre kategoriene Semantiske vansker, Vansker med å motta et budskap og Vansker med å formidle et budskap. I den enkle analysen i vedlegg 2, oppgis sumskårer på 20 spørsmål, men av plasshensyn angis ikke delskårer der.

## Case

Case 75, gruppe 3 (gutt på mellomtrinnet). Sumskåre: 62

- Semantiske vansker (67.5): Har problemer med å svare på spørsmål like raskt som andre (5), Bruker korte setninger når han/hun skal svare på spørsmål (4), Har vansker med å gjenfortelle en historie han/hun har hørt (4).
- Vansker med å motta (66,67): Har vansker med å huske ting (4), Har problemer med å huske beskjeder (5).

- Vansker med å formidle (50): Tar sjelden språklige initiativ (4).

Case 76, gruppe 3 (gutt på mellomtrinnet). Sumskåre: 77

- Semantiske vansker (90): Glemmer ord som han/hun vet hva betyr (4), Blander sammen ord med lik mening (4), Har vansker med å forstå hva vanlige ord betyr (4), har problemer med å svare på spørsmål like raskt som andre (5), Leter ofte etter de riktige ordene (5), Bruker ufullstendige setninger (5), Bruker korte setninger når han/hun skal svare på spørsmål (5), Har vansker med å gjenfortelle en historie han /hun har hørt (4).
- Vansker med å motta (80): Det virker som det han/hun lærer ikke fester seg (4), Har vansker med å huske ting (5), Har problemer med å huske beskjeder (5), Misoppfatter situasjoner og hendelser (4).
- Vansker med å formidle (56.67): Problemer med å gi uttrykk for behov (4), Har problemer med å uttale ord riktig (4).

Case 77, gruppe 3 (jente mellomtrinnet). Sumskåre: 54

- Semantiske vansker (62.50): Har problemer med å svare på spørsmål like raskt som andre (4)
- Vansker med å motta (40)
- Vansker med å formidle (56.57): Tar sjelden språklige initiativ (4).

Case 78, gruppe 3 (jente på mellomtrinnet). Sumskåre: 76

- Semantiske vansker (87.5): Glemmer ord som han/hun vet hva betyr (5), Har vansker med å forstå hva vanlige ord betyr (4), Har problemer med å svare på spørsmål like raskt som andre (5), Leter ofte etter de riktige ordene (5), Bruker ufullstendige

setninger (5), Bruker korte setninger når han/hun skal svare på spørsmål (5), Har vansker med å gjenfortelle en historie som han/hun har hørt (5).

- Vansker med å motta (76.67): Det virker som det han/hun lærer ikke fester seg (5), Har vansker med å huske ting (5), Misforstår muntlige instruksjoner og beskjeder (4), Misoppfatter situasjoner og hendelser (4).
- Vansker med å formidle (60): Tar sjelden språklige initiativ (5), Har problemer med å uttale ord riktig (5), Mestrer ikke dialog med jevnaldrende (5)

Case 79, gruppe 3 (gutt på mellomtrinnet). Sumskåre: 53

- Semantiske vansker (57.5): Har problemer med å svare på spørsmål like raskt som andre (4), Leter ofte etter de riktige ordene (5), Bruker ufullstendige setninger (5), Bruker korte setninger når han/hun skal svare på spørsmål (5).
- Vansker med å motta (73.33): Misforstår muntlige instruksjoner og beskjeder (4), Har problemer med å huske beskjeder (4), Misoppfatter situasjoner og hendelser (4).
- Vansker med å formidle (33.33): Tar sjelden språklige initiativ (4)

Case 80, gruppe 3 (jente på mellomtrinnet). Sumskåre: 60

- Semantiske vansker (70): Har vansker med å forstå hva vanlige ord betyr (4), Har problemer med å svare på spørsmål like raskt som andre (5), Leter ofte etter de riktige ordene (4), Har vansker med å gjenfortelle en historie han/hun har hørt (4).
- Vansker med å motta (73.33): Det virker som det han/hun lærer ikke fester seg (4), Har vansker med å huske ting (4), Har problemer med å huske beskjeder (4), Misoppfatter situasjoner og hendelser (4).
- Formidle (33.33)

**Konklusjon:** Vi ser av grafen over at barn med alvorlige språkvansker skiller seg tydelig fra barn med milde/moderate vansker og de normalspråklige barna ved å ha en høyere skåre på nesten alle utsagnene i spørreskjemaet. Når vi ser nærmere på barna med alvorlige vansker, oppdager vi at de har noe til felles: De har flest problemer av semantisk art, men de har også vansker innen de to andre kategoriene. Casene som er beskrevet viser skårene på 20 spørsmål for de seks barna som kan ha alvorlige språkvansker/spesifikke språkvansker ifølge lærerne som uttrykt ved 20 spørsmål. Alle seks barna har problemer med å svare på spørsmål like raskt som andre, fire av barna leter ofte etter de riktige ordene, tre av barna bruker ufullstendige setninger, fire av barna har vansker med å huske ting og fire av barna har vansker med å huske beskjeder.

Vi har nå sett hva som kjennetegner barna som kan ha SSV ifølge 20 spørsmål. I det følgende presenterer vi hvilke feil barna gjør på språktestene TROG og BPVS.

#### **4.2.9 Hva kjennetegner barn med mulig SSV ifølge TROG?**

**Case 62, gruppe 1,** (jente på mellomtrinnet): TROG (-1.28 SD).

Feil: I2: reversible over og under (koppen er under stjernen), K1: reversibel passiv (Kua blir jaget av jenta), M3: Pronomen - kjønn og tall (Hun peker på dem), O4: Verken eller (Verken jenta eller hunden sitter), S3: Relativsetning i objektet (Koppen er inni esken som er rød), T3: Innskutt relativsetning uten konjunksjon (Anda ballen ligger på, er gul).

**Case 76, gruppe 3,** (gutt på mellomtrinnet): TROG (-1.66 SD).

Feil: H1: Ikke bare x, men også y (blyanten er ikke bare lang, men også rød), L4 Nullanafor (Esken er inni koppen og er blå), O2: Verken eller (Verken blomsten eller skjerfet er langt), P3: x, men ikke y (Mannen, men ikke hesten hopper), R3: entall og flertall (Jentene står på stolen), S1: (Jenta løper etter hunden som hopper), S3: (Koppen er inni esken som er rød), S4: Relativsetning i objektet (Skjerfet ligger på blyanten som er blå), T1: (Sauen jenta ser på, løper), T3: (Anda ballen ligger på, er gul), T4: Innskutt relativsetning uten konjunksjon (Skjerfet boka ligger på, er blått)

**Konklusjon:** TROG identifiserer to elever på mellomtrinnet som barn med mulig SSV. Begge barna gjør feil på en av fire oppgaver som tester forståelsen av konstruksjoner med

”verken eller” (Eksempel: Verken jenta eller hunden sitter). Begge barna har også vansker med forståelsen av relativsetninger, men er ellers ikke like når det gjelder feiltyper.

#### **4.2.10 Hva kjennetegner barn med mulig SSV ifølge BPVS?**

**Case 76, gruppe 3** (gutt på mellomtrinnet): BPVS (-2.07).

Feil: Gjør feil i oppgavesettet for 8-9 år. b37 (tamburin), b43 (pelsdyr), b54 (begeistret), b55 (globus), b65 (snerre), b67 (belg), b69 (pedal), b74 (gevir), b75 (krets), b82 (reptil), b85 (arktisk), b86 (seilfly), b92 (isolasjon), b94 (komponere), b97 (parallell), b100 (staffeli), b101 (omfavne), b102 (redskap), b103 (kvartett), b104 (sitrusfrukt), b105 (indikator), b106 (bufe), b107 (søyle), b108 (timer). Har ikke klart flere oppgaver etter dette.

**Case 34, gruppe 1** (jente på småskoletrinnet): BPVS (-1.34 SD)

Feil: b37 (tamburin), b50 (tannlege), b57 (sump), b64 (intervjue), b65 (snerre), b67 (belg), b71 (balkong), b72 (forurensning), b74 (gevir), b75 (krets), b76 (kollisjon), b78 (applauderte), b79 (næringsrik), b81 (hodebunn), b82 (reptil), b83 (gjenoppliving), 84 (lenke).: Har ikke klart oppgaver etter dette.

**Case 12, gruppe 1** (jente på småskoletrinnet): BPVS (-1.40)

Feil: b37 (tamburin), b40 (teleskop), b41 (dryppe), b46 (grønnsak), b48 (væske), b50 (tannlege), b51 (røyskatt), b52 (trekke), b54 (begeistret), b55 (globus), b57 (sump), b61 (par), b64 (intervjue), b65 (snerre), b67 (belg), b69 (pedal), b70 (rovdyr), b74 (gevir), b75 (krets), b78 (applauderte), b79 (næringsrik), b80 (justerbar), b81 (hodebunn), b82 (reptil), b84 (lenke)

**Case 80, gruppe 3** (jente på mellomtrinnet): BPVS (-1.64)

b62 (komme), b63 (rørformet), b65 (snerre), b67 (belg), b72 (forurensning), b74 (gevir), b75 (krets), b77 (oppblåst), b78 (applauderte), b79 (næringsrik), b81 (hodebunn), b82 (reptil), b86 (seilfly), b94 (komponere), b100 (staffeli), b101 (omfavne), b102 (redskap), b103 (kvartett), b105 (indikator), b106 (bufe), b107 (søyle), b108 (timer).

**Konklusjon:** Det ser ut til at flere av barna har vansker med hverdagsord (som *hodebunn* og *tannlege*) og mer abstrakte ord (som *næringsrik*). Ett av barna må gå tilbake til oppgaver som er beregnet for yngre barn.

## Oppsummering av kapittel 4

Resultatene fra undersøkelsene viste at testene var tilnærmet normalfordelte, men at 20 spørsmål om språkferdigheter ikke var det. Spørreskjemaet 20 spørsmål tar utgangspunkt i barn som har vansker, og er derfor ikke ment å skulle være normalfordelt. Fordelingene legger føringer for hva slags analyser det har vært mulig å gjøre. Den manglende normalfordelingen til 20 spørsmål har lagt noen begrensninger for hvilke analyser som har kunnet foretas for å studere sammenhengen mellom spørreskjemaet og språktestene.

Det er foretatt reliabilitetsanalyse av 20 spørsmål om språkferdigheter i forbindelse med denne undersøkelsen, og testen viser seg å være svært reliabel. Det er ikke foretatt reliabilitetsanalyser av de andre måleinstrumentene i undersøkelsen, fordi disse har god dokumentasjon for reliabilitet fra før.

Korrelasjonsanalyser ble foretatt mellom språktestene TROG og BPVS og den nonverbale evnetesten Raven. Korrelasjonsmålet som ble brukt, var Pearsons  $r$ . Vi fant at det var signifikante sammenhenger mellom måleinstrumentene. Sammenhengen mellom både BPVS og TROG, og mellom TROG og Raven var signifikante, men styrken til korrelasjonene er relativt svak og forklarer lite av variansen. Det betyr at vi ikke kan trekke bastante slutninger om sammenhengen mellom variablene (McCauley, 2001). Korrelasjonen mellom TROG og BPVS er lav, mellom TROG og Raven er korrelasjonen lav/moderat og mellom BPVS og Raven er korrelasjonen nesten ubetydelig. Den sterkeste korrelasjonen er mellom TROG og Raven.

Ved hjelp av krysstabeller og en enkel deskriptiv analyse ble det avdekket at alle måleinstrumentene er enige om at hovedvekten av barna i utvalget er normalspråklige. Det er imidlertid et misforhold mellom hvilke barn som identifiseres som barn med mulig SSV av de ulike måleinstrumentene. Tabell 3 viste at to barn (2,5 %) har mulig SSV ifølge TROG og at fire barn (5 %) har mulig SSV ifølge BPVS. Tabell 4 viste at seks barn (7,5 %) har mulig SSV ifølge 20 spørsmål. Dette er et viktig funn som viser at flere barn identifiseres av 20 spørsmål enn av språktestene. BPVS fanger opp flere barn enn TROG. Blant barn som har mulig SSV, enes BPVS og TROG om ett, TROG og 20 spørsmål om ett og BPVS og 20 spørsmål om to barn. Beregning av prediktiv verdi for TROG og BPVS viste at testene hadde høy spesifisitet og lav sensitivitet. T-testen viste at de seks barna som lærerne angir har mulig

SSV, har svakere gjennomsnittsskåre på både BPVS og TROG enn barn som ifølge lærerne er normalspråklige.

I det følgende drøftes disse funnene.

# 5 DRØFTING

Formålet med denne undersøkelsen er å avdekke i hvilken grad TROG og BPVS identifiserer barn med SSV på en måte som er valid i forhold til lærernes beskrivelse av vanskene. For å bringe dette på det rene drøftes både kvantitative funn (i form av korrelasjoner, krysstabeller, t-test og beregning av prediktiv verdi) og kvalitative funn (i form av casestudier) i undersøkelsen. Kvantitative og kvalitative aspekter ved sammenhengen mellom de ulike måleinstrumentene som er benyttet, kan tenkes å gi indikasjoner om testenes begreps- og kriterievaliditet. Casestudiene benyttes også for å avgjøre hvorvidt et misforhold mellom testene og kriteriet skyldes at kriteriet kan sies å ha økologisk validitet, mens testene ikke har det. Om man kan stole på funnene i undersøkelsen, avhenger av om den er gjennomført på en valid og reliabel måte. Avslutningsvis kritiseres derfor forhold knyttet til operasjonalisering og gjennomføring av undersøkelsen som kan tenkes å ha påvirket validiteten og reliabiliteten til våre funn.

## 5.1 Drøfting av resultater i undersøkelsen

### 5.1.1 Måler TROG og BPVS alle aspekter ved begrepet SSV?

Dette spørsmålet handler om begrepsvaliditet. Begrepsvaliditet defineres som ”grad av samsvar mellom begrepet slik det er definert teoretisk og begrepet slik vi lykkes med å operasjonalisere det” (Kleven, 2002a, p. 176). Begrepsvaliditeten skal vurderes teoretisk og empirisk. Den teoretiske vurderingen av begrepsvaliditeten til TROG og BPVS gjøres i avsnitt 5.3 Drøfting av validiteten og reliabiliteten til undersøkelsen. I det følgende foretas en empirisk vurdering av begrepsvaliditeten til de to språktestene.

#### Empirisk vurdering av begrepsvaliditet

Første steg i den empiriske vurderingen består i å ta stilling til om vi har å gjøre med et homogent begrep. Dette kontrolleres ved en reliabilitetsmåling. I testmanualene til BPVS og TROG oppgis det at reliabiliteten til testene er svært høy (Bishop, 2009; Dunn, et al., 1997). Det antas at reliabiliteten til testene er høy også i denne undersøkelsen.



Andre steg i den empiriske vurderingen består i en ytre vurdering av hvordan måleresultatet oppfører seg i relasjon til målinger av andre begreper (Kleven, 2002a). Dette gjøres ved å vurdere den konvergerende validiteten (som undersøkes ved om resultatet korrelerer med andre målinger som det ut fra teorien kan tenkes å korrelere med) og den diskriminerende validiteten (som undersøkes ved å se om vårt måleresultat måler andre begrep som det ikke forventes å ha sammenheng med) (Kleven, 2002a). Dette gjøres i det følgende.

### **Konvergerende validitet: drøfting av korrelasjonen mellom TROG og BPVS**

TROG kan sies å være et mål for grammatisk forståelse og BPVS et mål for reseptivt vokabular. Siden vokabular og grammatikk er språkområder som barn med SSV ofte har problemer med, antas det at TROG og BPVS hver for seg kan identifisere barn med spesifikke språkvansker, og at testene derfor kan betraktes som mål på det samme underliggende begrepet. Vi forventet å finne en sammenheng mellom testene, fordi det i en annen studie er funnet høye korrelasjoner mellom TROG og BPVS (Bishop, 1989, ref. i Dunn, et al., 1997), og fordi disse språkområdene antas å representere ulike sider ved begrepet SSV. Dersom de to testene korrelerer rimelig høyt med hverandre, har de konvergerende validitet (Kleven, 2002a). Vi antar at korrelasjoner mellom de to måleinstrumentene kan gi indikasjoner på om dette stemmer.

Vi fant en signifikant korrelasjon mellom TROG og BPVS, men styrken på sammenhengen var svak. Det betyr at variansen på den ene testen forklarer lite av variansen på den andre. Det at det er en sammenheng mellom testene, indikerer at instrumentene måler ulike sider ved barns språkferdigheter. Det at sammenhengen mellom instrumentene bare er svak i vår undersøkelse, kan tolkes som at disse bare representerer en svært liten del av begrepet SSV, og at de to testene alene ikke kan brukes til å identifisere barn med SSV. Dette kan indikere at begrepet SSV er operasjonalisert på en utilstrekkelig måte i undersøkelsen, og redegjørelsen om begrepets kompleksitet i teoridelen støtter en slik tolkning. Drøftingen av dette spørsmålet vil bli utdypet i avsnitt 5.3.

Det kan for øvrig være flere årsaker enn dette til den svake korrelasjonen i vår undersøkelse. Et forhold som kan ha betydning, er alderen til barna i utvalget. Ved normeringen av norsk versjon av TROG, fant man at korrelasjonen mellom TROG og BPVS var høyest hos barn på 4-5 år (.57 og .53), og at den så gradvis avtok hos 7-år gamle barn (.45) og hos barn i aldersgruppen 8-16 år (.23). I testmanualen for norsk versjon av TROG forklares dette med at

den grammatiske kompetansen som måles med TROG, kan ha flatet ut i den høyeste aldersgruppen, samtidig som ordforrådet er i stadig utvikling (Bishop, 2009). Da barna i utvalget vårt befinner seg i aldersgruppen 7-11 år, kan dette forholdet tenkes å forklare den svake korrelasjonen også her.

### **Diskriminerende validitet: Drøfting av sammenhengen mellom Raven og språktestene**

Måler BPVS og TROG andre begreper som språktestene ikke forventes å ha sammenheng med? Som vi har sett i teoridelen er sammenhengen mellom intelligens og språkferdigheter omstridt. Dersom BPVS og TROG er rene språktester og Raven er et uttrykk for nonverbal intelligens, forventet vi at vi ikke skulle finne noen sterk sammenheng mellom Raven og språktestene.

Vi fant at korrelasjonen mellom BPVS og Raven er lav. Korrelasjonen mellom TROG og Raven kan betegnes som moderat, og det vil si at den er høyere enn korrelasjonen mellom TROG og BPVS. Det betyr at barnas skåre på Raven forklarer mer av variansen på TROG enn på BPVS. Dette er et overraskende funn, fordi vokabular i en annen undersøkelse (McCormick, 1994, ref. i Bishop, 2009) ble angitt å ha en ganske sterk forbindelse til kognitive ferdigheter. Den diskriminerende validiteten til BPVS ser ut til å være god i denne undersøkelsen, mens TROG har noe svakere diskriminerende validitet.

Det kan være flere ulike forklaringer på et slikt funn. BPVS og TROG har en enkel responsform der barna blir bedt om å peke på det bildet de tror er riktig ut fra det testleder spør om. Det kan føre til at det blir enklere å gjette, fordi man ikke behøver å gi noen forklaring. Det kan tenkes at barna lurer testen og bruker elimineringsmetoden når de skal angi riktig svar. Ifølge Bishop (2003) er imidlertid sannsynligheten for at en person svarer riktig på alle de fire oppgavene i en blokk ved en tilfeldighet minimal.

**Konklusjon:** Den empiriske vurderingen av testenes begrepsvaliditet tyder på at den er forholdsvis lav i vår undersøkelse. Kjernen i vurdering av begrepsvaliditet er imidlertid en undersøkelse av om det operasjonaliserte begrepet oppfører seg slik vi teoretisk skulle forvente. Hvis målingene består den prøven, brukes de til å validere våre teorier (Kleven, 2002a). Dette vil bli gjort i avsnitt 5.3 Drøfting av validiteten og reliabiliteten til undersøkelsen.

## 5.2 Gjør testene nøyaktige målinger?

Hovedspørsmålet i denne undersøkelsen er om TROG og BPVS gjør nøyaktige målinger. Dette spørsmålet handler om testenes kriterievaliditet, og er blitt undersøkt ved hjelp av krysstabeller, deskriptiv analyse og beregning av testenes spesifisitet og sensitivitet. Ved hjelp av krysstabellene fant vi ut at TROG og BPVS samsvarte godt med 20 spørsmål når det gjaldt å identifisere barn uten spesifikke språkvansker, men at det var misforhold mellom måleinstrumentene når det gjaldt hvilke barn som kan ha SSV. Det er naturlig at måleinstrumenter ikke samsvarer fullstendig, men det er et problem hvis tester som skal identifisere barn med spesifikke språkvansker, ikke fanger opp barn som oppfattes som barn med en alvorlig vanske i det daglige.

Hvor nøyaktig en test måler det den skal måle, angis vanligvis som testens sensitivitet og spesifisitet. Ved beregning av disse verdiene fant vi ut at testene var best til å identifisere alle individer uten vansker (det vil si spesifisitet), men var mindre gode til å identifisere de som hadde en mulig vanske (det vil si sensitivitet). BPVS har noe større sensitivitet enn TROG, og kan dermed anses for å være bedre til å identifisere hvilke barn i vårt utvalg som kan ha SSV, enn TROG når de to testene sammenlignes med kriteriet 20 spørsmål om språkferdigheter. Kravet til spesifisitet er innfridd for begge testene, men kravet til sensitivitet er langt fra innfridd. Dette skulle tilsi at de to språktestene ikke kan brukes alene i identifiseringen av SSV, noe som er i tråd med testmanualenes anbefalinger. Det er også i tråd med anbefalinger om å bruke flere tilnærminger ved kartlegging av barns språk (Bishop, 1997; Sattler, 2001).

Funnene over støttes av annen forskning på sammenhenger mellom testers og klinikers identifisering av barn med spesifikke språkvansker. Bishop & McDonald (2009) stiller i en studie spørsmål ved om tradisjonelle språktester mangler sensitivitet overfor klinisk viktige kjennetegn på språkvansker. Ifølge de to forskerne er det allment kjent at det kan være et ganske stort misforhold mellom hvilke barn med SSV som identifiseres ved hjelp av språktester, og hvilke barn som identifiseres klinisk. De mener at dette hovedsakelig kan skyldes to forhold: For det første har språktester varierende sensitivitet overfor klinisk viktige språkproblemer, for det andre fant de at barn av lav sosial status ikke like lett ble fanget opp av hjelpeapparatet, men at språktestene fanget opp at de hadde vansker. Som vi har sett i avsnitt 2.4 Forskning på testers sensitivitet og spesifisitet, har også flere andre studier vist at

tester ikke alltid har fullkommen sensitivitet og spesifisitet (Aram, et al., 1992; Gray, et al., 1999; Keegstra, et al., 2007).

Det er fire mulige utfall av en test målt mot et kriteriemål: sann positiv, falsk positiv, falsk negativ og sann negativ (som beskrevet i avsnitt 2.4 Forskning på testers sensitivitet og spesifisitet). I det følgende presenteres tre ulike case som illustrerer hvilke utfall testing med BPVS og TROG har gitt i vår undersøkelse.

### 5.2.1 Casestudier som viser ulike utfall av en test

#### a) Case 76. Sann positiv: testen plasserer barnet i risikogruppen, og dette samsvarer med kriterimålet.

Dette er en gutt på mellomtrinnet som med stor sannsynlighet har spesifikke språkvansker, fordi alle måleinstrumentene angir at dette barnet kan ha SSV. Barnet har en veldig svak skåre på BPVS (-2 SD), og ville dermed også ha blitt fanget opp av de litt strengere ICD-10 kriteriene hvis vi hadde brukt dem. Dette er det eneste barnet som skårer mer enn 2 SD under gjennomsnittet på en av testene.

Vokabular: Barnet ser ut til å ha en lavere språkaldre enn kronologisk alder, fordi han må reversere (gå tilbake til enklere oppgaver) pga. for mange feil i oppgavesettet for sin egen aldersgruppe. Han gjør også feil i oppgavesettet for den yngre aldersgruppen. Barn med SSV utvikler ordforrådet saktere enn sine jevnaldrende (Bishop, 1997; Leonard, 2000). Det kan være det vi ser et eksempel på her. I spørreskjemaet angir lærerne at barnet gjør flest feil innen kategoriene Semantikk og Vansker med å motta et budskap. Læreren antyder blant annet at gutten glemmer ord han vet hva betyr og blander sammen ord med lik mening samt at han har vansker med å forstå hva vanlige ord betyr. Dette stemmer godt overens med de typer feil gutten gjør på BPVS, der han har vansker med å forstå dagligdagse ord som omfavne, pelsdyr, pedal, søyle, begeistret og timer. Problemer av denne typen kan skyldes at eleven har svakere og færre semantiske assosiasjoner og mindre effektiv og nøyaktig gjenkalling av ord fra langtidsminnet. Dette er et trekk som ikke er uvanlige hos barn som kan ha SSV (Bishop, 1997; Leonard, 2000). På Språk 6-16 har barnet en svært lav skåre på deltesten Begreper (2). Deltesten Begreper undersøker det semantiske aspektet ved språket, det vil si barns kunnskap om ordenes betydning. Dette er en ekspressiv test som går ut på å forklare betydningen av ord. Testene BPVS og Språk 6-16 indikerer altså at barnet kan ha problemer med både

reseptivt og ekspressivt vokabular. Dette stemmer med Bishops (1997) oppfatning av at barn med ekspressive vansker ofte også har reseptive vansker.

Grammatisk forståelse: Feilene barnet gjør på TROG (-1.66) viser at barnet har størst vansker med å forstå relativsetninger. Det kan være flere årsaker til den svake skåren på TROG. En forklaring kan være at barnet har vansker med å forstå grammatiske strukturer, og denne tolkningen er trolig ofte nærliggende i forbindelse med lave skårer på TROG. Svake skårer på TROG kan imidlertid skyldes andre forhold også. Barnet kan ha problemer med å fastholde flere informasjonselementer i en setning, og dette kan være en forklaring på feilene i relativsetningene, som nettopp inneholder flere informasjonselementer. På Språk 6-16 har barnet en svak skåre på deltesten Ordspenn (5), og har også en skåre under normalområdet på deltesten Setningsminne (6). Dette stemmer godt overens med lærerens angivelse av at eleven har vansker med å huske ting og beskjeder. En tredje forklaringsmulighet er at barnet ikke forstår ordene i setningene. Dette harmonerer med lærerens beskrivelse av at gutten har vansker med å forstå hva vanlige ord betyr. TROG har imidlertid et ordforråd som er mye enklere enn BPVS, og som burde mestres allerede i førskolealder.

Fonologi: Læreren angir at barnet har problemer med å uttale ord riktig. Dette støtter at barn med språklydvansker ofte har svakheter også på flere andre språkområder. Det er derfor blitt stilt spørsmål ved om fonologiske vansker kan betraktes som en egen undergruppe av SSV, slik det gjøres i diagnosemanualen ICD-10 (Leonard, 2000).

Pragmatiske vansker: Læreren angir at barnet misoppfatter situasjoner og hendelser, og har problemer med å gi uttrykk for behov. Dette kan være tegn på pragmatiske vansker. Med pragmatiske vansker mener vi vansker med språket i bruk og tolkning av språklige ytringer i forskjellige situasjoner (Sveen, 2005). Barn som har ordforråd, språkforståelse og grammatisk kunnskap som ikke er aldersadekvat, kan streve med å bruke språket i sosialt samspill med andre (Rice, et al., 2005).

**Konklusjon:** Som vi har sett plasserer kriteriemålet 20 spørsmål om språkferdigheter barnet over i risikogruppen for spesifikke språkvansker, og både TROG og BPVS samsvarer med kriteriemålet. Dette er et utfall på testen som kan karakteriseres som sant positivt.

**b) Case 12. Falsk positiv: Barnet har ingen vanske, men en av testene viser at barnet har en vanske. Dette fører til overrapportering.**

Dette er en jente på småskoletrinnet som ifølge lærerens angivelser i 20 spørsmål ikke har noen tegn til å ha en vanske. På TROG (-.08) skårer hun også innen normalområdet.

Vokabular: BPVS gir en skåre på -1.4 SD. Dette er litt under cut-off-grensen vi har angitt for skåre på en språktest med tanke på identifisering av SSV. BPVS angir at barnet har problemer med hverdagsord som hodebunn, tannlege, grønnsak, dryppe, begeistret, pedal og lenke. Som vi har sett av beskrivelsen av det forrige caset, kan dette være et tegn på en mer alvorlig vanske, fordi det kan se ut til at barn med SSV utvikler ordforrådet saktere enn sine jevnaldrende (Bishop, 1997; Leonard, 2000). Eksemplene på ord som barnet har vansker med forståelsen av, er innholdsord som substantiv og verb, og slike ord kan være vanskeligere å lære enn funksjonsord som adjektiv og preposisjoner. Barn med SSV skårer ofte lavere enn sine jevnaldrende på ordforrådstester (Rice, 2004). Det kan stilles spørsmål ved sannsynligheten for at dette barnet kan ha SSV. Når læreren ikke gjør noen anmerkninger om dette barnet, fungerer det trolig godt i hverdagen. En annen mulighet er at barnet har en reseptiv vanske som ikke fanges opp av læreren. Det foreligger ikke data fra Språk 6-16 som kunne gi informasjon om dette barnets ekspressive vokabular.

**Konklusjon:** Som vi har sett, angir kriteriemålet 20 spørsmål om språkferdigheter at barnet ikke har en vanske, mens BPVS indikerer at hun har en vanske. Dette fører til overrapportering, og kan betegnes som et falskt positivt utfall på BPVS. (Utfallet på TROG kan betegnes som sant negativt.)

**c) Case 77. Falsk negativ: Barnet har en vanske, men testene fanger ikke dette opp. Dette fører til underrapportering.**

Dette er en jente på mellomtrinnet som har mulige spesifikke språkvansker ifølge 20 spørsmål (54), men som ikke fanges opp som et barn med SSV av TROG (-.39) og BPVS (-.53).

Lærerens vurdering tilsier at barnet kan ha vansker innen alle tre kategoriene på 20 spørsmål. Vanskene er størst innen kategoriene Semantiske vansker og Vansker med å formidle et budskap. Utsagnene som har høyest skåre, er at eleven har problemer med å svare på spørsmål like raskt som andre og sjelden tar språklige initiativ. Læreren fanger her opp forhold som ikke så lett kan fanges opp av en test. Ifølge Språk 6-16 har eleven en svak skåre

på deltestene Begreper (5) og Setningsminne (4), men skårer innen normalområdet på deltesten Ordspenn (7). Da deltesten Begreper er en ekspressiv test og kategorien Vansker med å formidle et budskap inneholder utsagn om ekspressivt språk, kan det tenkes at barnet først og fremst har vansker med ekspressivt språk, men har relativt god forståelse av språk. Det at BPVS og TROG ikke angir at barnet har vansker, tyder på at dette kan være en riktig tolkning.

**Konklusjon:** Som vi har sett, viser kriteriemålet 20 spørsmål om språkferdigheter at barnet har en vanske, mens ingen av språktestene fanger dette opp. Dette fører til at språktestene underrapporterer vansken, og kan betegnes som et falskt negativt utfall på begge testene.

d) **Sann negativ: Barnet har ingen vanske, og testene angir at barnet ikke har noen vanske.**

Som vi har sett i kapittel 4 (avsnitt 4.2.4) angir kriteriemålet 20 spørsmål om språkferdigheter i de fleste tilfeller at barna i utvalget ikke har en spesifikk språkvanke. I de aller fleste tilfellene samsvarer språktestene med kriteriemålet om dette. Dette utfallet på testene karakteriseres som sant negativt.

Som beskrevet over er det et misforhold mellom de ulike måleinstrumentene når det gjelder identifisering av barn som kan ha SSV. Ovenfor har vi bare gitt én casebeskrivelse for tre av de fire mulige utfallene av en test. Den siste kategorien (sann negativ) har vi ikke gitt noen casebeskrivelse av, fordi dette er barn som måleinstrumentene enes om ikke har noen vanske. I undersøkelsen vår har vi flere barn innen hver av disse kategoriene, slik vi har beskrevet i avsnitt 4.2.5 Prediktiv verdi.

Målet for diagnostisering bør være å etablere kriterier som reduserer antall barn i kategoriene falsk negativ og falsk positiv, og øker antall barn i kategoriene sann positiv og sann negativ. Det vil si at man øker sensitiviteten og spesifisiteten (Bishop & McDonald, 2009). Dette kan gjøres ved å vurdere hvilke cut-off-grenser som er best egnet til å identifisere vansken på hvert enkelt måleinstrument. Dersom grensen settes nær det som regnes som normalområdet på en test, kan det føre til overidentifisering, men hvis den settes langt fra normalområdet, kan det føre til underidentifisering. Av de fire mulige utfallene av en test regner vi underrapportering for å være det som kan få størst negative konsekvenser for barna, fordi underrapportering kan medføre at vansken ikke oppdages, og at barna dermed heller ikke får

den hjelpen de trenger. Overrapportering av vansker kan på den annen siden skape unødig bekymring hos dem det gjelder, men sannsynligvis vil ulempene ved underidentifisering kunne gi større skade. I vurderingen av hvor cut-off-grensen skal settes, bør denne avveiningen gjøres (Rydz, et al., 2005).

### 5.2.2 Drøfting av forhold som kan føre til underrapportering

På grunnlag av casene over er det mulig å gjøre noen antakelser om hvorfor språktestene i fire tilfeller underrapporterer vansker i forhold til kriteriet 20 spørsmål. Når vi skal svare på problemstillingen er det viktig å ta hensyn til slike forhold, fordi de kan gi kunnskap om hvorvidt underrapporteringen skyldes forhold knyttet til måleinstrumentene, eller om den skyldes forhold knyttet til vår operasjonalisering av begrepet SSV i undersøkelsen. Noen av disse antakelsene er blitt drøftet i forbindelse med hvert case, andre vil bli drøftet i fortsettelsen.

**Forhold knyttet til økologisk validitet:** Drøftingen av case over viser at spørreskjemaet 20 spørsmål om språkferdigheter fanger opp andre aspekter ved barnas språklige fungering enn det språktestene gjør. Spørreskjemaet kan sies å være mer økologisk valid enn språktestene, fordi det tar utgangspunkt i barnas fungering i naturlige situasjoner og over tid, mens testene bare gir et øyeblikksbilde av barnas prestasjon i en kunstig situasjon. Forhold ved den kunstige testsituasjonen kan gjøre at barnas yteevne enten blir dårligere eller bedre enn den faktisk er i skolesituasjonen.

Noen av utsagnene som er beskrevet i casene over (det at barn misoppfatter situasjoner og hendelser, har problemer med å uttrykke behov, og sjelden tar språklig initiativ) kan være et uttrykk for pragmatiske vansker, slik de er beskrevet i avsnitt 2.1.4 Lingvistiske kjennetegn hos barn med spesifikke språkvansker. Pragmatiske vansker fanges ikke så lett opp ved bruk av standardiserte tester, fordi de har å gjøre med hvordan barnet fungerer i samspill med andre. Hos barn med SSV blir de pragmatiske vanskene ofte sekundære i forhold til vansker disse barna har med semantikk og morfologi (Rice, et al., 2005). Da vi ikke har innhentet informasjon om tilleggsvansker for barna i utvalget vårt, kan vi ikke vite sikkert om eventuelle pragmatiske vansker er sekundære eller primære. Dette er et validitetsproblem for undersøkelsen.



**Forhold knyttet til testenes teoretiske utgangspunkt:** Tester som brukes til kartlegging av språklige vansker og ferdigheter, er ofte teoribaserte. Ulike tester har ulikt teoretisk utgangspunkt. TROG og BPVS er tester som trolig har sitt utspring i Brown og Hulmes teori (1996) om at vansker med språklig produksjon og persepsjon vil kunne føre til problemer med verbalt kortidsminne og språktilegnelse. Ifølge denne teorien kan ikke en spesifikk test pekes ut som en god markør, men tester på grunnleggende språklige funksjoner som semantikk og syntaks kan brukes.

Beregningen av sensitiviteten til TROG og BPVS viser at de ikke er godt egnet til å fange opp barn med mulig SSV i utvalget vårt. Ifølge Ottem & Lian (2008b) fungerer ikke spesifikke tester så godt som markør på SSV når de brukes på eldre barn. Det er mulig at dette forholdet delvis kan forklare underrapporteringen til språktestene i vår undersøkelse. Vi ser at deltestene til Språk 6-16 i flere tilfeller fanger opp vansker som ikke TROG og BPVS fanger opp.

Kan det tenkes at tester som baseres på fonologisk korttidsminne er bedre markører for SSV, enn spesifikke tester? I en studie fant Bishop & McDonald (2009) at blant annet vokabularoppgaver var mindre effektive enn verbalt korttidsminne til å skille klinisk henviste tilfeller fra andre tilfeller. Språk 6-16, som brukes som supplerende måleinstrument i denne undersøkelsen, er basert på teorien om at det er en sammenheng mellom språklig korttidsminne og evnen til å danne språklige begreper (Baddeley, 2003; Baddeley, et al., 1998). Siden ikke alle barna i utvalget er testet med Språk 6-16, kan vi ikke vite om den ville ha identifisert flere barn med SSV enn TROG og BPVS, men i tilfeller der data fra Språk 6-16 foreligger, ser den ut til å samsvare ganske godt med kriteriemålet 20 spørsmål. En mulig forklaring på at deltester i Språk 6-16 ser ut til å samsvare godt med 20 spørsmål om språkferdigheter, kan i tråd med Bishops funn være at fonologisk minne er en bedre markør for SSV enn spesifikke språktester.

En annen mulig forklaring på at Språk 6-16 og 20 spørsmål om språkferdigheter ser ut til å samsvare, kan være at begge favner et bredt spekter av ferdigheter. Selv om det ene instrumentet er en test og det andre et spørreskjema, har de det til felles at hukommelse og semantikk inngår som komponenter, og at skårene baseres på barns ekspressive språk. Dette kan også være en mulig forklaring på samsvaret.

Er grammatikk en god markør for SSV hos eldre barn? Vi har tidligere referert til en studie som fant at utviklingen av grammatiske ferdigheter hos normalspråklige barn avtar med økende alder (Bishop, 2009). Med det menes at barn tidlig mestrer de fleste grammatiske strukturer. På denne bakgrunn kan det stilles spørsmål ved om grammatisk forståelse som uttrykt ved TROG kan være en god markør for spesifikke språkvansker i utvalget vårt, som hovedsakelig består av barn som er sju år og eldre. Dette spørsmålet er i sin tur forbundet med spørsmålet om hvorvidt barn med SSV bare har forsinket språkutvikling, eller om de har selektive vansker. Selv om barn med SSV har forsinket utvikling i forhold til normalspråklige barn, vil de på enkelte områder aldri ta igjen normalspråklige barn (Leonard, 2000). Barn med spesifikke språkvansker har større problemer med morfologi enn normalspråklige barn, og studier peker på at barn med SSV som gruppe ikke alltid tar igjen sine jevnaldrende på dette området (Rice, 2004). Av den grunn skulle man tro at TROG var en god markør for SSV også for eldre barn. Ifølge Leonard (2000) tyder testprofilene til barn med SSV på at de har bedre forståelse av syntaktiske strukturer enn evne til å uttrykke slike strukturer i ekspressivt språk. Det er derfor mulig at vi ville ha identifisert flere barn med SSV dersom vi hadde benyttet et måleinstrument som testet barnas ekspressive grammatikkferdigheter. En annen mulig forklaring på at TROG ikke identifiserer like mange barn som BPVS, kan være at SSV er en svært kompleks vanske. Alle barn har ikke vansker med grammatisk forståelse, selv om de kan ha vansker på andre områder.

**Forhold knyttet til at ikke begge språkmodalitetene testes:** TROG og BPVS er reseptive språktester, og kan brukes til å kartlegge henholdsvis grammatisk forståelse og reseptivt vokabular. Beskrivelsen av begrepet SSV i kapittel 2 viser at det er et sammensatt begrep. TROG og BPVS kan sies å måle aspekter ved dette begrepet. Tyder våre funn på at de to reseptive språktestene alene kan brukes til diagnostisering av SSV? Som vi har sett i teoridelen er det fortsatt uenighet om hvordan SSV skal defineres, og det fører igjen til ulik operasjonalisering i forskning. I de diagnostiske kriteriene for forskning som er utgitt i forbindelse med ICD-10, angis ikke hvilken språktest som skal brukes til å identifisere SSV, selv om språktester kan ha ulik sensitivitet (Spaulding, et al., 2006). Dersom det er slik at barn med SSV kan ha en ren ekspressiv språkvanske, slik diagnosemanualen ICD-10 angir, er det en svakhet at det kun er brukt reseptive språktester i forbindelse med denne undersøkelsen. Ifølge ICD-10 innebærer en ren ekspressiv vanske at barns evne til å uttrykke seg verbalt er betydelig under det som er ventet på grunnlag av barnets mentale alder, mens språkforståelsen er normal (WHO, 2000). Ifølge Bishop (1997) vil imidlertid barn med en ekspressiv vanske

ofte også ha en reseptiv vanske som fanges opp dersom testene er sensitive nok. Hun hevder for øvrig også at foreldre og lærere lett oppdager kommunikasjonsvansker som medfører redusert forståelighet eller setningsstrukturer som høres ukorrekte ut, men ikke så lett fanger opp språkvansker som påvirker forståelse, ordforråd eller verbalt minne. Dersom Bishops argumentasjon legges til grunn, kan det forsvares at det bare er brukt reseptive språktester til identifisering av spesifikke språkvansker. En nærmere analyse av casene over viser imidlertid at situasjonen kanskje er mer sammensatt enn det de to påstandene kan gi inntrykk av:

I case 76 gis et eksempel på en vanske som er svært sammensatt. Måleinstrumentene indikerer at barnet har vansker med både ekspressivt og reseptivt vokabular samt grammatisk forståelse. Dette caset støtter derfor Bishops påstand om at ekspressive vansker ofte ledsages av en reseptiv vanske, og de reseptive testene fanger den opp. Case 77 er imidlertid ikke i samsvar med denne påstanden. I dette tilfellet angir 20 spørsmål om språkferdigheter i tillegg til deltestene Begreper og Setningsminne i Språk 6-16 at barnet har en vanske med ekspressivt språk, mens gode skårer på TROG og BPVS tyder på at barnet ikke har vansker med forståelsen av vokabular og grammatiske strukturer. Her støttes diagnosemanualens påstand om at en ekspressiv vanske kan opptre alene. Når kun reseptive tester brukes til identifisering av SSV, medfører det at vansken underrapporteres. Dette er et overraskende funn, fordi det virker logisk at man ofte forstår mer enn man kan gi uttrykk for i ekspressivt språk. Dersom deltesten Begreper bare tapper barns kunnskap om ordenes betydning, er det for øvrig overraskende at BPVS ikke fanger opp det samme. TROG og BPVS tester barns forståelse uten å kreve bruk av verbalt språk. På den måten tapper de trolig barns forståelse på en bedre måte enn deltesten Begreper i Språk 6-16, der ikke bare barnets forståelse, men også evnen til å uttrykke seg avgjør om det gis en god skåre. Den lave skåren på deltesten Setningsminne (som antas å være et mål på språklig minne og begrepsutvikling) kan kanskje stå i forbindelse med en ekspressiv vanske. I dette tilfellet ser vi trolig et eksempel på en vanske som ikke kan identifiseres ved en reseptiv test, men som fanges opp ved en ekspressiv test som Språk 6-16.

**Forhold knyttet til at testene er egnet til å skille mellom grupper med svake og sterke språkferdigheter, men ikke til å identifisere SSV:** Det er foretatt en t-test for uavhengige utvalg for å sammenligne gjennomsnittsskåren på TROG og BPVS hos barna som kan ha SSV og barna som ikke har SSV ifølge 20 spørsmål om språkferdigheter. T-testen viste at barn med mulig SSV skårer signifikant svakere enn de normalspråklige barna på både TROG og BPVS, men gjennomsnittsskårene er innenfor normalområdet for begge gruppene. Dette

funnet er i samsvar med en studie av den diagnostiske nøyaktigheten til fire vokabulartester som konkluderte med at det er viktig at statistisk evidens for validitet stemmer overens med det som er formålet med kartlegging av språkvansker (Gray, et al., 1999). Selv om barna med SSV som gruppe hadde lavere skårer enn normalspråklige barn på fire ulike vokabulartester, falt skårene innenfor det som vanligvis regnes for å være normalområdet. I studien hevdes det at dokumentasjon for begrepsvaliditet er sentralt hvis hensikten med en test er å beskrive språkferdigheter, men at det ikke betyr at testen egner seg til identifisering av barn med SSV. I studien hadde vokabulartestene evidens for samtidig validitet og diskriminerende validitet, men ingen av testene var godt egnet til å identifisere barn med SSV i utvalget (Gray, et al., 1999). I vår studie har testene belegg for kriterievaliditet ifølge testmanualene, men det vises ikke til studier som har undersøkt testenenes evne til å identifisere barn med SSV.

**Forhold knyttet til måleinstrumentenes omfang:** Når årsaker til ulikheter mellom måleinstrumentene skal analyseres, er en nærliggende tolkning at 20 spørsmål om språkferdigheter identifiserer flere barn med mulig SSV, fordi spørreskjemaet gir en bredere kartlegging enn hva språktestene gjør. 20 spørsmål tar ikke bare utgangspunkt i barns fungering i naturlige situasjoner og over tid, men dekker også et bredt spekter av barns symptomer på språklige vansker, i motsetning til TROG og BPVS, som måler et forholdsvis spesifikt område av barns språk. Når testene og 20 spørsmål skal skåres, summeres antall feil og anmerkninger til en sumskåre. En sumskåre over en valgt grense angir at barn er i en risikogruppe for språkvansker. Siden TROG og BPVS tester hvert sitt spesifikke område, skal vansken innen akkurat dette området være relativt stor dersom den skal gi utslag på språktestene. Siden 20 spørsmål favner et bredt område, trenger ikke barnet ha så store vansker innen hvert område for at det skal føre til en høy skåre og dermed indikere en vanske. Dette kan være en årsak til at 20 spørsmål fanger opp flere barn som kan ha SSV enn BPVS og TROG. For å illustrere dette kan vi se nærmere på profilmønsteret til de seks barna som har mulig SSV ifølge lærerne.

Profilmønsteret for barn som har mulig SSV ifølge lærerne som uttrykt ved 20 spørsmål (Figur 4, avsnitt 4.2.8) viser at de seks barna med en mulig spesifikk språkvanske i utvalget vårt har størst vansker med å svare på spørsmål like raskt som andre, ofte leter etter de riktige ordene, bruker ufullstendige setninger, bruker korte setninger når de skal svare på spørsmål og har vansker med å gjenfortelle en historie de har hørt. Barna med mulig SSV i utvalget vårt har også store vansker med å huske ting og beskjeder og med å ta språklig initiativ. Alle

vanskene som er beskrevet over, har det felles at de er lettest å observere i ekspressivt språk. Flere av dem ville trolig kunne fanges opp av ekspressive språktester av ulike typer, mens andre trolig er lettere å oppdage ved observasjon av barnet i samhandling med andre. Det at barna har vansker med å svare på spørsmål like raskt som andre, ville muligens ha blitt fanget opp i testsituasjonen dersom testleder hadde registrert hvor lang tid barna trengte for å avgi svar, men det er ikke blitt gjort i forbindelse med denne studien. En mulig årsak til misforholdet mellom tester og lærevurderinger kan være at flere av vanskene som angis i 20 spørsmål ikke fanges opp i en testsituasjon, og i hvert fall ikke av en reseptiv språktest.

### 5.2.3 Drøfting av forhold som kan føre til overrapportering

**Forhold knyttet til spørreskjemaet:** Ifølge Bishop (1997) er det vanlig at lærere ikke så lett fanger opp vansker som opptrer i reseptivt språk. Selv om de i spørreskjemaet kommenterer vansker som hører innunder kategorien semantiske vansker og vansker med å motta et budskap, baseres lærernes vurdering kun på barnas ekspressive språk. Det er ikke mulig for andre enn barna selv å vite helt sikkert hva de forstår. Det kan derfor tenkes at barna som fanges opp som barn med SSV av testene, men ikke av lærerne, faktisk har en språkvanske. Til støtte for en slik tolkning er det faktum at to av de tre barna som bare fanges opp av språktestene, er i tidlig skolealder. Det kan være vanskelig å skille mellom hva som er umodenhet og forsinket språkutvikling hos de yngste barna. Det foreligger ikke informasjon om hvor lenge læreren har kjent barna, men læreren har nødvendigvis dårligere kjennskap til de barna som har gått kort tid på skolen. En annen årsaksforklaring kan være at barn med språkvansker over tid kan utvikle adferdsmessige, sosiale, skolerelaterte og psykiske problemer (Hughes, et al., 2009; Ottem & Lian, 2008a), og dette kan gjøre at vansken blir mer synlig jo eldre barna blir.

Overrapporteringen kan også skyldes forhold knyttet til operasjonaliseringen av begrepet SSV i undersøkelsen. Dette drøftes i avsnitt 5.3 Drøfting av validiteten og reliabiliteten til undersøkelsen.

**Konklusjon:** Språktestene TROG og BPVS lykkes i å identifisere barn som har vansker med reseptivt vokabular og grammatisk forståelse, men fanger ikke opp barn med vansker som tilsynelatende er mest fremtredende i ekspressivt språk og språklig samhandling med andre. Årsaken til at TROG og BPVS ser ut til å underrapportere vansker som fanges opp av lærerne ved hjelp av spørreskjemaet, kan være at testene kun undersøker et begrenset område av barns

språkferdigheter. Barn med SSV er en heterogen gruppe, og har vansker som manifesterer seg svært forskjellig. Identifisering av spesifikke språkvansker krever derfor kartlegging med både reseptive og ekspressive tester i tillegg til observasjon av barnets fungering i det daglige.

## **5.3 Drøfting av validiteten og reliabiliteten til undersøkelsen**

Enhver undersøkelse har sine begrensninger, og det vil alltid være mulighet til forbedring. Hvis det er forhold ved undersøkelsen som kan svekke validiteten og reliabiliteten, er det viktig å være bevisst på dette, og ta slike forbehold med i vurderingen av funn i undersøkelsen.

### **5.3.1 Drøfting av statistisk validitet**

God statistisk validitet er en nødvendig betingelse for indre validitet, ytre validitet og begrepsvaliditet. For å oppnå dette forutsettes det at sammenhengen eller tendensen er statistisk signifikant og rimelig sterk. Det vil være forskjellig hva som regnes som ”rimelig sterk” innen det enkelte forskningsområdet (Kleven, 2002b). Innenfor samfunnsvitenskap kan man ikke forvente de helt høye korrelasjonene, både fordi det ikke er noen eksakt vitenskap og fordi det innenfor dette feltet ofte er mange variabler som virker sammen. Flere av disse kan være ukjente.

I vår undersøkelse ble det foretatt en korrelasjonsanalyse mellom TROG, BPVS og Raven. Alle korrelasjonene var signifikante og hadde lav til moderat styrke. Manglende styrke kan true den statistiske validiteten og dermed også de andre validitetstypene. Når det er sagt, kan korrelasjonenes lave styrke skyldes at vi hadde et lite utvalg med liten varians. Lav statistisk styrke er en trussel mot god statistisk validitet som øker sannsynligheten for å gjøre en Type I-feil, dvs. faren for å forkaste en sann nullhypotese (Lund, 2002b).

Det ble også foretatt en t-test for uavhengige utvalg som viste at de seks barna som lærerne angir har mulig SSV, har en signifikant svakere gjennomsnittsskåre på både TROG og BPVS, enn barn som ifølge lærerne er normalspråklige. Alle skårene er imidlertid innenfor normalområdet. Det at det er en signifikant forskjell, viser at vi kan stole på funnet om at TROG og BPVS skiller mellom barn med sterke og svake språkferdigheter, men at de ikke i samme grad som 20 spørsmål identifiserer barn med mulig SSV.

Det er en svakhet ved undersøkelsen vår at vi ikke har kunnet foreta korrelasjoner mellom de variablene som er viktigst for problemstillingen vår, det vil si mellom TROG og BPVS på en side og 20 spørsmål på en annen side. Ved hjelp av utregning av språktestenes spesifisitet og sensitivitet i forhold til kriteriet 20 spørsmål har vi likevel fått presis informasjon om forholdet mellom måleinstrumentene i vår undersøkelse.

### **5.3.2 Drøfting av begrepsvaliditeten**

Den empiriske vurderingen av begrepsvaliditeten til testene TROG og BPVS ble gjort i avsnitt 5.1.1. I det følgende foretas en teoretisk vurdering av begrepsvaliditet.

Det er blitt forsket mye på SSV som fenomen, men forskerne strides fortsatt om hvordan begrepet skal defineres. Barn med spesifikke språkvansker kan ha ulike typer vansker, og det er liten konsensus om hva som kjennetegner barn med SSV (Conti-Ramsden, et al., 1997). Det finnes ingen universell definisjon av SSV som forskerne er enige om (Aram, et al., 1992). Begrepet SSV er et abstrakt begrep, og begreper av denne typen kan ikke operasjonaliseres helt eksakt. Måten begrepet operasjonaliseres på får konsekvenser for hvilke og hvor mange barn som identifiseres.

### **Diskusjon av eksklusjons- og inklusjonskriteriet på tester i vår undersøkelse**

Begrepet SSV operasjonaliseres ved å måle barnets reseptive grammatiske forståelse og reseptive vokabular ved hjelp av henholdsvis TROG og BPVS. Ved identifisering av barn med SSV anbefales bruk av standardiserte tester med gode psykometriske egenskaper (Bishop, 1997), og de to testene i vår undersøkelse må kunne sies å innfri dette kravet. Ifølge ICD-10-kriteriene kan barn med SSV identifiseres ved å bruke en ekspressiv eller resektiv språktest, og det gis ingen retningslinjer for hvilken test som skal brukes, selv om språktester kan ha ulik sensitivitet (Spaulding, et al., 2006). Bishop (1997) anbefaler imidlertid bruk av måleinstrumenter som tester både ekspressive og reseptive språkfunksjoner. Som drøftingen av testenes underrapportering viser, kan manglende bruk av ekspressive tester i undersøkelsen være en årsak til at noen barn med mulig SSV ikke har blitt identifisert av testene. Det vil si at begrepsvaliditeten i undersøkelsen er svak som følge av at det abstrakte begrepet er underrepresentert.

Det er for øvrig knyttet andre problemer til operasjonalisering av begrepet SSV også, og disse har med bruken av uvaliderte cut-off-grenser å gjøre. Den varierende begrepsbruken og uenigheten om hvordan man karakteriserer barn med SSV, fører til ulik operasjonalisering av begrepet. Som vi har sett er det stor variasjon i hvilke cut-off-grenser for språkferdigheter og kognitive ferdigheter som brukes i ulike studier. Det finnes ingen gullstandard for diagnosen SSV (Bishop & McDonald, 2009) og cut-off-grensene som benyttes i forskning er ofte uvaliderte (Spaulding, et al., 2006).

Noen forskere foreslår at det i stedet for vilkårlige cut-off-skårer på språktester i stedet bør brukes evidensbaserte cut-off-skårer som er utledet for hver test (Spaulding, et al., 2006). Den empiriske cut-off-skåren skulle reflektere testenes sensitivitet og spesifisitet, slik at klinikeren har evidens for en tests evne til å skille mellom barn med og uten språkvansker (Spaulding, et al., 2006). Det ser ut til å være en tendens i nyere forskning til å benytte grenser som befinner seg i nærheten av normalområdet. Denne operasjonaliseringen av begrepet stemmer med en antakelse som er utbredt hos mange forskere (f.eks Leonard, 1991) om at barn med SSV befinner seg i forlengelsen av normalområdet, men det kan ikke fastslås med sikkerhet at denne antakelsen er riktig (Spaulding, et al., 2006). I vår undersøkelse har vi brukt cut-off-skårer nær normalområdet, og som vi har sett identifiserer språktestene tre barn med mulig SSV som ikke identifiseres ved bruk av spørreskjemaet. Dette har vi vurdert som et falskt positivt utfall av en test (det vil si overrapportering). Alle barna som testene overrapporterer som barn med SSV, har imidlertid skårer tett opptil normalområdet. Dersom vi hadde brukt cut-off-grensen som anbefales i ICD-10 i undersøkelsen vår, ville ingen av språktestene ha fanget opp vansker hos disse barna, og utfallet av begge testene ville ha blitt sant negativt. I ettertid ser vi at dersom grensen hadde vært satt til en skåre på  $-1,5$  SD på en språktest i denne undersøkelsen, ville begge testene ha fått fullkommen spesifisitet. Dette er imidlertid ikke nødvendigvis et mål i seg selv, fordi testene allerede har en spesifisitet som ligger langt over minstekravet. Dersom testene skal brukes til å identifisere barn med tanke på hvem som har behov for oppfølging, anser vi det for å være bedre om testene identifiserer for mange enn for få barn. Språkvansker kan ha en dyptgående virkning i et barns liv, og bare dersom vansken oppdages, kan barnet få hjelp til å redusere opplevelsen av et handikap i hverdagen. Hvis formålet er å sikre ”rene” grupper i forskning, vil det imidlertid være bedre med en strengere cut-off grense.



Det er et vanlig kriterium for diagnosen SSV at nonverbal IQ skal være innenfor det normale variasjonsområdet, og grensen settes ofte til over 85. Dette kriteriet har vi brukt i vår undersøkelse. Kriteriet er imidlertid omstridt, fordi vi foreløpig ikke vet nok om sammenhengen mellom IQ og språkvansker. Hvis en slik sammenheng er svak eller ikke finnes, bør man være forsiktig med å bruke nonverbal IQ som eksklusjonskriterium (Ottem & Lian, 2007). Forskere er uenige om hvor grensen bør settes. Plante (1998) anbefaler at grensen settes ved en nonverbal IQ på 70 eller 75, selv om dette medfører at barn som ikke har diskrepans mellom språklige og nonverbale ferdigheter, inkluderes i SSV-gruppen. I vår undersøkelse ville utvalget sannsynligvis ha blitt noe større dersom vi hadde fulgt Plantes anbefalinger. I så fall ville vi ha risikert at det ikke hadde vært noen diskrepans mellom verbale og nonverbale ferdigheter hos flere av barna i utvalget vårt. Dette ville ha vært i strid med den definisjonen av SSV som ligger til grunn for undersøkelsen vår, og ville ha økt risikoen for å identifisere barn med andre primære vansker enn språkvansker.

**Konklusjon:** Det ser ut til at operasjonaliseringen av begrepet SSV ved hjelp av de standardiserte testene i undersøkelsen vår er mangelfull. Det ville trolig ha vært mulig å oppnå et større samsvar mellom teoretisk og operasjonelt begrep dersom vi hadde benyttet en ekspressiv standardisert test i tillegg til reseptive tester.

### **Operasjonalisering av begrepet SSV i spørreskjemaet**

Spørreskjemaet 20 spørsmål er et ganske nytt instrument, og det har derfor ikke vært mange holdepunkter for hvor cut-off-grensen for språkferdigheter på dette måleinstrumentet skulle settes.

Som vi så i forrige avsnitt, er den valgte cut-off-grensen på språktestene uvalidert. Dette gjelder også cut-off-grensen som er satt på 20 spørsmål, men også den er blitt satt etter nøye overveielse. I en undersøkelse av sammenhengen mellom språktesten Språk 6-16 og 20 spørsmål om språkferdigheter fant Ottem (2009) at en skåre over 31 % av maksimal skåre forekom relativt sjelden hos typiske barn, og at skårer over 41 % av maksimal skåre forekom svært sjelden hos typiske barn. I undersøkelsen ble 31 % av maksimal skåre regnet som kritisk grense for å undersøke om det kan foreligge en språkvanske. I vår undersøkelse har vi satt grensen på 51 % av maksimal skåre av to årsaker: Ved å sette grensen såpass høyt ble misforholdet mellom testene og lærevurderingene ikke så stort som det ville ha blitt med en lavere grense, fordi 20 spørsmål ved en lavere grense ville ha identifisert mange flere barn.

Det er også større sannsynlighet for at de barna vi fant da vi satte grensen på 51 % av maksimal skåre, er barn med så store språklige vansker at deres fungering i skolehverdagen hemmes. Ved å sette kravet så høyt, økes sannsynligheten for at de barna som identifiseres, har språkvansker, men det er en fare for at ikke alle barn med språkvansker identifiseres.

### 5.3.3 Ytre validitet

Ytre validitet angår generalisering rettet mot en spesiell individpopulasjon, situasjon eller tid. Ytre validitet styrkes dersom utvalget består av relevante individer, situasjoner og tider (Lund, 2002b).

I undersøkelsen vår er utvalget valgt etter praktiske hensyn, og er dermed ikke tilfeldig trukket. Dette kan være en trussel mot ytre validitet. Utvalget er hentet fra to Oslo-skoler, og det kan stilles spørsmål ved om disse er representative for hele landet. Et av seleksjonskriteriene i undersøkelsen vår var at elevene skulle ha norsk som morsmål. Det betyr at en stor andel av minoritetsbarna er tatt ut. Variansen i utvalget vårt er derfor ikke så stor, slik at vi sitter igjen med barn uten de store vanskene. Dersom utvalget skulle ha vært representativt for populasjonen, måtte vi ha valgt et utvalg som var representativt for hele landet (både med hensyn til foreldrenes utdanning, økonomi og bosted). Så strenge krav til utvelgelse er ressurskrevende og har ikke vært mulig å gjennomføre innenfor rammene til vårt prosjekt. Hvis det er slik at utvalget vårt er lite representativt, er det noe overraskende at forekomsten av barn med mulig SSV som identifiseres av de ulike instrumentene i undersøkelsen, stemmer så godt med tall fra forskningslitteraturen. Det kan tenkes at utvalgets manglende representativitet ikke har hatt så stor betydning for funnene i denne undersøkelsen, fordi SSV er en vanske som ifølge de mest kjente årsaksteoriene ikke har sammenheng med miljøfaktorer (se avsnitt 2.1.7).

I vår undersøkelse har vi en lav statistisk styrke, og derfor kan vi ikke generalisere til populasjonen. Som vi har vist i diskusjonen over, støttes imidlertid mange av våre funn av forskningslitteratur på feltet.

### 5.3.4 Andre forhold som kan ha påvirket validiteten og reliabiliteten

**Prosjekttilknytning:** Vi har vært en del av et større forskningsprosjekt der premissene for gjennomføringen var satt på forhånd. Tilknytningen til KiSP-prosjektet har på flere måter

vært svært positivt. Den har gitt oss tilgang til data fra en stor gruppe med informanter, og mange av formalitetene som må ordnes i forbindelse med et forskningsprosjekt, var ordnet på forhånd. På den annen side var det faktum at hovedprosjektet har lagt føringer på hva slags informasjon som har blitt innhentet, ført til at vi ikke har kunnet operasjonalisere prosjektet vårt på en måte som er ideell for å svare på vår problemstilling. Ideelt sett ville vi for det første ha testet alle elever i utvalget med en ekspressiv språktest for å få en bedre operasjonalisering av begrepet SSV i undersøkelsen. Vi ville for det andre også ha forsøkt å få to lærere til å fylle ut spørreskjemaet for hver elev. I tillegg ville vi for det tredje ha innhentet opplysninger om tilleggsvansker for å sikre at barna som identifiseres ved hjelp av testene, har en primær språkvanske. Ingen av disse tiltakene har vært mulig å gjennomføre innenfor rammene av vårt prosjektet, og dette har trolig redusert validiteten til undersøkelsen. Når vi for eksempel ikke har tilgang til informasjon om tilleggsvansker, kan det tenkes at noen av barna som er identifisert som barn med mulig SSV, er elever med adferdsvansker, konsentrasjonsvansker og andre vansker av ikke-språklig art som har en språkvanske som er sekundær. I denne forbindelsen må det bemerkes at slik differensialdiagnostisering er krevende, og at mange barn med slike primære vansker ikke har fått noen diagnose. Som nevnt beskriver lærerne ved hjelp av 20 spørsmål om språkferdigheter i noen tilfeller vansker som kan oppfattes som pragmatiske. Det kan imidlertid være vanskelig å avgjøre om den pragmatiske vansken er primær eller sekundær, fordi man ofte ikke kan vite om det er språkvansker som hemmer barnets interaksjon med andre eller om det er pragmatiske vansker som er årsaken til dette.

**Vurderingskriteriet:** Vi har brukt 20 spørsmål om språkferdigheter som det kriteriet som testene skal måles opp mot. I motsetning til standardiserte tester som TROG og BPVS, kan 20 spørsmål sies å være økologisk valid i den forstand at det omfatter en bredere del av barnets språklige og kommunikative fungering enn det tester gjør. Lærerne vurderer eleven i en naturlig kontekst uten at barnet skal prestere noe under bestemte betingelser. En forutsetning for et kriterium er som tidligere nevnt at det har adekvate psykometriske egenskaper, det vil si at det bør la seg måle, være nøytralt og relevant for testens formål. Språkferdigheter lar seg måle ved at lærerne i spørreskjemaet skal gradere elevenes funksjonelle språkferdigheter på en skala fra 1 til 5. Det er også relevant for testens formål, fordi det inneholder utsagn om barns språkferdigheter, og dermed er valid i forhold til folks oppfatning av språkvansker. Som nevnt tidligere kan det være problemer knyttet til kriterievaliditet i den forstand at kriteriene ikke nødvendigvis har fullkommen reliabilitet og validitet. Ifølge Kleven (2002a) kan for

eksempel en lærervurdering ha dårligere reliabilitet enn testen. 20 spørsmål er ikke standardisert og normert, noe som kan bety en større fare for subjektive vurderinger som sies å være en trussel mot kriteriets reliabilitet. Faren for en subjektiv vurdering økes ved at skjemaet er fylt ut av bare én lærer. Det hadde styrket reliabiliteten og validiteten til vurderingskriteriet om skjemaet hadde blitt fylt ut av flere lærere og eventuelt også foreldre. Det at flere nærpersoner vurderer barnets språk, vil kunne gi et sannere bilde av barnet, fordi man observerer barnet i flere naturlige settinger. Det er en fordel for validiteten dersom læreren kjenner barnet godt, og har gode forutsetninger for å si om barnet har språklige vansker sammenlignet med andre barn. I motsatt fall svekkes validiteten.

Videre ser man ofte at en språkvanske kan tilsløres av andre vansker (for eksempel adferdsvansker), og det kan være en utfordring for den enkelte lærer å vite hva som er det egentlige problemet. Den kunnskapen som hver enkelt lærer har om et barn, kan være mer eller mindre bevisst. Læreren har gjennom sin yrkeskarriere tilegnet seg erfaringer og kunnskaper som overføres til vurderinger og handling i den daglige samhandlingen med hver enkelt elev. Det kan være personlig kunnskap, men kan også dreie seg om lite eksplisitt kunnskap som er allmenn innenfor yrkesgruppen (Hiim & Hippe, 1998). Dette er såkalt taus kunnskap som ifølge Polanyi (1967) ikke er verbalisert, men som det kan settes ord på. Spørreskjemaet 20 spørsmål om språkferdigheter kan hjelpe læreren til å sette ord på taus kunnskap, slik at kunnskap som ellers ville ha forblitt uuttalt, kan brukes til identifisering av barn med spesifikke språkvansker. Ifølge Hiim & Hippe (1998, 2001) kan en språkliggjøring av denne kunnskapen føre til økt faglig bevissthet om egen praksis og for lærergruppa som helhet.

### **5.3.5 Oppsummering av undersøkelsen**

Problemstillingen som vi søker å besvare med denne undersøkelsen, er som følger:

*I hvilken grad identifiserer TROG og BPVS barn med spesifikke språkvansker på en måte som er valid i forhold til læreres beskrivelse av vanskene?*

For å besvare problemstillingen stilte vi noen forskningsspørsmål som er blitt besvart underveis. Nedenfor gis det et kort sammendrag av svaret på hvert spørsmål.

*Gjør testene nøyaktige målinger?* Dette spørsmålet handler om kriterievaliditet. For å besvare dette spørsmålet måtte vi for det første se hvordan TROG, BPVS og 20 spørsmål om språkferdigheter samsvarte. Krysstabellene viste at TROG og BPVS i stor grad samsvarte med 20 spørsmål om hvilke barn som er normalspråklige, men at det var et misforhold mellom alle måleinstrumentene når det gjaldt hvilke barn som kan ha SSV. Ifølge BPVS er dette 5 %, ifølge TROG 2,5 % og ifølge 20 spørsmål 7,5 % av barna. Det er naturlig at måleinstrumenter ikke samsvarer fullstendig, men for å undersøke om dette misforholdet er akseptabelt, beregnet vi for det andre testen spesifisitet og sensitivitet. Våre analyser viste at testene innfrir kravet til spesifisitet med god margin, og altså er svært gode til å identifisere barn som ikke har vansker. Testene har imidlertid relativt lav sensitivitet, og er altså ikke godt egnet til å identifisere barn med spesifikke språkvansker i vår undersøkelse. Dette kan forklares både med forhold knyttet til testen teoretiske utgangspunkt, at ikke begge språkmodalitetene testes, at måleinstrumentene har ulikt omfang og at begrepet SSV ikke er operasjonalisert på en tilstrekkelig måte i vår undersøkelse. Det kan også forklares med forhold knyttet til testen begrepsvaliditet og økologiske validitet.

*Måler testene TROG og BPVS alle aspekter ved begrepet spesifikke språkvansker?* Dette spørsmålet handler om begrepsvaliditet. Vi kontrollerte begrepsvaliditeten empirisk og teoretisk. Den empiriske vurderingen av testen begrepsvaliditet tyder på at den er forholdsvis lav i vår undersøkelse. Av den teoretiske vurderingen av begrepsvaliditeten i undersøkelsen, konkluderte vi med at operasjonaliseringen av begrepet SSV ved hjelp av de standardiserte testene i undersøkelsen vår er mangelfull. Det ville trolig ha vært mulig å oppnå et større samsvar mellom teoretisk og operasjonelt begrep dersom vi hadde benyttet en ekspressiv standardisert test i tillegg til reseptive tester.

*Kan et misforhold mellom språktestene og spørreskjemaet 20 spørsmål om språkferdigheter forklares med forhold knyttet til økologisk validitet?* Casestudiene tyder på at testene er gode til å måle det de skal måle (dvs. reseptivt vokabular og grammatisk forståelse), men at det er noe de ikke måler: De fanger ikke opp barns språk i sosial samhandling. Spørreskjemaet 20 spørsmål måler et mye bredere spekter av barns språkferdigheter enn det testene gjør.

I denne undersøkelsen ser vi at lærerne fanger opp flere barn med mulig SSV enn testene, og det tyder på at testene underrapporterer SSV.

Undersøkelsen har svak validitet, og funnene kan derfor ikke generaliseres til populasjonen. Dersom utvalget hadde vært større og mer variert, ville dette sannsynligvis ha styrket validiteten til undersøkelsen. Som vi har vist i teoridelen er det i flere studier dokumentert at psykometriske tester ikke identifiserer barn som av klinikere oppfattes som barn med SSV (Aram, et al., 1992; Bishop & McDonald, 2009; Keegstra, et al., 2007). Vårt hovedfunn stemmer godt overens med disse.

### 5.3.6 Konklusjon

Riktig identifisering av barn med spesifikke språkvansker danner grunnlag for å ta avgjørelser om oppfølging som kan ha stor betydning i barns liv. Det er derfor viktig at instrumenter som brukes til identifisering av barn med slike vansker, er valide med tanke på identifisering. Testene TROG og BPVS er mye brukt til identifisering av barn med spesifikke språkvansker, både i forskning og klinisk sammenheng, men testmanualene gir ikke opplysninger om testenes validitet for identifisering av barn med SSV. Denne undersøkelsen retter derfor søkelyset mot de to testenes validitet til dette formålet. Som vi har sett, dreier testvaliditet seg om to sentrale spørsmål: hva måler testen og hvor godt måler den det? Validitet er ikke noe absolutt, men noe en test har i ulik grad.

Dersom TROG og BPVS anses som mål på reseptivt vokabular og grammatisk forståelse, og skal brukes til kartlegging av barns individuelle språkkompetanse, tyder funnene i denne undersøkelsen på at TROG og BPVS har svært høy validitet. For barna i vår undersøkelse har BPVS en noe høyere validitet enn TROG. Testene har også svært god validitet når det gjelder å identifisere hvilke barn som *ikke* har spesifikke språkvansker. Validiteten er størst innenfor et monologisk språksyn, der fokuset rettes mot individuell språkkompetanse, og kommunikasjon anses som isolerte handlinger som utveksles mellom individer.

Dersom TROG og BPVS anses som mål på om barn har spesifikke språkvansker, og et dialogisk språksyn legges til grunn, tyder funnene i denne undersøkelsen på at TROG og BPVS har relativt lav validitet. Spesifikke språkvansker er et komplekst fenomen, og denne undersøkelsen indikerer at identifisering av slike vansker krever at både reseptive og ekspressive språkferdigheter kartlegges. Identifisering med tester er imidlertid ikke tilstrekkelig dersom et dialogisk språksyn legges til grunn. Innenfor et dialogisk språksyn er ikke kommunikasjon et individuelt fenomen, men noe som skjer i interaksjon med andre og i en kontekst. Språktester kan ikke fange opp hvordan barns kommunikasjon fungerer i sosialt

samspill med andre, og har således lav økologisk validitet. Det er et problem dersom tester som skal identifisere barn med spesifikke språkvansker, ikke fanger opp barn som oppfattes som barn med en alvorlig vanske i det daglige. For å øke den økologiske validiteten til identifiseringsprosessen, bør tester derfor suppleres med en økologisk tilnærming til kartlegging. Denne undersøkelsen tyder på at synliggjøring av læreres tause kunnskap gjennom spørreskjemaet 20 spørsmål om språkferdigheter i så måte kan være et nyttig supplement.

## 6 VEIEN VIDERE

Bruk av tester er som vi har sett den vanligste tilnærmingen til identifisering av barn med spesifikke språkvansker, men slike tester fanger ikke opp viktige forhold knyttet til hvordan barn fungerer i hverdagen. Fortsatt leter man etter metoder for å identifisere barn med denne vansken på en stabil, konsistent og nøyaktig måte (Plante, 1998). Utarbeiding av evidensbaserte testkriterier kan tenkes å forbedre testenes evne til identifisering av barn, men det virker urealistisk å tro at tester noen gang kan bli så gode at de blir økologisk valide.

Denne undersøkelsen bekrefter at det bør forskes mer på gruppen av barn med spesifikke språkvansker. Først når man får vite mer om dem, er det mulig å finne gode metoder for identifisering. Mer forskning på økologiske tilnærminger til identifisering, og på forholdet mellom tester og folks oppfatning av barns vansker vil kunne øke validiteten og reliabiliteten til identifiseringsprosessene både i klinisk sammenheng og i forskning.



# Register

# Litteraturliste

- AERA. (1999). *Standards for educational and psychological testing*. Washington: American Educational Research Association.
- APA. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.
- Aram, Morris, R., & Hall, N. E. (1992). The Validity of Discrepancy Criteria for Identifying Children with Developmental Language Disorders. *Journal of Learning Disabilities*, 25(9), 549-554.
- Aram, Morris, R., & Hall, N. E. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research*, 36, 580-591.
- Aram, & Nation, J. E. (1975). Patterns of language behaviour in children with developmental language disorders. *Journal of Speech and Hearing Research*, 18.
- Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36, 189-208.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Befring, E. (2007). *Forskningsmetode, etikk og statistikk*. Oslo: Samlaget.
- Bele, I. V. (2008). Tilnærminger til språkvansker og læring - språk og makt. In I. V. Bele (Ed.), *Språkvansker Teoretiske perspektiver og praktiske utfordringer* (pp. 9-29). Oslo: Cappelen Damm.
- Bishop, D. V. M. (1997). *Uncommon understanding: development and disorders of language comprehension in children*. Hove: Psychology Press.
- Bishop, D. V. M. (1998). Development of Children's Communication Checklist (CCC): A Method for Assessing Qualitative Aspects of Communicative Impairment in Children. *Journal of Child Psychology and Psychiatry*, 39(6), 879-891.
- Bishop, D. V. M. (2003). *Test for Reception of Grammar - Second Edition (TROG-2)*. London: Harcourt Assessment.
- Bishop, D. V. M. (2006). What causes Specific Language Impairment in Children? *Current directions in psychological science* 15(5), 217-221.
- Bishop, D. V. M. (2009). *Test for Reception of Grammar - Second Edition (TROG-2) - norsk versjon*. Stockholm: Tryck AB.
- Bishop, D. V. M., Laws, G., Adams, C., & Norbury, C. (2006). High Heritability of Speech and Language Impairments in 6-year-old Twins Demonstrated using Parent and Teacher Report. *Behavior Genetics*, 36(2), 173-184.
- Bishop, D. V. M., & McDonald, D. (2009). Research Report: Identifying language impairment in children: combining language test scores with parental report. *International Journal of Language & Communication Disorders*, 44(5), 600-615.
- Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, 37, 391-403.
- Bjerkkan, K. M. (2000). *Verbal morphology in specifically language impaired children: evidence from Norwegian*. Unpublished Akademisk avhandling, Det humanistiske fakultet, Universitetet i Oslo, Oslo.
- Botting, N., McCracken, W., & Conti-Ramsden, G. (2002). Standard assessment tasks and children with a history of specific language impairment: Evidence of residual learning difficulties in relation to junior school placement. *Educational and Child Psychology*, 19(4), 4-15.

- Braut, G. S. (2009). Prediktiv verdi. *Store Norske Leksikon* Retrieved 5. februar, 2010, from [http://www.snl.no/sml\\_artikkel/prediktiv\\_verdi](http://www.snl.no/sml_artikkel/prediktiv_verdi)
- Brown, G. D. A., & Hulme, C. (1996). Nonword repetition, STM, and age-of-acquisition versus pronunciation-time limits in immediate recall for forgetting-matched acquisition: A computational model. In S. E. Gathercole (Ed.), *Models of short-term memory*. Hove: Psychology Press.
- Christophersen, K.-A. (2009). *Databehandling og statistisk analyse med SPSS*. [Oslo]: Unipub.
- Cohen, N. J., Davine, M., & Kelly-Meloche, M. (1989). Prevalence of unsuspected language disorders in a child psychiatric population. *Journal of the American Academy of Child and Adolescent Psychiatry* 28, 107-111
- Connolly, P. (2007). *Quantitative data analysis in education: a critical introduction using SPSS*. London: Routledge.
- Conti-Ramsden, G., & Botting, N. (1999a). Classification of children with specific language impairment: longitudinal considerations. *Journal of Speech, Language, and Hearing Research*, 42, 1195-1204.
- Conti-Ramsden, G., & Botting, N. (1999b). Characteristics of children attending language units in England: a national study of 7-year-olds. *International Journal of Language & Communication Disorders*, 34, 359-366.
- Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The Extent to Which Psychometric Tests Differentiate Subgroups of Children With SLI. *Journal of Speech, Language, and Hearing Research* 40(4), 765-777.
- Cook, T. D., Campbell, D. T., Reichardt, C. S., McCain, L. J., & McCleary, R. (1979). *Quasi-experimentation: design & analysis issues for field settings*. Boston: Houghton Mifflin Co.
- Cristophersen, K. A. (2006). *Databehandling og statistisk analyse med SPSS*. Oslo: Unipub A/S.
- Cristophersen, K. A. (2009). *Databehandling og statistisk analyse med SPSS*. Oslo: Unipub A/S.
- Dunn, Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale - Second Edition (Testbook)*. London: Nfer-Nelson Publishing Company Ltd.
- Dunn, M., Flax, J., Sliwinski, H., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence *Journal of Speech and Hearing Research*, 39(3), 643-654.
- Fey, M. E., Long, S. H., & P.L, C. (1994). Reconsideration of IQ criteria in the definition of specific language impairment. In R. V. Watkins & M. L. Rice (Eds.), *Specific language impairments in children* (pp. 161-178). Baltimore: Paul H. Brookes.
- Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Los Angeles, Calif.: Sage.
- Fletcher, P., & Ingham, R. (1995). Grammatical impairment. In P. Fletcher & MacWhinney (Eds.), *The handbook of child language* (pp. 603-622). Oxford: Blackwell.
- Florin, M., & Johannessen, K. S. (1988). *Dialoger: 6/88 tyst kunskap*. [Stockholm]: Carlsson Bokförlag.
- Fraenkel, J. R., & Wallen, N. E. (2008). *How to design and evaluate research in education*. New York: McGraw-Hill.
- Friel-Patti, S. (1999). Specific Language Impairment: Continuing Clinical Concerns. *Topics in language disorders*, 20, 1-13.

- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: an introduction*. Boston: Allyn and Bacon.
- Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech & Hearing Services in Schools, 30*(2), 196-206.
- Haywood, H. C. (1997). Interactive assessment. In R. Taylor (Ed.), *Assessment of individuals with mental retardation* (pp. 108-129). San Diego: Singular.
- Hiim, H., & Hippe, E. (1998). *Kompendium, hovedfag i yrkespedagogikk: profesjons- og yrkeskunnskap, utviklingsforskning, profesjons- og yrkesdidaktikk*. Oslo: Høgskolen i Akershus. Yrkespedagogisk institutt.
- Hiim, H., & Hippe, E. (2001). *Å utdanne profesjonelle yrkesutøvere*. Oslo: Gyldendal akademisk.
- Hughes, D., Turkstra, L. S., & Wulfeck, B. B. (2009). Parent and self-ratings of executive function in adolescents with specific language impairment. *International Journal of Language & Communication Disorders, 44*(6), 901-916.
- Hulme, C., & Roodenry, S. (1995). Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*, 373-398.
- Johnston, J. R. (1988). Specific language disorders in the child. In N. J. Lass (Ed.), *Handbook of speech-language pathology and audiology* (pp. 685-715). St. Louis, MO: C. V. Mosby.
- Johnston, J. R. (1994). Cognitive abilities of children with language impairment. In R. W. W. M. L. Rice (Ed.), *Specific language impairments in children* (pp. 107-123). London: Paul H. Brookes Publishing Co.
- Kamhi, A. (1998). Trying to Make Sense of Developmental Language Disorders. *Language, Speech & Hearing Services in Schools, 29*(1), 35-44.
- Keegstra, A. L., Knijff, W. A., Post, W. J., & Goorhuis-Brouwer, S. M. (2007). Children with language problems in a speech and hearing clinic: background variables and extent of language problems. *International Journal of Pediatric Otorhinolaryngology, 71*, 815-821.
- KiSP. (2009). Kunnskapsgenerering i det spesialpedagogiske forskningsfeltet (KiSP). Retrieved 8. mars, 2010, from <http://www.uv.uio.no/forsk/forsk-grupper/kunnskapsgenerering/index.html>
- Kleven, T. A. (2002a). Begrepsoperasjonalisering. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 141-183). Oslo: Unipub forlag.
- Kleven, T. A. (2002b). Ikke-eksperimentelle design. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 265-286). Oslo: Unipub forlag.
- Kleven, T. A. (2007a). Hvordan er begrepene operasjonalisert? Spørsmålet om begrepsvaliditet. In T. A. Kleven (Ed.), *Innføring i pedagogisk forskningsmetode* (pp. 120-138). Oslo: Unipub forlag.
- Kleven, T. A. (2007b). Statistikk. In T. A. Kleven (Ed.), *Innføring i pedagogisk forskningsmetode* (pp. 84-119). Oslo: Unipub forlag.
- Kruuse, E. (2005). *Kvantitative forskningsmetoder i psykologi og tilgrensendefag*. [København]: Dansk psykologisk Forl.
- Leonard. (2000). *Children with specific language impairment*. Cambridge, Mass.: MIT Press.
- Leonard, L. B. (1991). Specific language impairment as a clinical category. *Language, Speech and Hearing Services in Schools, 22*, 66-74.
- Lie, F., & Tetzchner, S. v. (2010). Dynamisk utredning og tiltak i skolen. *Skolepsykologi, 1*, 15-30.
- Lind, M. (2002). *Conversational cooperation: the establishment of reference and displacement in aphasic interaction*. Oslo: Unipub AS.

- Linell, P. (1998). *Approaching Dialogue. Talk, interaction and contexts in dialogical perspectives*. Amsterdam: John Benjamins publishing Co.
- Lund, T. (2002b). Metodologiske prinsipper og referanserammer. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 79-123). Oslo: Unipub forlag.
- Lund, T. (2002c). Avslutning. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 323-333). Oslo: Unipub forlag.
- Lyster, S.-A. H. (1994). *Språkrelaterte lærevansker hos barn og ungdom: kartlegging og tiltak*. Oslo: Universitetsforl.
- Lyster, S.-A. H. (2008). Barns språkvansker - generelle og spesifikke tiltak. In I. V. Bele (Ed.), *Språkvansker. Teoretiske perspektiver og praktiske utfordringer* (pp. 147-164). Oslo: Cappelen.
- McArthur, G. M., & Bishop, D. V. M. (2004). Which people with specific language impairment have auditory processing deficits? *Cognitive Neuropsychology*, 21(1), 79-94.
- McCauley, R. J. (2001). *Language Disorders in Children*. London: Lawrence Erlbaum Associates, Publishers.
- McCauley, R. J., & Demetras, M. J. (1990). The Identification of Language Impairment in the Selection of Specifically Language-Impaired Subjects. *Journal of Speech and Hearing Disorders*, 55, 468-475.
- Midtbø, T. (2007). *Regresjonsanalyse for samfunnsvitere: med eksempler i SPSS*. Oslo: Universitetsforl.
- Montgomery, J. W. (2005). Effects of input rate and age on the real-time language processing of children with specific language impairment. *International Journal of Language & Communication Disorders*, 40(2), 171-188.
- NESH. (2009, 11.12.2009). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi. Forskningsetiske komiteer. Retrieved December, 11, 2009, from <http://www.etikkom.no/no/Forskningsetikk/Etiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>
- Ottem, E. (2004). Diagnostisering av spesifikke språkvansker hos barn; inklusjons- og eksklusjonskriterier. *Nevropsykologi*, 7(2), 3-8.
- Ottem, E. (2009). 20 spørsmål om språkferdigheter - en analyse av sammenhengen mellom observasjonsdata og testdata. *Skolepsykologi*, 1, 11-27.
- Ottem, E., & Lian, A. (2007). Spesifikke språkvansker hos barn og unge. *Skolepsykologi*, 4(2007), 3-12.
- Ottem, E., & Lian, A. (2008a). Spesifikke språkvansker I. In I. V. Bele (Ed.), *Språkvansker: teoretiske perspektiver og praktiske utfordringer* (pp. s.31-42). Oslo: Cappelen akademisk.
- Ottem, E., & Lian, A. (2008b). Spesifikke språkvansker II: Teori og empiri i kognitiv psykologisk forskning. In I. V. Bele (Ed.), *Språkvansker. Teoretiske perspektiver og praktiske utfordringer* (pp. 43-58). Oslo: Cappelen Akademiske Forlag.
- Plante, E. (1998). Criteria for SLI: The Stark and Tallal Legacy and Beyond. *Journal of Speech, Language, and Hearing Research*, 41, 951-957.
- Plante, E., & Vance, R. (1994). Selection of preschool language tests: a data-based approach. *Language, Speech & Hearing Services in Schools*, 25, 15-24.
- Polanyi, M. (1967). *The tacit dimension*. Garden City, N. Y.: Doubleday.
- Rapin, I., & Allen, D. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling*. New York: Academic Press.
- Raven, J. C., Court, J. H., & Raven, J. (1986). *Manual for Raven's progressive matrices and vocabulary scales*. London: H. K. Lewis.

- Rice. (2004). Growth models of developmental language disorders. In M. L. Rice & S. F. Warren (Eds.), *Developmental language disorders: from phenotypes to etiologies* (pp. 207-240). Mahwah, N.J.: Lawrence Erlbaum.
- Rice, Warren, S. F., & Betz, S. K. (2005). Language symptoms of developmental language disorder: An overview of autism, Down syndrome, fragile X, specific language impairment and Williams syndrome. *Applied Psycholinguistics*, 26, 1-27.
- Rydz, D., Shevell, M. I., Majnemer, A., & Oskoui, M. (2005). Developmental Screening. *Journal of Child Neurology*, 20(1), 4-21.
- Rygvold, A. L. (2008). Språkvansker hos barn. In E. Befring & R. Tangen (Eds.), *Spesialpedagogikk* (pp. 229-249). Oslo: Cappelen akademisk forl.
- Sattler, J. M. (2001). *Assessment of children: cognitive applications*. San Diego, Calif.: J.M. Sattler.
- Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility Criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech & Hearing Services in Schools*, 37(1), 61-72.
- Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders*, 46, 114-180.
- Svartdal, F. (2004). *Psykologiens forskningsmetoder: en introduksjon*. Bergen: Fagbokforl.
- Svartdal, F. (2009). *Psykologiens forskningsmetoder: en introduksjon*. Bergen: Fagbokforl.
- Sveen, A. (2005). Pragmatikk. In K. E. Kristoffersen, H. G. Simonsen & A. Sveen (Eds.), *Språk: en grunnbok* (pp. 95-120). Oslo: Universitetsforl.
- Tallal, P. (1976). Rapid auditory processing in normal and disordered language development. *Journal of Speech, Language, and Hearing Research*, 9, 182-198.
- Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Najjarajan, S., et al. (1996). Language Comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271, 81-84.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40, 1245-1260.
- Wangensteen, B. (2005). *Bokmålsordboka: definisjons- og rettskrivningsordbok*. Oslo: Kunnskapsforlaget.
- Whithurst, G. J., & Fischel, J. E. (1994). Practitioner Review: Early Developmental Language Delay: What, If Anything, Should the Clinician do about It? *Journal of Child Psychology and Psychiatry*, 35(4), 613-648.
- WHO. (1993). *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic criteria for research*. Geneva: World Health Organization.
- WHO. (2000). *ICD-10: den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer : systematisk del og alfabetisk indeks*. Oslo: Statens helsetilsyn.
- WHO, KITH, & helsedirektoratet, S.-o. (2004). Internasjonal klassifikasjon av funksjon, funksjonshemming og helse. Retrieved 1. april, 2010, from [http://www.kith.no/upload/3423/setup\\_ICF2004\\_v2\\_pc.zip](http://www.kith.no/upload/3423/setup_ICF2004_v2_pc.zip)

# Vedlegg 1: 20 spørsmål

20 spørsmål om språkferdigheter					
Sett i forhold til barn på samme alder, hvor godt passer følgende utsagn for barnet? Bruk skalaen fra 1 til 5. Vennligst kryss av for hvert utsagn. Prøv å svare på alt selv om du ikke er helt sikker eller synes utsagnet virker rart.					
			Helt galt		Helt riktig
			1	2	3 4 5
<b><i>Semantisk evne</i></b>					
1. Glemmer ord som han/hun vet hva betyr			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Blander sammen ord med lik mening (for eksempel skjorte, genser, jakke)			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Har vansker med å forstå hva vanlige ord betyr			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Har problemer med å svare på spørsmål like raskt som andre			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Leter ofte etter de riktige ordene			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Bruker ufullstendige setninger			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Bruker korte setninger når han/hun skal svare på spørsmål			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Har vansker med å gjenfortelle en historie han/hun har hørt			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sum					
<b><i>Vansker med å motta et budskap</i></b>					
9. Det virker som om det han/hun lærer ikke fester seg			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Har vansker med å huske ting			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Vansker med å forstå hva andre sier			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Misforstår muntlige instruksjoner og beskjeder			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Har problemer med å huske beskjeder			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Misoppfatter situasjoner og hendelser			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sum					
<b><i>Vansker med å formidle et budskap</i></b>					
15. Eleven er vanskelig å forstå			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Problemer med å gi uttrykk for behov			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Blir ikke forstått av andre			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Tar sjelden språklige initiativ			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Har problemer med å uttale ord riktig			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Mestrer ikke dialog med jevnaldrende			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sum					
Sum totalt					

## Vedlegg 2: enkel deskriptiv analyse

Gruppe i 20 spørsmål	Case	Kjønn	Trinn	Z-skåre på TROG	Z-skåre på BPVS	20 spørsmål	Begreper	Ordspenn	Setningsminne	Raven
Gruppe 1 (20%-30% av maksimal skåre)	1	Gutt	småskole	1.75	2.03	20				1.78
	2	Jente	småskole	.13	-.04	20				-.72
	3	Jente	småskole	.91	.81	27				-.47
	4	Jente	småskole	.92	1.40	20				1.80
	5	Jente	småskole	-.98	.64	25				-.72
	6	Gutt	småskole	-.41	.17	20				-.12
	7	Gutt	småskole	.60	-1.08	20				.10
	8	Jente	småskole	-.98	.30	28				-.12
	9	Jente	småskole	.60	-.51	20				-.10
	10	Gutt	småskole	1.41	-.37	20				1.64
	11	Gutt	småskole	1.03	-.02	20				.79
	12	Jente	småskole	-.08	-1.40	20				-.89
	13	Gutt	mellom	.35	1.36	29				.35
	14	Jente	mellom	-.56	-.33	22				-.58
	15	Jente	mellom	.66	-.17	26				-.58
	16	Jente	mellom	-.39	.44	22				.42
	17	Jente	mellom	1.75	-.30	20				1.36
	18	Jente	mellom	.91	-.87	20				1.18
	19	Jente	mellom	.51	.30	20				.83
	20	Gutt	mellom	.21	.05	21				-.54
	21	Jente	mellom	.13	-.17	20				-.12
	22	Gutt	mellom	.13	.54	20				-.57
	23	Jente	mellom	-.08	.42	20				-.94
	24	Jente	mellom	2.40	2.40	20				1.11
	25	Gutt	småskole	.91	1.16	20				.38
	26	Gutt	småskole	-.41	.44	20				-.38
	27	Gutt	småskole	1.92	-.10	20				-.94
	28	Gutt	småskole	.51	1.59	20				.47
	29	Jente	småskole	.91	-.44	20				-.12
	30	Jente	småskole	.60	-.69	24				.54
	31	Jente	småskole	.60	-1.08	24				.31
	32	Jente	småskole	.51	.81	20				.83
	33	Gutt	småskole	.21	.87	20				-.31
	34	Jente	småskole	-.41	-1.34	20				.21
	35	Jente	småskole	-.41	-.59	20				1.02
	36	Gutt	småskole	.13	.09	20				.21
	37	Gutt	småskole	-.41	-.75	20				.57
	38	Jente	småskole	-.41	1.00	20				-.12
	39	Jente	småskole	.04	.24	20				-.55
	40	Jente	småskole	-1.08	-.40	24				.59
	41	Jente	småskole	-.35	-.17	21				-.35
	42	Jente	småskole	.04	-.40	20				2.05
	43	Jente	småskole	1.36	.76	20				1.36



	44	Gutt	småskole	-.69	1.51	25				-.80
	45	Jente	småskole	.69	-.88	21				.59
	46	Jente	småskole	.69	.62	20				.37
	47	Jente	småskole	.04	-.24	20				-.69
	48	Jente	mellom	1.7	.72	20				.86
	49	Jente	mellom	.69	.44	20				.12
	50	Jente	mellom	-1.08	.94	20				-.55
	51	Gutt	mellom	.69	.44	20				.12
	52	Gutt	mellom	.69	-.04	25				1.12
	53	Jente	mellom	-.57	.25	20				.66
	54	Jente	mellom	.00	1.24	20				1.35
	55	Gutt	mellom	.66	.71	20				.00
	56	Jente	mellom	-.98	-.82	25				.42
	57	Jente	mellom	1.48	1.73	20				1.06
	58	Gutt	mellom	1.66	.94	25				1.82
	59	Gutt	mellom	.69	.94	20				1.49
	60	Jente	mellom	1.48	.00	25				-.58
	61	Jente	mellom	-1.28	.94	21				.00
	62	Jente	mellom	.66	-.51	25				.00
	63	Jente	mellom	-.39	-.04	43	8	10	9	-.92
	64	Gutt	mellom	.35	.35	42	12	12	13	.76
	65	Jente	mellom	.69	-.67	47	9	10	14	-.12
	66	Jente	mellom	.04	2.07	40	9	12	8	1.12
	67	Gutt	mellom	.35	-.17	47	12	3	7	.00
<b>Gruppe 2 (31%- 50% av maksimal skåre)</b>	68	Gutt	mellom	.32	-.04	49				.21
	69	Gutt	mellom	-.39	-.88	36	9	5	9	.12
	70	Gutt	mellom	.69	.12	39				-.28
	71	Gutt	mellom	.04	-.04	35				.12
	72	Gutt	mellom	1.66	1.21	39	10	7	11	1.49
	73	Gutt	mellom	.00	.51	34	11	12	8	.85
	74	Gutt	mellom	.66	.25	33	4	7	7	-.58
<b>Gruppe 3 (51%- 100% av maksimal skåre)</b>	75	Gutt	mellom	-.39	-.67	62	6	3	3	.86
	76	Gutt	mellom	-1.66	-2.07	77	2	5	6	-.12
	77	Jente	mellom	-.39	-.53	54	5	7	4	.59
	78	Jente	mellom	-1.01	-.76	76				-.76
	79	Gutt	mellom	.69	1.5	53				-.28
	80	jente	mellom	-1.08	-1.64	60	5	9	10	.37

# Vedlegg 3: beskrivelse av TROG-2

## Test for Reception of Grammar (TROG 2)

TROG 2 (Bishop, 2009) er norsk versjon av en britisk test som opprinnelig ble utviklet av Dorothy Bishop på slutten av 1970-tallet. Testen er oversatt og standardisert til norske forhold, og er beregnet for barn fra 4-16 år. TROG-2 er en reseptiv språktest for kartlegging av grammatisk forståelse som kan gi et godt grunnlag både for å finne en persons funksjonsnivå og for å kunne diagnostisere spesifikke grammatiske vansker. I den britiske versjonen av manualen vises det til studier der den opprinnelige versjonen av TROG har vært brukt. Ifølge disse studiene har testen vist seg å være et godt redskap ved utredning av blant annet spesifikke språkvansker (Bishop, 2009). I den norske manualen oppgis det at testen ikke bør brukes alene i diagnostiseringen av språkvansker, men kan inngå som en del av et testbatteri.

Testen er individuelt administrert og består av 80 flervalgsoppgaver. Oppgavene er organisert i blokker med fire oppgaver for hvert grammatisk forhold som undersøkes. (En blokk er en betegnelse for en gruppe med oppgaver som tester barnets forståelse av en bestemt grammatisk konstruksjon.) Hver oppgave har fire svaralternativ som presenteres i form av bilder. Testleder skal vise en oppgave til barnet og si: ”Her har vi en bok med fire bilder på hver side. Nå skal jeg si noe, og så skal du peke på det bildet som passer best til det jeg sier.” Barnet skal så peke på det som passer med setningen testleder leser fra en stimulusbok, og behøver altså ikke å bruke språk for å avgi svar. For hver oppgave er det ett riktig svaralternativ. De tre andre svaralternativene betegnes som distraktorer, og skiller seg fra korrekt svar kun ved at man har endret et grammatisk eller leksikalt element. Eksempel: I oppgave K1 ”Kua blir jaget av jenta” representerer svaralternativ 1) ”Jenta blir jaget av kua” en grammatisk distraktor. Ordrekkefølgen er svært viktig for forståelsen, og distraktorene er konstruert på en slik måte at man kan oppdage feilmønstre hos testpersonen. Vanskegraden er økende, og testingen avsluttes når 5 blokker etter hverandre er galt besvart. Dersom barnet svarer galt på en oppgave innen en blokk, regnes hele blokken som feil.

# Vedlegg 4: beskrivelse av BPVS II

## **British Picture Vocabulary Scale II (BPVS II)**

British Picture Vocabulary Scale II er en britisk test som er utviklet av Dunn, Dunn, Whetton og Burley (1997). Det er en individuelt administrert test som kartlegger reseptivt vokabular for aldersgruppen 3-15 år. Ifølge Dale & Reichert (1957) referert i Dunn & Dunn, Whetton & Burley (1997) kan testen predikere skolefaglig fungering. Testen er ikke oversatt og standardisert til norske forhold, så vi har benyttet den britiske versjonen av testen.

Testen består av 14 oppgavesett med 12 ord i hvert sett. Hver oppgave har fire svaralternativ som presenteres i form av bilder. Barnet skal velge det bildet som passer til stimulusordet som testleder leser opp. Testlederen sier: ”Jeg skal si et ord, og så vil jeg at du skal peke på det bildet som passer best til det jeg sier”. Man starter med det oppgavesettet som tilsvarer barnets alder (evt. forventet alder). Har barnet mer enn én feil i første oppgavesett, må tidligere oppgaver administreres for å finne barnets nedre grense. Dersom barnet har åtte eller flere feil i et oppgavesett, avsluttes testingen. Dette er barnets øvre grense. Testen går fra høy- til lavfrekvente ord, og fra konkrete til abstrakte ord og begreper. Vanskegraden er økende, og verbal respons er ikke nødvendig.

# Vedlegg 5: beskrivelse av Raven SPM

## Raven Standard Progressive Matrices (SPM)

Raven SPM (Raven, et al., 1986) er en nonverbal test som er ment for bedømmelse av generelle evner. Testen er beregnet for aldersgruppen 6 til 17 år, men kan også administreres til voksne. Administreringen gjøres individuelt eller i gruppe. I vår undersøkelse er testen gitt som en gruppetest. Raven SPM består av 60 testledd med varierende vanskegrad og ulike krav til logisk tenkning. Hvert ledd består av en figur eller en matrise hvor et element mangler, og oppgaven går ut på å identifisere det manglende element blant fem alternativer. Barnet skal sammenlikne mønstre og mønsterkombinasjoner og foreta analogiske slutninger. Testen er en såkalt matrisetest som måler evnen til å kunne foreta en logisk utledning av nye relasjoner og mønstre på grunnlag av kjent informasjon. Den gjør det mulig å vurdere problemløsningskapasitet uten at det benyttes verbalt språk.