

# TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER

*Hva er dagenens praksis, hvilke feilkonklusjoner kan oppstå og kan dynamiske tester øke reliabel informasjon i utredningen av elever fra språklige minoriteter?*

**Inger-Johanne Gundersen**



Hovedoppgave i pedagogikk – Pedagogisk forskningsinstitutt.

UNIVERSITETET I OSLO

11.04.2006

---

# Innhold

<b>INNHOOLD</b> .....	<b>2</b>
<b>SAMMENDRAG AV HOVEDOPPGAVEN. TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER.</b> .....	<b>7</b>
<b>1. KAPITTEL 1. REDEGJØRELSE OG BEGREPSAVKLARINGER</b> .....	<b>10</b>
1.1 BAKGRUNN FOR HOVEDOPPGAVENS PROBLEMSTILLINGER.....	11
1.2 SPRÅKLIGE MINORITETER.....	12
1.2.1 Tospråklighet og intelligens .....	15
1.2.2 Utbredelse av språklige minoriteter.....	15
1.2.3 Hva er språklige minoriteter .....	20
1.2.4 Isfjellmodellen .....	22
<i>Sveitserost - metaforen</i> .....	23
1.2.5. Erfaringsbakgrunn og sosiale forskjellig.....	23
1.3 TEST OG TESTING .....	24
1.3.1 Hva er en test?.....	25
1.3.2 Hvorfor teste?.....	26
1.3.3 Hva måler testen?.....	27
1.3.4 Test kompetanse.....	28
1.4 PSYKOMETRI.....	29
1.4.1 Reliabilitet .....	30
1.4.2 Validitet .....	31
1.4.3 Begrepsvaliditet.....	31
1.4.4 Prediktiv validitet .....	33

1.4.5	<i>Bias</i> .....	33
1.4.6	<i>Bias og validitet</i> .....	34
1.4.7	<i>Standardisering og normering</i> .....	35
1.5	INTELLIGENS OG MENNESKESYN.....	36
1.6	PRESENTASJON AV TESTENE, WISC-III, LEITER-R, LEM.....	39
1.6.1	<i>WISC-III</i> .....	41
1.6.2	<i>LEITER-R</i> .....	43
1.6.3	<i>LEM</i> .....	45
1.7	SAKKYNDIG UTREDNING.....	50
1.7.1	<i>Det er viktig å skille mellom testing og utredning</i> .....	50
1.7.2	<i>Forholdet mellom tilpassetopplæring og spesialundervisning</i> .....	50
1.7.3	<i>Når blir språket et problem for språklige minoriteter og hva skjer når problemet oppstår?</i> 52	
1.7.4	<i>Sakkyndig vurdering</i> .....	54
1.7.5	<i>Elevens fremtidsutsikter</i> .....	56
1.7.6	<i>Kompetansekrav til de ansatte i PPT</i> .....	57
1.8	KILDEREFLEKSJONER.....	58
1.9	OPPGAVENS OPPBYGNING.....	58
1.10	AVSLUTNING.....	59

## **2. KAPITTEL 2 PROBLEMSTILLING NR. 1: HVILKE PROBLEMER INNEBÆRER DAGENS PRAKSIS MED BRUK AV STATISKE TESTER PÅ ELEVER FRA SPRÅKLIGE MINORTITETER?**

.....	<b>62</b>	
2.1	ANALYSE AV SAKKYNDIGE VURDERINGER.....	63
2.1.1	<i>Har de sakkyndige vurderingene vektlagt elevens minoritetsspråklige bakgrunn? ...</i>	66

2.1.2	<i>Hvilke type data er innhentet og er data innhenting preget av en helhetlig - økologisk tenkning?</i> .....	67
2.1.3	<i>Hvilke tester er brukt i vurderingen?</i> .....	67
2.1.4	<i>Hvordan er testene tolket?</i> .....	68
2.1.5	<i>Har det vært bruk av dynamiske tester?</i> .....	68
2.1.6	<i>Min oppsummering av de 6 sakkyndige vurderingene</i> .....	69
2.2	FAGPERSONER UTTALER SEG OM DAGENS PRAKSIS .....	70
2.3	HVA SIER DE SAKKYNDIGE OM EGEN PRAKSIS? .....	72
2.4	STEREOTYP BEHANDLING AV ELEVER FRA SPRÅKLIGE MINORITETER. HVA SIER FORELDRE OM DAGENS PRAKSIS.....	74
2.5	HVILKE PROBLEMER ER DET MED STATISK TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER?75	
2.5.1	<i>Elever fra språklige minoriteter stemples som mentalt retarderte.</i> .....	75
2.6	PROBLEMENE LØSES IKKE VED Å FJERNE TESTENE ELLER SAKKYNDIGE VURDERINGER.....	76
2.7	AVSLUTNING .....	77
<b>3.</b>	<b>KAPITTEL 3. PROBLEMSTILLING NR. 2: HVORDAN KAN MAN ELIMINERE FOREKOMSTEN AV FEILKONKLUSJONER I TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER?</b> .....	<b>78</b>
3.1	HVORDAN TESTENE BØR BRUKES.....	79
3.2	TOLKNING.....	80
3.2.1	<i>Hvordan tolke testresultater?</i> .....	80
3.2.2	<i>Tolkningens historie.</i> .....	81
3.2.3	<i>Hva sier WISC-III manualen om tolkning?</i> .....	82
3.2.4	<i>Hva sier LEITER-R manualen om tolkning?</i> .....	84
3.2.5	<i>Hva sier LEM om tolkning?</i> .....	86

---

3.2.6	<i>Tolkningstyper</i> .....	86
	<i>Psykometriske forhold ved tilpassning av tester</i> .....	87
	<i>Økologisk tenkning</i> .....	88
	<i>Ipsativ kartlegging og tolkning</i> .....	90
	<i>Kritisk tenkning ved tolkning av tester</i> .....	91
	<i>Intelligent testing med WISC-III</i> .....	93
3.2.7	<i>Tolkning i et flerkulturelt samfunn</i> .....	95
3.2.8	<i>Hvordan bli en god testtolker</i> .....	97
3.3	OPPLÆRING OG FAGLIG OPPDATERING .....	100
3.4	KULTURREDUSERTE TESTER .....	102
3.5	BENYTTTE DYNAMISKE TESTER .....	103
3.6	AVSLUTNING .....	104
<b>4.</b>	<b>KAPITTEL 4. PROBLEMSTILLING NR.3: KAN DYNAMISKE TESTER REDUSERE FOREKOMSTEN AV FEILKONKLUSJONER OG ØKE RELIABEL INFORMASJON I UTREDNINGEN AV ELEVER FRA SPRÅKLIGE MINORITETER?</b> .....	<b>106</b>
4.1	TEORIGRUNNLAGET I DYNAMISKE TESTER .....	107
4.2	HVORFOR KOM DYNAMISKE TESTER .....	108
4.3	HVA ER DYNAMISKE TESTER? .....	110
4.4	HOVEDGRUPPER AV DYNAMISKE TESTER OG TESTFORMAT .....	113
4.4.1	<i>Metakognitive intervensjoner</i> .....	115
4.4.2	<i>Lære med testen</i> .....	115
4.4.3	<i>Restukturere testsituasjonen</i> .....	117
4.4.4	<i>Trening av en enkelt kognitiv funksjon</i> .....	119
4.4.5	<i>Testenes format</i> .....	120

---

4.5	NÅR BØR MAN BRUKE DYNAMISK TESTING?.....	121
4.6	EVALUERING AV DYNAMISKE TESTER.....	122
4.7	EMPIRI OM REDUKSJON AV FEILKONKLUSJONER VED BRUK AV DYNAMISKE TESTER PÅ SPRÅKLIGE MINORITETER.....	125
4.8	HVORFOR DYNAMISKE TESTER ER LITE BRUKT.....	131
4.9	AVSLUTNING.....	133
<b>5.</b>	<b>KAPITTEL 5. KONKLUSJON .....</b>	<b>135</b>
<b>6.</b>	<b>KILDELISTE .....</b>	<b>137</b>

---

## **SAMMENDRAG AV HOVEDOPPGAVEN. TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER.**

**Hva er dagens praksis, hvilke feilkonklusjoner kan oppstå og kan dynamiske tester øke reliabel informasjon i utredningen av elever fra språklige minoriteter?**

Testing av elever fra språklige minoriteter har lenge vært et mye diskutert og vanskelig område. Det har i hovedsak vært bruken av statiske tester som har vært kritisert. Disse testene måler noen av kunnskaps- og ferdighetsområdene som eleven har lært til nå. Dette har vært og er et godt estimat for hva *norske elever* har lært til nå og hvordan de vil gjøre det i fremtiden. Hva med elever som ikke har lært det vi forventer at de skulle ha lært til nå? Dette gjelder en del elever fra språklige minoriteter. En betydelig andel av disse elevene har ofte ikke den samme erfaringsbakgrunnen som en norsk gjennomsnittselev, som de statiske testene er normert etter. De tradisjonelle testene som brukes av PPT måler ikke elevens læringspotensial. Måles ikke læringspotensialet, men kun noe av det eleven burde ha lært til nå, kan den prediktive validiteten derfor blir svak for de elevene som ikke tilhører testens normeringsgrunnlag. Ved bruk av dynamiske tester i utredningen av elever fra språklige minoriteter ville man kunne få målt elevens læringspotensial og ikke bare målt utvalgte områder av det eleven har lært til nå. Det har vist seg at dynamiske tester har en høyere prediktiv validitet for elever fra språklige minoriteter enn statiske tester.

Jeg har valgt å se på følgende tre problemstillinger.

1. Hvilke problemer innebærer dagens praksis med bruk av statiske tester på elever fra språklige minoriteter?
2. Hvordan kan man eliminere forekomsten av feilkonklusjoner i testing av elever fra språklige minoriteter?
3. Kan dynamiske tester redusere forekomsten av feilkonklusjoner og øke reliabel informasjon i utredningen av elever fra språklige minoriteter?

Hensikten med oppgaven er at den kan være en bevisstgjøring og forhåpentligvis til praktisk nytte for PP-rådgivere, lærere og elever fra språklige minoriteter.

Min metodiske tilnærming er litteraturstudier. Jeg har benyttet meg av normative – og preskriptive kilder. De normative kildene er Opplæringsloven, Håndbok for PP-tjenesten og Etske prinsipper for PP-rådgivere. De preskriptive kildene er bøker, artikler, testmanualer og sakkyndige vurderinger. For å belyse dagens praksis har jeg brukt 6 sakkyndig vurderinger fra et tilfeldig valgt PPT-kontoret i Oslo. I litteraturstudiene og analysen har jeg drøftet testteori/ testbruk i forhold til praksisfeltet ved testing av elever fra språklige minoriteter.

Det er utbredt enighet om behovet for å gjøre noe med testpraksis for elever fra språklige minoriteter. Dagens testpraksis kan føre til feilkonklusjoner som igjen fører til feilplasseringer og uegnede tiltak som kan og har resultert i sosial reproduksjon og stemping av elever fra språklige minoriteter.

Problemer ved dagens praksis er at man ikke får skilt mellom språkforskjeller og språkvansker. Mange av de testene som brukes bygger sine poengskalaer på hva som er normalvariasjonen i målpopulasjonens kulturelle og språklige erfaringsbakgrunn. Problemet er at elever fra språklige minoriteter ikke har den samme kulturelle og språklige erfaringsbakgrunn som norske elever. Det vil si at testen ofte måler hvor godt eleven kan språket og ikke om eleven har lærevansker.

Test- og språkkompetanse hos de sakkyndige kan også variere. Det er den sakkyndige som velger og legger premissene for hvilke tester som skal brukes og hvordan testen brukes, samt tolkningen av testresultatene. I de sakkyndige vurderingene jeg undersøkte var WISC-III, som er en statisk test, brukt i alle tilfellene. Det vil si at statiske tester brukes ved testing av elever fra språklige minoriteter. Testmanualen til WISC-III er tydelig på at testen ikke er egnet for elever fra språklige minoriteter. Brukes statiske tester på disse elevene bør valg av test og tolkning begrunnes. Jeg hevder at det ikke er den valgte testen som er årsaken til feilkonklusjoner, men hvordan testens resultat er tolket. Det er derfor viktig å skille mellom

---

test og testing, fordi det ofte er testen som blir beskyldt for å være dårlig når det er testingen som blir kritisert. Jeg ønsket å finne ut om testene var så dårlige som de ble beskyldt for i kritikken som var rettet mot dem. Etter å ha lest grundig gjennom testmanualene til WISC-III, Leiter-R og LEM for å se hva de sier de skal måle, hvem testen er beregnet for og tolkninger av testresultatene, kom jeg fram til at testen ikke alltid ble brukt i henhold til manualens spesifikasjoner. En test er et verktøy som ikke blir bedre enn de hendene som bruker den. Jeg mener at tester kan brukes på andre måter enn det som står i testmanualen, men da må det fremkomme hvorfor man har valgt å bruke testen på en annen måte. Hvordan testresultat tolkes har en avgjørende betydning for de konklusjoner man kommer frem til. Jeg har derfor vektlagt tolkning av tester da jeg mener det er her feilkonklusjoner ofte oppstår.

Dynamiske tester kan redusere faren for feilkonklusjoner ved testing av elever fra språklige minoriteter. Vygotskys teori om ”Sonen for den nærmeste utvikling” er teorigrunnlaget i dynamiske tester. Testresultatet gir et mål på hva som er elevens læringspotensial.

Dynamiske tester måler eleven i forhold til seg selv, mens en statistisk test måler eleven i forhold til en normeringspopulasjon. Det vil si at dynamiske tester måler hva eleven kan lære mens statiske tester måler hva eleven har lært til nå innenfor noen kunnskaps og ferdighets områder. Internasjonale studier bekrefter dette. Elever som skårer lavt på en statistisk test kan ha et høyt læringspotensial, men det kan være skjult av språkforskjeller.

Dagens praksis ved bruk av statiske tester på elever fra språklige minoriteter er ikke god nok, noe må gjøres. Løsningen er ikke å fjerne testene og de sakkyndige vurderingene, det vil bare gjøre vondt verre. Det er testmetodene som må forbedres. Kunnskap om testteori, tolkningsmetoder og tospråklig utvikling er av stor betydning for hvordan tester brukes og hvordan de blir tolket. Dette kan være med å redusere faren for feilkonklusjoner. I tillegg vil bruk av dynamiske tester ytterligere redusere faren for feilkonklusjoner i utredninger av elever fra språklige minoriteter.

# 1. KAPITTEL 1. REDEGJØRELSE OG BEGREPSAVKLARINGER

## INNLEDNING

I dette kapitlet tar jeg opp bakgrunnen for hovedoppgavens problemstillinger om testing av elever fra språklige minoriteter. Videre redegjør jeg for problemene og plasserer dem i sin rette kontekst. Dette er et relativt stort kapittel, men jeg mener det er viktig å avklare begrepene og plassere problemene i sin rette kontekst. Jeg har valgt å vektlegge begrepsavklaringer fordi det ikke alltid gjøres i den litteraturen jeg har lest. Innholdet i begrepene, ”språklige minoriteter, test og testing, psykometri og intelligens”, som er brukt i denne oppgaven kan variere.

Historien er av betydning for dagens praksis. Jeg starter derfor kapitlet med et historisk tilbakeblikk på hvordan tester ble brukt / misbrukt på språklige minoriteter.

Forståelse av det som blir sagt og gjort i testsituasjonen er avgjørende for elevens testresultat. Elever fra språklige minoriteter kan derfor bli testet for språkkompetanse og ikke det testen er ment å skulle måle. Tester har blitt kritisert for å være kulturavhengige fordi språklige minoriteter har skåret lavt. Er det testene som er kulturavhengige eller er det testingen som ikke har holdt seg til manualen? Jeg har lest grundig gjennom testmanualene for å finne ut hvordan testen skal brukes og sett på testens psykometriske kvaliteter. Testmanualene jeg har sett på er WISC-III, Leiter-R og LEM .

Testing er en av flere metoder som brukes i en utredning, dette er noe av grunnlaget for en sakkyndig vurdering. Hvilke kompetansekrav stilles til de sakkyndige i PPT?

Jeg avslutter kapitlet med metodiske refleksjoner og oppgavens videre oppbygging.

---

## 1.1 Bakgrunn for hovedoppgavens problemstillinger.

Da jeg var i praksis i PPT møtte jeg en elev som hadde bodd hele sitt liv i Norge uten å ha lært norsk. Dette var en elev som hadde vokst opp med kun å snakke morsmålet sitt. Da denne eleven begynte i 1. klasse satt eleven stille og forskremt i klassen. Det var læreren som etter hvert oppdaget at eleven ikke kunne norsk. Eleven hadde mye uro i kroppen og var derfor henvist til PPT for en sakkyndig vurdering. Jeg synes ikke det er rart om eleven har mye uro i kroppen når den sannsynligvis ikke forsto alt som blir sagt. Denne eleven ble testet med WISC-III, men kunne nesten ikke svare for seg. Dette gjorde et sterkt inntrykk på meg.

Praksiserfaring fra PPT og diskusjoner med medstudenter viste en utstrakt bruk av testing i utredningsarbeidet. Testene som ble brukt var statiske og normert for norske elever. Elever fra språklige minoriteter skåret under gjennomsnittet på slike tester, derfor mente jeg at statiske tester undervurderte potensialet hos elever fra språklige minoriteter. Jeg mente at elevene skåret lavt fordi de ikke forstod alt som ble sagt i testsituasjon. Problematikken med testing av elever fra språklige minoriteter opptok meg og jeg diskuterte dette mye med andre studenter. Det måtte da finnes tester som ikke kun målte hva eleven hadde lært til nå?

Jeg fikk to artikler av en medstudent, disse mente hun at burde være av interesse for meg. Det var en artikkel av Hansen (2000) "Hva innebærer dynamisk testing" og en annen artikkel av Pihl (2002) "Sakkyndig konstruksjon av avvik og stigma i en interetnisk relasjon". Disse to artiklene gav meg ideen om hva jeg ville skrive om i min hovedoppgave. Jeg ville skrive om dynamiske tester og reduksjon av feilkonklusjoner i testing av elever fra språklige minoriteter. Ville faren for feilkonklusjoner bli redusert ved å ta i bruk dynamiske tester i utredningen av elever fra språklige minoriteter?

Jeg tok kontakt med den ene artikkelforfatteren, Hansen, for å høre om han kunne gi meg tips om relevant litteratur om dynamiske tester. Hansen kunne mye om tema og jeg fikk flere boktips av han. De fleste bøkene Hansen anbefalte meg fantes ikke på universitetsbiblioteket og måtte bestilles fra utlandet. Dette tolker jeg slik at dynamiske tester ikke kan være et stort tema i Norge.

Den andre artikkelforfatteren, Pihl, tok jeg også kontakt med for å høre om hun hadde skrevet mer om tema. Jeg fikk Pihl sin publikasjonsliste slik at jeg kunne lese andre artikler hun hadde skrevet.

I samarbeid med min veileder, Øzerk, delte jeg opp det første utkastet til problemstilling i tre problemstillinger. Disse tre problemstillingen har fulgt meg gjennom søk i litteraturen og skriveprosessen. Takket være problemstillingene mener jeg at jeg har greid å holde meg til relevant litteratur.

#### PROBLEMSTILLINGENE.

1. Hvilke problemer innebærer dagens praksis med bruk av statiske tester på elever fra språklige minoriteter?
2. Hvordan kan man eliminere forekomsten av feilkonklusjoner i testing av elever fra språklige minoriteter?
3. Kan dynamiske tester redusere forekomsten av feilkonklusjoner og øke reliabel informasjon i utredningen av elever fra språklige minoriteter?

## 1.2 Språklige minoriteter.

Språklige minoriteter har i hovedsak vært omtalt som innvandrere eller immigranter. Vi hører ofte snakk om den økende innvandringen. Denne økende innvandringen fører blant annet med seg språklige minoriteter, fordommer og utfordringer. Historisk sett har det vært fordommer, som for eksempel troen på at innvandrere er mindre intelligente. Tester har vært brukt som dokumentasjon på at innvandrere var mentalt defekte. (Schultz og Shultz 1996).

---

”Since the time of Galton, mental abilities such as intelligence have been regarded as innate and stable. Standardized tests of mental ability have shown that, as a group, people from higher socioeconomic classes have higher scores than people from lower classes. One explanation for this phenomenon has been in terms of genetic differences”. (Samuda og Kong 1989:2).

Genetiske forskjeller som nevnes i sitatet over, ble brukt som bevis på at det var intelligens forskjeller mellom folkegrupper. De kulturelle variasjonene mellom folkegruppene ble ikke vektlagt. Det var kun resultatene fra testingen som ble vektlagt for å dokumentere intelligens forskjell mellom folkegrupper. Et eksempel på hvordan statiske tester ble brukt for å stemple immigranter som mentalt defekte var den utstrakte IQ testing av immigranter som kom til Amerika rundt 1920 tallet. Resultatet på IQ testen var avgjørende for om de fikk komme inn i landet. IQ testingen ble utført for å hindre at mentalt defekte mennesker innvandret til Amerika (Schultz og Schultz 1996). Det vil si at IQ testen ble sett på som et universelt verktøy som kunne sortere folk i grupper som mentalt defekte eller normalt fungerende. Tolkene som oversatte testens engelske spørsmål for innvandrerne som ikke kunne engelsk, reagerte på testbruken. Immigrantene som ikke kjenner den Amerikanske kultur kan umulig kunne svare på en del av spørsmålene hevdet tolkene. (Schultz og Schultz 1996). Tolkenes innvendinger ble det ikke tatt hensyn til.

Ved testing av innvandrere i Nord Amerika og Canada har det vist seg at de er dumme og de har ingen akademisk fremtid, (Samuda og Kong 1989) men det er ikke funnet empiriske bevis på at det er forskjellig IQ mellom raser. (Samuda og Kong 1989). Med utgangspunkt i dette hevder jeg at denne testingen på innvandrere ikke har holdt mål. Samuda og Kong (1989) har brukt begrepet ”rase”, dette er et vanlig begrep i USA. Implisitt i begrepet ”rase” ligger det at det er forskjellige menneskeraser, derfor brukes ikke dette begrepet i Norge. Her til lands bruker vi ”etniske minoriteter” når vi refererer til forskjellige folkegrupper.

Jeg har valgt å bruke WISC-III som et eksempel på statiske tester, da WISC-III er den mest brukte og studerte statiske testen ved utredning av barn. ”WISC-R är ett av de mest utforskade psykologiska testen och ett av de mest citerade testen i facitlitteraturen. Det er publicerats hundretals studier sedan 1974 och många ger stöd for validiteten i tolkningar og tillämpningar av WISC-R.” (WISC-III:54). Det kan stilles spørsmålsteget ved sitatet da det er WISC-III selv som sier dette om seg selv.

Det er ikke rart at det har vært skapt en tro på forskjeller i intelligensen da de fleste har blitt testet med WISC-R. (Pihl 2002. McCloskey og Athanasiou 2000). Språklige minoriteter som testes med WISC-R får ofte en skåre under 70IQ. (Pihl 2002. Hessels 1996). Med en IQ under 70 diagnostiseres man som mentalt retardert og barnet er derfor ikke opplæringsdyktig i vanlig skole. En IQ test alene er ikke nok for å stille diagnosen mentalt retardert, men det er en av kriteriene for diagnosen.( ICD-10). I stede for begrepet mentalt retardert brukes begrep som prestasjonsnivået er ekstremt lavt.(Flanagan og Kaufman 2004). Kaufman (1994) omtaler de som skårer under 69 IQ poeng som ”Intellektually deficient”(1994: 98).Det vil si at de har et mangelfullt intellekt. (Kunnskapsforlagets blå ordbøker 1996). Med andre ord kan man si at skårer man under 70 IQ poeng er det stor sannsynlighet for at barnet får spesialundervisning etter §5.1. i opplæringsloven.

I det internasjonale diagnosesystemet ICD-10 diagnostiseres man som ”Mild mentalt retardation” når IQ er under 70. I ICD-10 står det;

“Approximate IQ range of 50 to 69 (in adults, mental age from 9 to under 12 years). Likely to result in some learning difficulties in school. Many adults will be able to work and maintain good social relationships and contribute to society.“

IQ tester generelt er normert til en bestemt populasjon. WISC-III er tydelige på at språklige minoriteter ikke er en del av denne normeringen. ”Om barnet kommer från en annan kultur och/eller har bristfälliga kunskaper i svenska, kan testet knappast fungera så som det är avsett.” (WISC-III:89). Det vil si at slike tester ikke er pålitelig for språklige minoriteter.

---

### 1.2.1 Tospråklighet og intelligens

Hemmer tospråklighet den kognitive utviklingen? Sagt på en annen måte, er de tospråklige mindre intelligente? ”Det var først i begynnelsen av dette århundre at tospråklige ble gjort til gjenstand for undersøkelser...Hovedinteressen lå i å finne ut om tospråklighet hadde negative eller positive effekter på tenkning og intelligens målt i form av IQ”. (Øzerk 1992:132). Øzerk (1992) deler forskningen om tospråklighet og intelligens i to perioder, den negative- og den positive perioden. Den negative perioden strakte seg fra rundt 1900 til 1960. Etter 1960 regnes som den positive perioden. Med negativ periode menes en negativ utvikling av intelligensen ved tospråklighet. Forskning fra 1960 tallet har motbevist den negative forbindelsen mellom tospråklighet og intelligens. (Øzerk 1992). ”At det fremdeles finnes mennesker som tror på at tospråklighet fører til halvering av intellektuell utvikling, er det ingen tvil om”. (Øzerk 1992:134). Selv om det er motbevist at språklige minoriteter er dumme eller at det hemmer den kognitive utviklingen, er de fremdeles overrepresentert i spesialundervisning. (Tabell 1.)

### 1.2.2 Utbredelse av språklige minoriteter.

”Pedagogisk-psykologisk tjeneste får stadig flere henvisninger som gjelder elever med minoritetsbakgrunn.” (Pihl 2002c:604).

Er det slik at det er en større andel blant minoritetsspråklige elever enn det er blant total antallet elever, som får spesialundervisning etter § 5.1 i opplæringsloven?

Hvor mange elever får særskilt språkopplæring etter §2.8 i opplæringsloven?

Hvilke forskjeller er det mellom Osloskolene og skolene på landsbasis, når det gjelder fordeling av spesialundervisning etter §5.1 og særskilt språkopplæring §2.8?

Tabell 1. viser fordelingen av det total antallet elever og minoritetsspråklige elever som får Spesialundervisning etter § 5.1, og elever som får opplæring etter §2.8.

Tallene jeg har brukt for minoritetsspråklige elever i Osloskolen og på landsbasis kan ikke i alle tilfeller sammenlignes da de definerer og kategoriserer minoritetsspråklige på ulike måter. I Oslo opereres det med et tall for hvor mange minoritetsspråklige elever det er i grunnskolen. På landsbasis er det ikke et tall for antall minoritetsspråklige, men de defineres ut fra hvor mange som får tospråklig fagopplæring og særskilt norskopplæring.

Fordelingen av elever i spesialundervisning etter § 5.1 i Norge er det pr. 01.10.2004 totalt 620.620 elever i grunnskolen. Av det totale antallet elever er det 35.243 som får spesialundervisning etter §5.1. Det vil si at 5,67 % av elevene i den norske grunnskole får spesialundervisning etter § 5.1. Jeg sammenlignet fordelingen av elever som får spesialundervisning etter §5.1 på landsbasis med de som får det i Osloskolen.

Tabell 1. Fordelingen av det total antallet elever og minoritetsspråklige elever som får spesialundervisning etter § 5.1, og elever som får opplæring etter §2.8.

<b>Forelingen av elever i spesialundervisning etter § 5.1 opplæringsloven.</b>	Hele landet med Oslo	Hele landet uten Oslo	Oslo
Total antallet elever.	620 620	568 216	52 404
Total antallet elever som får tildeling etter §5.1?	35 243	32 300	2 943
Hvor mange % av den totale elevmassen får spesialundervisning etter § 5.1?	5,67 %	5,68 %	5,62 %
Hvor mange minoritetsspråklige elever får spesialundervisning etter § 5.1?	-----	-----	10,5 %

<b>Elever som får opplæring etter § 2.8. Særskilt språkopplæring for elever fra språklige minoriteter.</b>	Hele landet med Oslo	Hele landet uten Oslo	Oslo
Elever med morsmålsopplæring og/eller tospråklig fagopplæring	19 730	15 611	4 119
Hvor mange % av den totale elevmassen får morsmålsopplæring og / eller tospråklig fagopplæring?	3,18 %	2,75 %	7,86 %
Elever med særskilt norskopplæring (norsk som andrespråk)	35 807	24 918	10 889
Hvor mange % av det total antallet elever utgjør elever med særskilt norskopplæring?	5,77 %	4,39 %	20,78 %
Antall minoritetsspråklige elever i grunnskolen.	-----	-----	17 100
Hvor mange % av elevene i grunnskolen er minoritetsspråklige.	-----	-----	32,63 %

Prosent andelen elever som får spesialundervisning etter §5.1 er henholdsvis 5,67% for hele landet, 5,68 % for hele landet uten Oslo og 5,62% for Osloskolen. Med andre ord er andelen elever som får spesialundervisning etter §5.1 så og si likt for Osloskolene som for hele landet sett under et.

Hvor stor del av de som får spesialundervisning etter §5.1 er elever fra språklige minoriteter? Det er ikke lett å gi et enkelt tall til svar på dette spørsmålet. Gruppen av elever som går under betegnelsen språklige minoriteter er ikke en homogen gruppe, de varierer blant annet i hvor godt de kan det norske språket og hvor integrert de er i den norske kulturen. Grunnskolenes informasjonssystem ([www.wis.no/gsi](http://www.wis.no/gsi)) har en inndeling som gir et bilde av hvor mange elever som kommer fra språklige minoriteter. Inndelingene er “Elever med

morsmålsopplæring og /eller tospråklig fagopplæring” og “Elever med særskilt norskopplæring“.

Antallet elever som får særskilt norskopplæring utgjør på landsbasis 35.807 elever, det er 5,77 % av den totale elevmassen. 5,77 % av den totale elevmassen høres ikke så mye ut sett i forhold til andre grupper som har lære- og tilpassningsvansker i skolen.

Når jeg sammenlignet fordelingen av elever som får særskilt norskopplæring etter § 2.8 i Osloskolen med fordelingen på landsbasis kommer forskjeller til syne. Det er 5,77 % av elevene i Norge som får særskilt norskopplæring etter § 2.8 og 20,78 % av elevene i Osloskolen. Trekket Osloskolens tall ut av total antallet for hele landet blir forskjellen ennå større. I Osloskolen får 20,78 % av elevene særskilt norskopplæring og det er 4,39 % i landet forøvrig. Det vil si at Osloskolen har en betydelig høyere prosentandel med elever fra språklige minoriteter enn resten av landet. Å ha mange elever fra språklige minoriteter er i seg selv ikke ensbetydende med mange henvisninger fra denne gruppen til PPT.

Jeg sendte et par spørsmål til Utdanningsetaten i Oslo kommune for å finne ut:

- 1) Hvor mange elever fra språklige minoriteter får spesialundervisning etter § 5.1?
- 2) Hvor mange prosent av de som henvises til PPT er elever fra språklige minoriteter?

Svar 1).”Jeg har en intern statistikk om dette-den er ikke offentliggjort, men tallene for skoleåret per 1.10.2004 er som følger; 10,5 % av alle minoritetsspråklige elever har spesialundervisning.” (e-post fra utdanningsetaten 19.01.05. Navnet er anonymisert).

Svar 2).”Av ca 7500 saker i arbeid pr. år er min.språk andelen 35 % av henviste klienter til PPT. Jeg kan IKKE garantere for at alle disse elevene har norsk som annet-språk da denne statistikken kun er basert på elevenes opprinnelsesland.” (e-post fra utdanningsetaten 31.01.05. Navnet er anonymisert).

---

Svarene jeg fikk fra utdanningsetaten er klare i sin tale, det er flere elever fra språklige minoriteter som får spesialundervisning etter §5.1 enn norske elever. 10,5 % av elevene fra språklige minoriteter i Oslo får spesialundervisning etter §5.1 mens det totalt i Osloskolen er 5,62 % som får spesialundervisning.

Svar nummer 2 viser at 35 % av henvisningene PPT mottar er for minoritetsspråklige elever. Ser vi på prosentandelen minoritetsspråklige i Osloskolen som er 32,63 %, er differansen 2,37 %. Dette utgjør 1242 flere minoritetsspråklige elever som henvises til PPT. Det vil si at det er en overrepresentasjon av denne gruppen elever i henvisningene til PPT.

Overrepresentasjonen av elever fra språklige minoriteter mener jeg utgjør en betydelig andel av elevmassen. Hvis denne andelen blir testet med uegnede tester som igjen danner grunnlaget for tildeling av spesialundervisning etter §5.1 mener jeg det er fare for at mange elever får uegnete støttetiltak som ikke stemmer overens med de behov som elevene virkelig har. Faren for uegnete støttetiltak kan føre til stempling og feilplassering av elever som blant annet kan være med å opprettholde den sosiale reproduksjon. Eksempler på hvordan elever fra språklige minoriteter har vært utsatt for uegnede tester blir utdypet i kapitel 2, *Problemer med dagens praksis*. I kapitel 3, *Hvordan eliminere forekomsten av feil konklusjoner i testing av språklige minoriteter*, kommer jeg med forslag til tiltak som kan være med å forebygge feilaktig bruk av tester på elever fra språklige minoriteter.

Det har kommet en endring i loven som åpner opp for å gi elever fra språklige minoriteter spesialundervisning. Elever med rett til særskilt språkopplæring etter §2.8 kan også få spesialundervisning etter §5.1., ref. rundskriv F-008-04. Kap1.7.3. ”Når blir språket et problem for språklige minoriteter og hva skjer når problemet oppstår?”, omtaler dette rundskrivet.

Norge er ikke alene om å ha et økende antall språklige minoriteter som henvises for utredning av eventuelle lærevansker. Internasjonalt ser man også en stadig økning av språklige minoriteter i spesialundervisning. Hessels sier følgende ”...since the beginning of the eighties, the percentage of minority children in special education in the *Netherlands* has more than doubled.....”(2000:110). Cummins og Swain viser til den raske veksten av

minoritets språklige elever i *Canada*. De skriver blant annet; “.an increasing number of minority students are being referred for educational and psychological assessment.” (1989:184). Scheuneman og Oakland skriver at det er flere elever fra språklige minoriteter som henvises for utredning enn det er *Amerikanske* elever. “The rate of referrals for special educational services is generally higher for minority group children than for White children.” (1998:79).

### 1.2.3 Hva er språklige minoriteter

I vår dagligtale kan det være store variasjoner i hva hver enkelt av oss legger i forskjellige begrep. Variasjonen i dagligtalens begreper gir oftest ikke større innvirkning enn misforståelser mellom noen mennesker, mens det innen vitenskap er viktig å definere begreper. Har man avklart hva som menes med et begrep har man også avklart at man snakker om det samme. Vi snakker da om begrepsvaliditet.(Cozby1997). Begrepsvaliditet viser til at begrepet virkelig reflekterer det man ønsker å beskrive eller vise vitenskapelig. Innen vitenskapen kan svak begrepsvaliditet påvirke mange mennesker og føre til utilsiktet virkning.

Språklige minoriteter, minoritetsspråklige, fremmedspråklige, etniske grupper, fremmed kulturelle, utenlandske, tospråklige og så videre, mange navn på samme fenomen eller fritt valg av begrep? Sammenblanding av begrep uten kunnskap om hva som ligger i begrepene, er å tro at man snakker om samme fenomen, men legger forskjellige innhold / betydning i begrepet. Det skaper lett misforståelser og konflikter. Det er derfor viktig å være enige om hva som menes med begrepene og være konsekvent i hvilke begrep man velger å bruke. I lovttekster og offentlige dokumenter brukes begrepet ”språklige minoriteter”. Jeg har derfor valgt å bruke begrepet ”språklige minoriteter”.

I Opplæringsloven §2-8. Særskild språkopplæring for elever fra språklege minoriteter, står det: ”Elevar i grunnskolen med anna morsmål enn norsk og samisk har rett til særskild

---

norskopplæring til dei har tilstrekkeleg dugleik i norsk til å følge den vanlege opplæringa i skolen.” Det vil si at elever som har annet morsmål enn norsk og samisk er elever fra språklige minoriteter.

Lovteksten bruker begrepet morsmål som om det var klart hva som ligger i begrepet. Men hva som legges i begrepet `morsmål` er ikke alltid presist.(Øzerk 1992). Det jeg ser som en dekkende beskrivelse av morsmål er: Når språket har bygd opp et begrep for mor og det er mor som er den første som snakker til deg kalles det morsmål. (Skutnabb-Kangas 1981). Det vil si at førstespråket ditt er morsmålet ditt. Hva da med tilfeller der man har to førstespråk eller der man ikke lærer morens språk? Og tilfeller der man lærer morsmål et par år for så kun og omgås andrespråket samtidig som morsmålet blir helt fraværende. En enkel forklaring av morsmål finnes ikke.

Jeg har tatt meg den frihet å konsekvent oversette engelske begrep som ”bilingual”, ”second language learners”, ”ethnic minorities” og ”linguistically diverse” med ”språklige minoriteter”. Dette har jeg gjort fordi jeg har savnet klare begrepsavklaringer i artikler og bøker som har brukt de ovenfor nevnte engelske begrep.

Slik jeg har forstått det, blir språklige minoriteter i hovedsak fremstilt som om de er en homogen gruppe. Men språklige minoriteter er ikke en homogen gruppe. Jeg har valgt ut tre fenomen som er med å belyse mangfoldet i gruppen av språklige minoriteter.

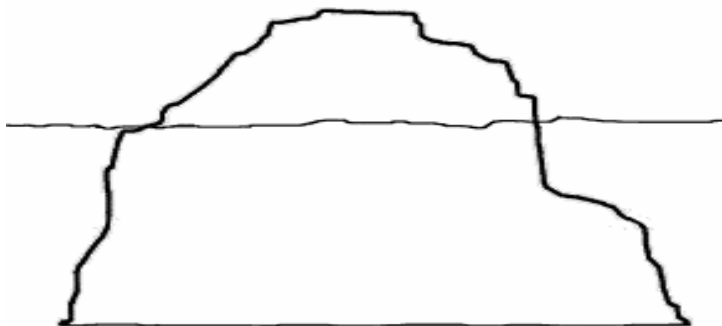
For det *første* er det en stor gruppe individer med sine særegne individuelle variasjoner.

For det *andre* er det forskjellige tospråklighetstyper.(Øzerk 1992). Tidspunktet for innlæring av morsmålet og andrespråket er avgjørende for tospråklighetstypen. Hvis innlæring av morsmål og andrespråket foregår samtidig kalles det for *simultant tospråklig* oppvekst. Læres derimot morsmålet først og andrespråket noen år senere kalles det for *suksessiv tospråklig* oppvekst. (Øzerk 1993). Simultant tospråklig oppvekst gir som regel en funksjonell tospråklighet, der begge språkene er godt utviklet, mens en suksessiv tospråklig oppvekst fører til et morsmåls - dominant tospråklighet, det vil si at morsmålet er best utviklet. I noen tilfeller er morsmålsinnlæringen så mangelfull at det er andrespråket som er best utviklet, da kalles det andrespråk - dominant. Tospråklighetstypen viser til hvilket av språkene som eleven behersker best. Selv om elvene har samme tospråklighetstype kan de ha varierende grad av språkforståelse.

Det *tredje* fenomenet er språkforståelsen hos språklige minoriteter. Språkforståelse er et komplekst fenomen (Øzerk 1992. McCloskey og Athanasiou 2000. Cummins 1989). Siden språkforståelsen er et komplekst fenomen vil det si at elever fra språklige minoriteter har forskjellig grad av språkforståelse. Språket kan deles i fire språklige delferdigheter. (Øzerk 1992) Å forstå, snakke, lese og skrive språket. En elev kan snakke språket uten å kunne lese, skrive eller forstå språket. Når det gjelder å lese språket kan det deles i å avkode og å forstå. Det vil si at eleven kan lese, avkode en tekst, uten å forstå noe av innholdet. I en test kan det i samtalen med eleven virke som om eleven kan kommunisere på språket men forstår eleven det den leser? For å illustrere hva vi ser og ikke ser i en samtale med en minoritetsspråklig elev, har jeg valgt en modell som viser språkets oppbygging.

### 1.2.4 Isfjellmodellen

Det finnes forskjellige modeller for hvordan man mener at språket er oppbygd og hvordan innlæringen av andrespråk påvirker eller drar nytte av førstespråket. (Bråten 1998. Øzerk 1992. McCloskey og Athanasiou 2000. Cummins 1987). Jeg har valgt å bruke isfjellmodellen som jeg mener gir en god illustrasjon av språkets oppbygging.

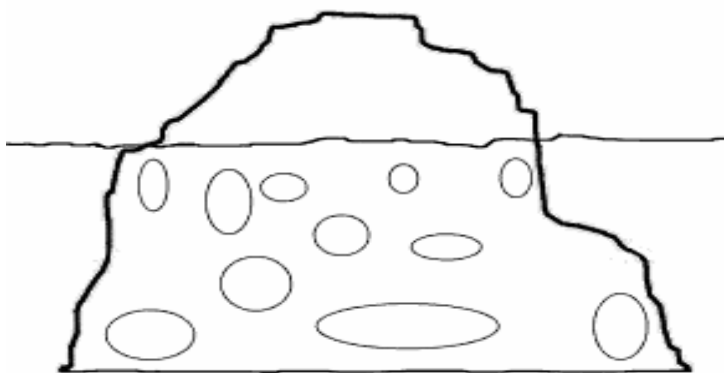


Figur 1. Isfjellmodellen

Isfjellet har en synlig og en usynlig del, det samme har språket. Den delen av isfjellet som er over vannet er den synlige delen av språket. Mens den usynlige delen av isfjellet er språkets usynlige del og fundamentet. Fundamentet består av forståelsen og erfaringsbakgrunn.

### **Sveitserost - metaforen**

Når man snakker med en språklig minoritets elev er det isfjellets synlige del man ser. Hva vet vi om elevens usynlige del? Kanskje den usynlige delen ser ut som en sveitserost? (Øzerk 2003). Sveitserost - metaforen viser til at det er et svakt språklig, begrepsmessig og kunnskapsmessig fundament hos en del tospråklige. Sveitserost - metaforen er et bilde på elevens problem når det er for lang utredningstid og tid for å bevilge resurser for å hjelpe eleven.



Figur 2. Sveitserost - metaforen

### **1.2.5. Erfaringsbakgrunn og sosiale forskjellig.**

Språket er ikke nøytralt, det er sosialt betinget. Det vil si at vårt morsmål er med å bestemme vårt verdensbilde. (Skutnabb-Kangas 1981). Å beherske et språk er en forutsetning for å kunne delta aktivt i samfunnet. (Øzerk 1992). Språket er avgjørende for i det hele tatt å få tilgang til mulighetene i samfunnet. Elever med svak språkforståelse har blitt testet og plassert slik at de ikke har fått tilgang til mulighetene i samfunnet. (Pihl 2002a. Samuda et.al.1989. Hessels 1996). Slik testpraksis og plassering av språklige minoriteter har vært

med på å opprettholde sosiale forskjeller i samfunnet. Opprettholdelse av sosiale forskjeller gjelder ikke bare for elever fra språklige minoriteter. Jeg ser det som like aktuelt for elever fra forskjellige samfunnslag.

Bernstein og Vygotsky har på hver sin måte beskrevet hvordan språket er sosialt konstruert. (Gundersen 2001).

Språkforskjeller og kulturelle variasjoner er med å opprettholde klasseskiller og sosial status. Selv innenfor samme språk er det store variasjoner mellom de som lykkes og ikke lykkes på skolen. Om eleven lykkes på skolen avhenger blant annet av hvilke kultur eller skal vi si erfaringsbakgrunn eleven møter skolen med. Passer elevens erfaringer med skolens kultur?

I hvilken grad er skolen dyktig til å tilpasse seg elevens behov og gi dem muligheter til å realisere sine potensialer, slik at eleven får lære og utvikle seg i samsvar med sine potensialer i skolens læringsfellesskap?

Hvordan finne og kartlegge elevens problemområder slik at tiltak kan iverksettes? Tester er et av verktøyene som ofte benyttes i kartleggingen av elevens problemområder. Neste avsnitt tar derfor for seg test og testing.

## 1.3 Test og testing

Det finnes mange former for tester, men i denne oppgaven er det kun snakk om psykologiske tester.

Jeg ser det som viktig å skille mellom begrepene test og testing. En test er et verktøy som brukes ved testing. Eksempler på tester er WISC-III, Leiter-R og LEM. Testen ser jeg som et

---

nøytralt verktøy, da den har klargjort i testmanualen hva den er ment å skulle måle. Når testen tas i bruk for å teste en elev kalles det testing. Tester som henges ut burde fått kritikk for hvordan de er brukt og ikke kritikk av selve testen. (Jackson 2002). Testmisbruk kan skade testens troverdighet. Testingen ser jeg ikke som nøytral, fordi den er utsatt for større variasjoner og det er her kritikken burde vært rettet. Når testen tas i bruk av testadministrator er det ikke garantert at testen alltid er i like kyndige hender. Testadministrator er ikke en ensartet person. Utdannelse, trening og erfaring er noen av faktorene som er av avgjørende betydning for hvordan testingen utføres. Testingen starter allerede når testadministrator velger ut testen som skal brukes på eleven. Her ved valg av test kommer vi til den første faren for at testresultatene får bias. Velges tester som ikke er egnet for den eleven som skal testes? Testadministrator må ha kjennskap til testene og den eleven som skal testes for å sikre at det velges en egnet test. Jeg mener det er under testingen at fare for bias i testresultatene er størst. “Derfor har det ikke stor verdi at diskutere test uten at tåge i betraktning, hvordan de anvendes.” (Jackson 2002:27)

### 1.3.1 Hva er en test?

En test er et verktøy som kan måle forskjellige egenskaper og ferdigheter i et individ, mellom individ eller i en gruppe. Det har vært en nesten blind tro på at det testen måler er en sannhet, men tester er kun en konstruksjon vi har skapt som hjelp til å forstå og systematisere verden. Vi klassifiserer og kategoriserer for å skape orden i verden slik at verden blir lettere å forstå. (Flanagan, McGrav og Ortiz 2000).

Tester består av en samling oppgaver som skal være representativt for det man ønsker å vite noe om. Testens oppgaver skal være et representativt utsnitt av de egenskaper man er interessert i å undersøke. Jackson (2002) definerer test som; “En systematisk metode til at beskrive adferd og bagvedliggende egenskaper ved anvendelse af numeriske skalaer eller fastlagte kategorier.” (Jackson 2002:226). Jeg tolker denne definisjonen slik at en test er en systematisk metode som kommer frem til noen tall som sier noe om eleven. Tester kan sees

på som nøkler til å finne eller forstå elevens skolevansker. Derfor er det viktig at nøklene brukes på en riktig måte.

En test kan også sees på som en systematisk, standardisert og objektiv stikkprøve av personens evner og / eller ferdigheter. Med standardisert menes at testen skal være lik for alle som blir testet, uavhengig av testadministrator. Objektivitet viser til at det er overensstemmelse mellom en persons skårer uavhengig av testadministrator. Når en test er lik for alle kan det hevdes at testen er rettferdig. Testmanualen og trening i testen skal være med å sikre at alle som testes med en bestemt test blir behandlet på samme måte. Dette er den ideelle versjonen, men virkeligheten er ikke så enkel. Selve testingen er ikke en enkel mekanisk prosess. Her er det mange faktorer å ta stilling til. Det først man må ta stilling til er hvem som skal testes og hvorfor skal denne eleven testes.

### **1.3.2 Hvorfor teste?**

Når en test tas i bruk bør man stille seg spørsmålet; Hvorfor teste og hva er hensikten med å anvende akkurat denne eller disse testene? Tester er utviklet for å måle våre ferdigheter og eller prestasjoner. Man bør ha avklart disse spørsmålene før man begynner testingen.

Tester er en del av dokumentasjonen de sakkyndige bygger sine slutninger og anbefalinger på. Tester spiller en viktig rolle når det gjelder legitimeringen av lærevansker eller for å stille en diagnose.

“De fleste er først og fremmest interessert i at anvende test som hjelp til at træffe beslutninger. Når man tager beslutninger, forudsiger man egentlig noget om fremtiden.” (Jackson 2002:38).“Hovedårsagen til at anvende test er, at vi formoder, at resultatene vil være til hjælp til at tage en beslutning. Dvs. vi tror, at de informasjoner, vi får fra testscoren, vil hjelpe os med at tage en eller anden slags forutsigelse.” (Jackson 2002:75).

Barns fremtid kan være bygget på testresultat, derfor er det ekstra viktig å vite hva testen kan og ikke kan si noe om.

Uten testing kunne man hevde at de sakkyndige slutningene hadde vært ren synsing. Synsing er noe alle og en hver kan gjøre, men synsing kan ikke dokumenteres. Selv om det kan være fare for feilslutninger ved bruk av tester, kan det være enn nesten større sjanse for feilslutninger ved ikke å bruke tester. Sakkyndige vurderinger er ikke narrativer, det vil si historiefortellinger (Pihl 2002). Poenget er å vite hvordan testen skal brukes. Det er viktig å se på helheten og ikke en test alene. En enkelt test isolert gjør mer skade enn nytte. (Skutnabb-Kangas 1981).

Fordi tester har vært brukt på feilaktig måte (Schultz og Schultz) har det ført til en skeptisk holdning til tester. Det er forskjellige holdninger til testing av språklige minoriteter. Jeg mener Skutnabb – Kangas (1981) har en negativ og pessimistisk beskrivelse av problemet med testing av språklige minoriteter.

”Tests are unreliable, they give a distorted picture of reality, they reduce the individual, they can never measure or capture what one is capable of doing with language in real situations, and tests can never catch all the important features anyway. The devising and use of tests is something one ought not to get involved with.” (1981:217).

Hva slags verktøy skal man da ta i bruk ved utredninger av minoritetspråklige elever? Man behandler ikke feber ved å kaste termometeret.(Jensen 1980).

### **1.3.3 Hva måler testen?**

Det finnes ulike typer tester. Noen tester måler elevens generelle evner og eller ferdigheter mens andre måler spesifikke evner og / eller ferdigheter. Jeg har her valgt å bruke begrepet evnetester og ikke intelligens tester fordi `intelligens` begrepet er verdiladet. (Rust og Golombok 1994). Felles for testene er at de prøver å si noe om elevens funksjonsnivå på et eller flere områder. Jeg har i denne oppgaven valgt å forholde meg til tester som skal måle generelle evner. Generelle evnetester som blir nevnt i denne oppgaven er; WISC-III, Leiter og LEM. De generelle evnetestene kan sees på som en screening av elevens evner og ferdigheter, derfor har jeg valgt å forholde meg til disse i oppgaven. Jeg ser denne screeningen som en god måte å bli bedre kjent med eleven på. Deltestene kan gi hint om mer spesifikke problemområder. Hvis noen deltester utpeker seg mer enn andre kan det være indikasjon på spesifikke lærevansker hos eleven. De deltestene som upeker seg bør derfor undersøkes videre. "Ingen test skal stå alene, og det er testbrugerens ansvar, at dette ikke sker i praksis,...." (Jackson 2002:13). Tar man en avgjørelse på grunnlag av en enkelt test er det fare for at avgjørelsen kan inneholde bias. Det er viktig å ta i bruk flere kilder før avgjørelser om eleven tas, da det kan påvirke elevens fremtid. Det vil si at de målte resultatene fra en test kan ha avgjørende betydning for elevens fremtid. Blir eleven stemplet som mentalt retardert, dyslektiker, hyperaktiv eller normalt fungerende? Hvilke stempel eleven får på seg har betydning for hvilke tilbud eleven får i skolen. "En test kan likeledes være nyttig i en utdanningsmessig sammenheng, hvis det innebærer, at barnene får en mer hensigtsmessig undervisning, og at skolen er i stand til at udnytte undervisningsressourcerne på en mer effektiv måte." (Jackson2002:41). Skal en test vær nyttig slik som sitatet til Jackson sier, er det viktig at testen utføres på en korrekt måte av kompetente personer.

#### **1.3.4 Test kompetanse.**

"Ingen test er bedre end den person, som benytter den,...." (Jackson 2002:13). Hvis ingen test er bedre enn den som benytter den, ser jeg det som veldig viktig at de som benytter en test er kompetente. Hvordan kan man bli kompetent til å teste? Det jeg ser som de viktigste punktene for å være en kompetent testadministrator er utdanning, test trening og lese testmanual. Jeg mener det også er viktig å sikre seg at testadministrator har formell

---

kompetanse. Dette kan gjøres ved å stille krav til formell kompetanse hos de sakkyndige i PPT. Se kap.1.7.6 kompetansekrav til PPT.

## 1.4 Psykometri

Mye av testkritikken som kommer frem i media mangler informasjon om psykometriske kvaliteter. (Jensen 1980). Det er en gammel kilde jeg har brukt som referanse for denne påstanden. Jeg mener referansen er like gyldig den dag i dag. Psykometri gjelder for alle tester. Det er en av grunnene til at jeg har valgt å ta med et avsnitt om psykometri. Avsnittet er ment å tydeliggjøre hva seriøse tester må forholde seg til. Dette avsnittet skal også være med å støtte opp under testkvaliteten, samt vise at de testene jeg bruker i oppgaven oppfyller psykometriske krav.

Rust og Golombok (1999) definerer psykometrien som den delen av psykologien som driver med målbare faktorer. ”There are good assessments, and bad assessments, and there is a science of how to maximize the quality of assessment. That science is psychometrics. There is no other aspect of the field of psychology that has such an impact on individuals in their daily lives”. (Rust and Golombok 1999:4). Med andre ord kan man si at psykometri er vitenskapen som måler psykologiske fenomen. Psykologiske fenomen er ikke av den håndfaste typen slik det er i fysikkens verden. Fysikken forholder seg til eksakte enheter som lengde og vekt. Det er ikke stor uenighet i verden om hvordan lengde og vekt måles. Derimot er det ikke full enighet om hva psykologiske fenomen som intelligens og lærevansker er, og hvordan de måles. Psykologiens enheter er begreper. (se avsnitt 1.4.3. begrepsvaliditet). Siden begreper ikke er eksakte enheter, må psykologiske tester finne en måte å sikre testkvaliteten på. For å sikre testkvaliteten har vi psykometri, det handler om hvordan gjøre

testene så gode som mulig. Følges de psykometriske parametere gir det et bilde av at det ikke er selve testen det er noe galt med, hvis testresultatene inneholder bias.

Alle tester bør ha en testmanual som sier hva testen er ment å måle og hvem den er beregnet på. Testmanualene omtaler også testens psykometriske aspekter. Psykometrien har flere aspekter som er med å øke kvaliteten på tester. To av de viktigste aspektene ved en tests kvalitet er reliabilitet og validitet. Kan man stole på testresultatene og måler de det de er ment å måle?

### **1.4.1 Reliabilitet**

Oversettes det engelske begrepet `reliabel` til norsk, får man ordet `pålitelig`. Da kan man enkelt sagt si at tester som er reliable er pålitelige. (Dette er en forenkling av begrepet.) Det betyr at en test man kan stole på, måler det samme hver gang. I testmanualene skal man kunne finne hva testens reliabilitet er og hvordan den er beregnet. Da vet vi hvor pålitelig testresultatene er. IQ tester generelt viser til en reliabilitet på 0.9 (Rust and Golombok 1999). Det er høy reliabilitet.

Testen kan være reliabel uten å være valid, men den kan ikke være valid uten at den samtidig er reliabel. Det betyr at testen kan måle det samme hver gang, men den måler ikke det den sier den skal måle. Vi har eksempler fra testingens begynnelse på 1900-tallet, der den britiske vitenskapsmannen Galton målte sensorisk kapasitet og mente det viste til personens mentale kapasitet. Galton fikk samme resultat hver gang han brukte sine psykologiske instrument, testen var reliabel. Problemet oppstod når sensorisk kapasitet og mental kapasitet ikke korrelerte, testen var ikke valid. (Schultz & Schultz 1996). Ustabile og ikke-reliable testskårer er den mest fatale testbias. (Jensen 1980).

### **1.4.2 Validitet**

Validitet er det mest sentrale begrepet i hele testforetaket. (Jensen 1980). Vi så Galton som mente at sensorisk kapasitet kunne vise til mental kapasitet. (Schultz & Schultz 1996).

Hadde Galton fått testresultatene til å korrelere med den mentale kapasiteten hadde testen hans vært valid. Men testen hans målte ikke det den sa den skulle måle. Validitet handler om gyldigheten av det testen sier den skal måle. Hvordan finne ut om testen måler det den sier at den måler? Det er flere typer validitet. (Rust and Golombok 1999). Validitetstypene viser til forskjellige områder av testen. Jeg har valgt ut det jeg ser som de viktigste validitetstypene i tester for elever. Det er begrepsvaliditet og prediktiv validitet.

### **1.4.3 Begrepsvaliditet**

Begrepsvaliditet viser til om den operasjonaliserte definisjonen av variabelen reflekterer variabelens teoretiske mening. (Cozby 1997). Innen forskningen kaller man objekter,

hendelser, fenomen eller situasjoner for variabler. Variabler er abstrakte begrep som må operasjonaliseres slik at de kan studeres. Begrepene `intelligens`, `kognitive evner` og `lærevansker` må operasjonaliseres for å være sikker på hva som legges i begrepet.

Intelligenstester har vært kritisert for ikke å måle det de er ment å skulle måle. Slik kritikk mener jeg ofte kunne vært unngått hvis det hadde vært enighet og forståelse for hva som legges i begrepet. Noen ord kan være så verdiladet at fagfeltet velger å bytte begrepet med mindre verdiladde ord.

”Once these problems with the way the concept of intelligence was being used were recognized, several strategies were developed to overcome them. First, the word `ability` was generally substituted for `intelligence` in existing tests. This is more than just a semantic quibble. Ability is an accurate word for the description of the trait that IQ tests measure, is less far-reaching in its implications, and does not carry with it the automatic connotations of `intelligence` as moral worth. Second, attention focused increasingly on what a person with high ability was actually able to do - the term was used transitively rather than intransitively”. (Rust and Golombok 1994:21).

Her ser vi et eksempel på at fagfeltet velger å endre begrep som de ser som stigmatiserende og hemmende for de som blir testet. Jeg tolker det slik at begrepet `evner` har i seg muligheten til å si noe mer om personens fremtidsutsikter enn begrepet `intelligens`. (Fremtidsutsikter, se neste avsnitt om prediktiv validitet). Med andre ord kan man si at `evner` er et dynamisk ord mens `intelligens` er et statisk ord. ”When ability rather than intelligence is considered, it seems much more apparent that the environment, in particular school and university, must have an impact on what a person is able to do.” (Rust and Golombok 1994:21). Hva skolen tror at eleven er i stand til å gjøre er av stor betydning for elevens fremtidsutsikter.

#### 1.4.4 Prediktiv validitet

Prediktiv validitet sier noe om fremtiden. Dette kan vises ved korrelasjon mellom testskåren og hvor godt eleven gjør det i et spesifikt område. Eksempel på prediktiv validitet kan være; "Intelligence tests certainly measure something, and this something is highly correlated with academic achievement". (Rust and Golombok 1999:19). Skårer eleven høyt på intelligenstester er det stor sannsynlighet for at eleven gjør det bra på skolen. Her kan man stille seg spørsmålet om det er høna eller egget som kom først; Er det høy IQ som gir gode skolerresultat, eller er det gode skolerresultat som gir høy IQ?

#### 1.4.5 Bias

Bias i en test viser til systematiske feil. (Jensen 1980). De systematiske feilene påvirker noen testpersoner mer enn andre. Bias er et begrep som kan gjelde for enkelte testspørsmål eller for hele testen. (Jackson 2002). Mange mennesker tenker på bias kun i forbindelse med diskriminering av visse grupper. Det er da snakk om kulturell bias. Kulturen elevene har vokst opp i kan være av så forskjellig art at en elevs erfaringsbakgrunn virker negativt inn på testen. Testresultatet gjenspeiler da ikke elevens virkelige kunnskaper og ferdigheter. Tester har som oppgave å diskriminere. Det vil si at tester skal vise individuelle forskjeller, innenfor en bestemt populasjon, nærmere bestemt den populasjonen som ble brukt for testens normeringsgrunnlag.

Jeg mener at hvis normeringsgrunnlagets populasjon ikke er representativt for elevens oppvekstkultur bør testen ikke brukes på eleven. Om testen blir brukt på eleven må

testadministrator være seg dette bevisst. En slik bruk av testen som avviker fra testmanualen må komme tydelig frem i testrapporten / den sakkyndige vurderingen.

Hvis testadministrator konsekvent tester elever med en test som ikke er egnet for en gitt gruppe elever kan vi snakke om systematiske feil, det vil si bias.

### **1.4.6 Bias og validitet**

I dette avsnittet ønsker jeg å vise hvor viktig det er å lese testmanualen nøye. To hypoteser kan gi en og samme test med lik skåre enten testbias eller testvaliditet. Det vil si at en test kan være både valid og ha test bias, men det er testmanualen som avgjør hvordan man skal forstå testen. Dette kan illustreres med to hypoteser.

Hypotese 1: Minoriteter skårer lavere en majoriteten på denne testen.

- Testen har bias for minoriteter.

Hypotese 2: Det er forventet at minoriteter skårer lavere på denne testen.

- Minoriteter skårer lavere, det vil si at testen har validitet.

Disse to hypotesene mener jeg er et godt eksempel som viser viktigheten av å lese testmanualen før testingen og påfølgende tolkning settes i gang. I testmanualer står det hva testen er ment å skulle måle og hvem testen er beregnet på. For eksempel står det i WISC-III manualen at testen er et psykologisk måleinstrument som bedømmer den intellektuelle kapasiteten hos svensktalende barn i alderen 6 til 16 år. (Det står svensktalende barn fordi Norge bruker den svenske manualen som bygger på svensk normeringsgrunnlag). Kritiseres WISC-III for å ha bias for minoriteter, hypotese 1, kan det hevdes at kritikken er fremsatt på et sviktende grunnlag. Manualen har redegjort for at testen ikke er beregnet på elever som ikke behersker språket. Det kan dermed hevdes at hypotese 2, det er forventet at minoriteter skårer lavere på testen, er den aktuelle hypotese for WISC-III.

---

“The fact that tests are not biased, however, does not guarantee that they will be used properly or wisely.” (Jensen 1980:715). Videre sier Jensen at det er lettere å finne feil bruk av testen en feil ved selve testen.

“The main faults with mental testing generally involve abuses and inadequacies in the use of tests rather than bias or other psychometric deficiencies in the tests themselves.” (Jensen 1980:738).

“Most current standardized tests of mental ability yield unbiased measures for all native-born English-speaking segments of American society today, regardless of their sex or their racial and social-class background.” (Jensen 1980:740).

### 1.4.7 Standardisering og normering

Standardisering av en test er ment å skulle sikre at testpersonene blir behandlet på en og samme måte slik at det ikke blir forskjellsbehandling. Testmanualene har en beskrivelse av hvordan testen skal administreres, utregnes og tolkes. Dette skal sikre at alle som blir testet blir behandlet på samme måte. Problemer kan oppstå når testadministrator ikke har satt seg skikkelig inn i testmanualen og bruker testen på en uhensiktsmessig måte. For å kunne sammenligne testresultatene er det en forutsetning at testen utføres på den foreskrevne måten, slik at alle blir behandlet likt.

Testens skårer sier noe om hvordan testpersonen skårer i forhold til en gitt populasjon. Tester er utviklet med forskjellige målpopulasjon i tankene. Det vil si at en bestemt populasjons skåringer blir normen som andre måles i forhold til. Resultatene på en test er tall som kalles råskåren. Råskåren regnes om til mer håndterlige tall som kalles standardskåren.

Standardskåren på IQ tester er normalfordelt. (Jackson 2002). Standardskåren til en person kan si noe om hvordan denne personen skårer i forhold til normalfordelingen. Ligger personens skåre over eller under gjennomsnittet eller skårer personen som gjennomsnittet?

Skårer eleven langt under gjennomsnittet er det en indikasjon på at eleven kan ha en spesifikk eller generell lærevanske.

## 1.5 Intelligens og menneskesyn.

I dette avsnittet vil jeg se på hva som legges i begrepet intelligens som intelligens- / evnetester måler.

Hva er intelligens? Hva er kilden til intelligens? Kan intelligens måles? Disse spørsmålene har fått forskjellige svar og de har vært og er mye diskutert. Når det gis svar på de ovenfor nevnte spørsmålene om intelligens, gjenspeiler svarene et grunnleggende menneskesyn. Det menneskesynet som kommer frem er av avgjørende betydning for hvilke utviklingsmuligheter man tror at eleven har. Nygård (1993) skriver om mennesket som aktør eller brikke. Er mennesket en aktør som kan ta styring i eget liv eller er det en brikke som blir styrt av de ytre forhold? *Jeg tolker Nygårds aktør og brikke begrep til å passe det jeg kaller et dynamisk og statisk menneskesyn.* Der aktøren representerer det dynamiske menneskesynet og brikken det statiske menneskesynet. Det statiske menneskesynet ser ikke på mennesket som et modifiserbart vesen slik det dynamiske menneskesynet gjør. Det statiske menneskesynet ser intelligens som en medfødt og fast egenskap. Mens det dynamiske menneskesynet ser intelligens som en fleksibel egenskap hvor det medfødte utvikler seg i samspill med sine omgivelser. Det vil si at et statisk menneskesyn setter eleven og dens fremtidsutsikter i en bås som ikke kan endres. Utviklingsmulighetene til en elev i det dynamiske menneskesynet gir rom for endring. Jeg mener det dynamiske menneskesynet bør være gjeldene, slik at dårlige oppvekstvilkår ikke skal være et hinder for videre utviklingsmuligheter.

---

For å vise noe av bredden i intelligensdebatten har jeg valgt ut 3 representanter som har skrevet om intelligens. Disse tre representantene mener jeg gir et innblikk i debatten om intelligens. Høsten 1994 utkom boken "The Bell Curve" (Herrenstein og Murray). Denne boken argumenterer for at intelligens går i arv. Det vil si at når intelligens går i arv kan det hevdes at intelligens, IQ, er å finne i genene. "The Bell Curve", omhandler hvordan intelligens virker inn på alt fra skoleprestasjoner, ekteskapsproblemer, arbeidsledighet, kriminalitet og foreldrerollen. Det vil si at boken prøver å forklare alle sider av det Amerikanske samfunnet ut i fra intelligens. (Herrnstein og Murray 1994). Dette synes jeg er en litt vel enkel forklaring på samfunnsforhold. Jeg kan ikke se at Herrnstein og Murray har tatt stilling til andre grunner for den sosiale reproduksjon en arv. Eksempel på at feilaktig bruk av tester som kan stemple elever og forhindre fremtidige utviklingsmuligheter kan jeg overhode ikke se at de har vurdert. Jeg mener Herrensteins og Murray har et statisk menneskesyn.

Det er flere som reagerer sterkt på Herrnstein og Murry's bok. De to andre representantene jeg har valgt ut, er kritikere til boken "The Bell Curve". Gould (1997) var en av de som reagerte svært kritisk til "The Bell Curv". Som motsvar kom Gould med boken "The mismeasure of man" i 1997. Gould har et eget kapittel som slakter begrunnelsene som "The Bell Curve" bygger sine påstander på. Slik jeg ser det tar Gould (1997) et oppgjør med den biologiske determinismen og arvelighets forkjempere. Han slår et slag for fleksibilitet. Jeg mener Gold har et dynamisk menneskesyn.

Den andre som reagerte kritisk til "The Bell Curve" var Gottfredson. Gottfredson's (1997) reaksjon til "The Bell Curve" endte i en 25 punkts liste om hva som var hovedstrømmen innen intelligensforskning. Listen ble undertegnet av 52 anerkjente forskere.

Jeg vil nå ta en gjennomgang av noen punkter (1, 2, 3, 5, 7, 9) i Gottfredson's 25 punkts liste, for å vise hva som menes med intelligens i dag. Punktene er valgt ut fra hvor relevante de er for denne oppgavens problemstillinger.

I det første og andre punktet defineres intelligens som en generell og sammensatt kapasitet som kan måles.

1. "Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather it reflects a broader and deeper capability for comprehending our surroundings -- "catching on," "making sense" of things, or "figuring out" what to do."

2." Intelligence, so defined, can be measured, and intelligence tests measure it well...".(Gottfredson 1997:14).

I punkt tre vises det til at det er flere typer intelligens tester og at de alle måler den samme intelligensen.

3." While there are different types of intelligent tests, they all measure the same intelligence. Some use words or numbers and require specific cultural knowledge (like vocabulary). Other do not, and instead use shapes or designs and require knowledge of simple, universal concepts (many/few, open/closed, up/down)."(Gottfredson 1997:14).

Det vil si at de tre testene, WISC-III, Leiter og LEM, jeg har valgt som eksempler i denne oppgaven måler intelligens. Noen tester krever spesifikk kulturell kunnskap. Andre tester krever at man kan språket som testen bruker, dette nevnes i punkt fem.

Punkt fem viser til at intelligens tester ikke har kulturell bias for engelsktalende svarte.

5." Intelligence tests are not culturally biased against American blacks or other native-born, English- speaking peoples in the U.S." (Gottfredson 1997:14).

---

Jeg ser det som viktig at `engelsk-talende` er spesifiser. Jeg mener at språket er en av hovedårsakene til at intelligenstagere blir anklaget for å ha kulturell bias.

I korte trekk kan man si at det er viktig å lese testmanualen for å vite hvem testen er beregnet for.

I punkt syv avviser de det som var grunnlaget for tidligere tiders rasehygiene.

7. "Members of all racial-ethnic groups can be found at every IQ level. The Bell Curve of different groups overlap considerably, but groups often differ in where their members tend to cluster along the IQ line." (Gottfredson 1997:14).

9." IQ is strongly related, probably more so than any other single measurable human trait, to many important educational, occupational, economic, and social outcomes." (Gottfredson 1997:13-14).

Det vil si at IQ er av betydning for det utbytte eleven har av skolen. Av dette kan man dra følgende slutning om at en elev som får en høy IQ på en evnetest vil gjøre det bra i skolen. Blir en elev testet med en uegnet test er det fare for at eleven skårer lavere enn det ville gjort med en egnet test. På den måten blir elevens fremtidsutsikter påvirket av testens resultat. Fordi et lavt IQ skåre kan virke stigmatiserende og føre til et selvoppfyllende profeti.

Siden elevens fremtidsutsikter blir påvirket av testresultatet, mener jeg det er veldig viktig at det er et dynamisk menneskesyn som ligger til grunn for intelligensbegrepet. Slik at eleven får muligheter til å endre og utvikle sitt potensial.

## 1.6 Presentasjon av testene, WISC-III, Leiter-R, LEM.

Hvorfor jeg har valgt WISC-III, Leiter-R, og LEM?

Wechsler Intelligence Scale for Children-III (WISC-III), Leiter-R og Learning Potential Test for Ethnic Minorities (LEM) er alle evnetester. Disse tre testene måler en elevs evnenivå på forskjellige måter. WISC-III og Leiter-R er eksempler på statiske tester mens LEM er eksempel på dynamisk test. Jeg ønsker å se på de testene som brukes mest ved utredning av elever i Norge. WISC-III og Leiter-R er de mest brukte evnetestene ved norske PP-kontor. (Kapittel 2.1. analyse av sakkyndige vurderinger). Testene kan betraktes som to ytterpunkter der WISC-III regnes for en verbaltest, mens Leiter-R regnes for en nonverbaltest. Begge testene er standardiserte og normerte og gir dermed en indikasjon på hvordan eleven fungerer i forhold til sinne medelever.

Jeg har pr. dags dato ikke funnet noe PP-kontor som bruker dynamiske evnetester. LEM har jeg valgt fordi den er spesielt utviklet for etniske minoriteter. Det kan hevdes at LEM er en nonverbaltest fordi den ikke tar i bruk forståelsen av ord, men kun bruker non-ord. Jeg mener at LEM ikke er en nonverbal test fordi den lager og bruker nye ord for objekter som er uten gjenkjennelse. Har eleven språkvansker og brudd i den fonologiske lupen (Badderley 2002) vil den ha vansker med å holde nye ord i kortidsminnet, derfor mener jeg at LEM tester verbale evner.

Jeg vil nå gi en kort redegjørelse for WISC-III, Leiter og LEM. Testene blir redegjort for etter følgende spørsmål.

- \* Når ble testen utviklet?
- \* Hva er bakgrunn for at testene er laget?
- \* Hva måler testene?
- \* Hvem er testen beregnet for?
- \* Er normeringsgrunnlaget som brukes på norske elever, laget for andre land?
- \* Hvordan er testens oppbygging?
- \* Brukervennlighet, administrasjon og tidsbruk?
- \* Hva er testens reliabilitet og validitet?

---

Jeg har brukt testmanualene til de respektive testene som kilde for redegjørelsene av testene i dette kapitlet. Når det gjelder LEM testen har jeg også basert meg på artiklene til Hessels (1996, 1997, 2000, 2002) som kilder for redegjørelsen av LEM. LEM testen har forlaget sluttet å produsere fordi den ikke solgte nok. Jeg har fått tilgang til en Nederlandsk versjon av LEM testen.

På slutten av dette kapitlet er det satt opp en tabell, tabell 1.3, som skjematisk viser sammenligning av WISC-III, Leiter-R og LEM.

### **1.6.1 WISC-III**

WISC-III har sine røtter tilbake til 1930 tallet. Wechsler utviklet WISC fra Wechsler-Bellevue Intelligence Scale som var beregnet for voksne. Denne skalaen hadde egenskaper som gjorde den individuelle administreringen av intelligenstesting enklere enn de tester som var tilgjengelige på den tiden. I dag er det den tredje versjonen av WISC, WISC-III, som brukes. Grunnlaget for revideringen av testen har blant annet vært å oppdatere normene og “modernisere” innholdet i deltestene. I Norge har psykologiforeningen valgt å bruke den Svenske utgaven av WISC-III testmanual. Jeg mener det er en svakhet at Norge ikke har en norsk utgave av testmanualen. Svensk og norsk er to relativt like språk, men det er ikke alle de svenske ordene jeg forstår, noe jeg sannsynligvis ikke er alene om. Det skal sies at skåringsprotokollen finnes i en norsk utgave. Da WISC-III er den mest brukte psykologiske testen for barn mener jeg at også testmanualen burde vært oversatt til norsk, for å forhindre missforståelser.

”Wechsler Intelligence Scale for Children (WISC), som nu föreligger i sin tredje version (WISC-III), är ett av världens mest använda psykologiska test för barn. Det främste skälet till att WISC-testet fått en så stor spridning är att det ger ett tämligen brett och stabilt mått på barns intelligens, eller annorlunda uttryckt, deras kognitiva funktionsnivå. Det finns ett positivt samband mellan intelligenstestpoäng och hur man klarar utbildning, yrkesliv och andra utmaningar i livet.” (WISC-III:86).

Testen er et psykologisk måleinstrument som bedømmer den intellektuelle kapasiteten hos barn i alderen 6 til 16 år. Testmanualen presiserer at intelligens er mye mer enn det som kan testes. WISC-testen kan sees på som en stikkprøve på intelligens. Wechsler selv sa at intellektuell begavelse er bare en del av intelligensen. Videre står det i testmanual at deltestene er valgt ut for at man ved hjelp av dem skal kunne beskrive mange ulike mentale egenskaper. Til sammen gir deltestene et bilde av barnets generelle intellektuelle evne. Med andre ord er det viktig å skille testresultatet og barnets kognitive potensialet, fordi det ikke er det samme.

Normene i den Svenske versjonen av WISC-III er representative for den svenske barnepopulasjonen. Stratifisert utvalg, sosialgruppedeling, utdanningsnivå, alder og kjønn er fordelt. Språklig tilhørighet er beskrevet slik: ”Grundkravet för alla barn som deltagit i normeringstestningarna har varit att de skal klara vanlig svensk förskola/skola.” (WISC-III:30). I normeringsgruppen var det 87 % av barna som snakket svensk hjemme. Testing med WISC er ikke tilrådelig i følgende situasjoner: ”Om barnet kommer från en annan kultur och/eller har bristfälliga kunskaper i svenska, kan testet knappast fungera så som det är avsett.” (WISC-III:89). Testens verbale del er mest følsom viss barnet kommer fra en annen kultur. Utføringsdelen kan brukes som et grovt og ufullstendig mål, men bør suppleres med ikke-verbale tester som Ravens matriser, Bender Visual Motor Gestalt test, Leiter-R eller Draw-a-man. Barnet bør utredes i forhold til sin egen kulturelle bakgrunn og eventuelt anvende tolk, en slik utredning bør skreddersys for hvert enkelt barn. Jeg forstår WISC-III slik at testen ikke gir valide målinger for elever fra språklige minoriteter med et annet morsmål enn svensk. Derimot er det svenske normeringsgrunnlaget brukt for norske barn. *Er barna i det svenske og norske samfunn like?*

Deltestene i WISC-testen, valgte Wechsler ut på grunnlag av sine erfaringer om at de var klinisk sensitive. WISC-III består av 13 deltester som er delt i en verbaldel og en utføringsdel.

Manualen har en generell oversikt over hva hver enkelt deltest måler, men det står tydelig i innledningen at dette er et utgangspunkt for refleksjon rundt barns testresultat og profil. Det vil si at deltestene ikke skal tolkes bokstavelig.

Flere steder i WISC-III manualen vises det til at testen er effektiv og tidsbesparende metode for å få objektiv informasjon om barnet. Testen brukes mye fordi den gir mye informasjon på kort tid, det tar ca 2 timer å gjennomføre hele testen. En annen grunn for at testen brukes mye, mener jeg er fordi det er den testen vi har hatt tilgjengelig lengst.

Testmanualen viser til studier og litteratur som gir en klar støtte til testens validitet og reliabilitet.

## **1.6.2 LEITER-R**

Russell Leiter startet utviklingen av sin test i 1927. Den første kommersielle Leiter-testen kom først ut i 1940. Leiter ønsket å lage en test som kunne måle intelligens på tvers av forskjellige kulturer. Han så at det var kulturelle problemer med de tradisjonelle intelligenstestene.

“Desiring to create a measure of intelligence with fair, cross cultural applicability.... The goal of developing this instrument was to construct a reliable and valid nonverbal measure of intellectual ability, memory and attention that could be used to assess children, adolescents that young adults who could not be reliably and validly assessed with traditional intelligence tests.”(1997:1)

Leiter-R er en nonverbal evne test, beregnet for personer i alderen 2 år og 0 måneder til 20 år og 11 måneder. Leiter-R er utviklet for blant annet å kunne teste minoriteter og elever med engelsk som andrespråk. Jeg forstår det slik at Leiter-R er en test som kan brukes på alle elever uavhengig av språklig tilhørighet eller kulturell bakgrunn. Jeg mener at den uavhengige språktilhørigheten er en selvmotsigelse i Leiter-R. Fordi alle deltestene har et spillnavn, dette er ment som en motivasjonsfaktor. Denne motivasjonsfaktoren kan ikke

gjelde for de som ikke kan språket. Dette ser jeg ikke på som en alvorlig bias, men man bør være klar over dette forholdet.

Leiter-R som brukes i dag, ble revidert og forbedret i 1997. Testens normeringsgrunnlag består av samme prosentvise folkegruppesammensetning som U.S. folketelling i 1993. Testen er ikke normert etter norske forhold. *Betyr det at amerikanske og norske barn regnes som like?*

Leiter-R består av 20 deltester som er organisert i fire områder: resonering, visualisering, hukommelse og oppmerksomhet. Leiter-R er delt i to subtestgrupper, der den ene måler den generelle intellektuelle evnen og den andre måler hukommelse og oppmerksomhet.

(VR) Visualisering og resonering, regnes som den intellektuelle delen av Leiter.

Brief screen IQ tar 25 minutter og full IQ skala tar 40 minutter.

(AM) Oppmerksomhet og hukommelse

Kort versjon tar 25 minutter og hele tar 40 minutter.

Det vil si at det tar ca 2 timer å gjennomføre hele testen. Testmanualen sier ikke noe om at dette er en komplisert test å administrere. Erfaring viser at testadministrator bør ha trent en del på testing med Leiter-R før han begynner å teste. Jeg tror den kompliserte administrasjonen av testen er en av grunnene til at testen ikke benyttes så mye som jeg mener den burde vært. Jeg mener Leiter-R burde vært brukt mer fordi: "When the discrepancy between general cognitive ability and academic achievement must be assessed, particularly for children with English as a Second Language or those with communication disorders, Leiter-R provides a valuable source of estimates of nonverbal intellectual ability".(1997:3)

Leiter-R regnes for å være en kulturreduert og solid nonverbal test. "Over the years, the Leiter International Performance Scale has been recommended in the professional literature for evaluation of a number of speech and language deficient groups because of its unique response format which is expressively nonverbal." (1997:ix).

---

Testmanualen viser til forsknings studier som bekrefter at Leiter-R er en valid og reliabel test. Korrelasjonsstudier med andre tester viser god korrelasjon. Studier har vist at det er en høyere korrelasjon mellom Leiter-R og utføringsdelen på WISC-III, enn det er mellom Leiter-R og den verbale delen på WISC-III. Leiter-R fullskala IQ og WISC-III viste en korrelasjon på .86. Leiter-R fullskala og Utføringsdelen på WISC-III hadde en korrelasjon på .85. Leiter-R fullskala og den verbale delen på WISC-II hadde en lavere korrelasjon på mellom .77 og .80. Disse korrelasjonsstudiene viser at Leiter-R og WISC-III måler relativt likt. Det jeg savner i korrelasjonsstudiene er en redegjørelse for bakgrunnen til personene som ble brukt i disse studiene? Var det minoritetsspråklige eller en gjennomsnitt person som ble brukt ved korrelasjonen?

### **1.6.3 LEM**

I Lidz (2000) sin oversikt over hva som finnes av dynamiske tester var LEM den eneste som var spesifikt beregnet for minoritetsspråklige. Lidz skriver følgende om LEM, "Main purpose is to provide a fairer measure of intelligence in children from minority cultures." (2000:8). Dette var bakgrunnen for at jeg valgte LEM som eksempel på en dynamisk test.

LEM ble utgitt i Nederland i 1991 av Hamers, Hessels og Van Luit. De begynte arbeidet med testen i 1987. Hensikten med LEM var å utvikle en test som kunne beregne generelle kognitive evner hos tyrkiske og marokkanske barn, i alderen 5 til 8 år, som var bosatt i Nederland. LEM var ment å skulle være et bedre verktøy for å beregne generelle kognitive evner hos minoritets barn, enn det tradisjonelle intelligenstester gjorde. LEM er utviklet fra et læringspotensielt perspektiv. (Kapittel 4). Det betyr at LEM måler elevens læringspotensial, LP. (Kapittel 4.2.) Testen måler hvilke nytte eleven kan dra av den hjelp den får under testingen. I motsetning til de etablerte intelligenstestene som måler hva eleven kan greie uten hjelp.

”Het testen beperkt zich namelijk niet tot de bepaling van datgene wat het kind reeds aan kennis en vaardigheden heeft verworen, doch beoogt een schatting te geven van het leervermogen, gebaseerd op de leerprestaties van het kind op algemene intelligentietaken. Dit weerspiegelt de definitie dat `intelligentie het vermogen tot leren is` (zie ook Budoff 1987b)”. (LEM 1991:71).

Testingen begrenses nemlig ikke til en fastsetting av det barnet allerede har tilegnet seg av kunnskap og ferdigheter, men tar sikte på å gi et anslag av læreevnen, basert på barnets læreprerastasjoner på generelle intelligensoppgaver. Dette gjenspeiler definisjonen at ”intelligens er evnen til å lære.” (Min oversettelse fra LEM 1991:71).

Forskning har vist at LEM også kan brukes på nederlandske barn i samme aldersgruppe. Det har vært utført studier av LEM i andre land enn Nederland, disse studiene har vist gode resultat.(Hessels og Hessels-Schlatter, 2002) Det vil si at LEM er en kognitiv test som kan brukes på barn i alderen 5 -8 år.

Hvordan ville testresultatene fra LEM vært på barn bosatt i Norge? Jeg mener man ikke kan slutte seg til at LEM er en universell test, uten først å ha undersøkt hvordan testen fungerer i de respektive land den skal brukes i. Det burde vært gjennomført studier av LEM i Norge. Hvis det viser seg at LEM gir valide testresultat på barn (språklige minoriteter og majoriteter) i Norge, mener jeg vi har et godt verktøy som er kulturet redusert og lett å administrere. Vi har da tilgang på en kognitiv test som gir valid og reliabel informasjon om barns læringspotensial, uavhengig av barnets språklige tilhørighet.

Testing av læringspotensial er det som har vist seg å være mest lovende for minoritetsspråklige elever. De tradisjonelle intelligensstestene undervurderer ofte språklige minoritetselevens kognitive evner... “...findings provide some support for the hypothesis that a `dynamic` process like learning cannot be adequately predicted from a `stativ` measure and general cognitive abilities but, instead, requires a `dynamic` measure of cognitive functioning.” (Hessels 1997:122).

LEM tilhører den dynamiske tilnærmingen som kalles for, 'læring med testen'. (Kapitel 4). Læringen i LEM er standardisert og den består av å gjenta oppgaver. Testadministrator skal ikke begynne selve testingen før eleven har forstått deltesten. Skåringen beregnes i forhold til hvilke nytte eleven har av hjelpen i læringsdelen. Det vil si at jo mer hjelp eleven trenger, desto lavere skåre får eleven. Elevens råskåre regnes om i en skala som er kompatibelt med IQ poeng. I LEM kalles skåren for læringspotensial skåren (LP-skåren) og ikke IQ skåre. LP-skåren har en normalfordeling som er lik IQ skåren i tradisjonelle evnetester. I Testmanualen til LEM er det en oversikt over normalfordelingen i "Tabel 5.4 Indeling in leerpotentieel-categorieën" (LEM 1991:71).

Tabell 2. Inndeling av læringspotensial – kategorier. (Min oversettelse).

<b>LEM LP-resultat</b>	<b>%under normalfordelingen</b>	<b>Kvantitative beskrivelse</b>	<b>Praktisk beskrivelse</b>
> 130	2.2	Høyt læringspotensial	Begavet
115 – 130	13.6	Over middels læringspotensial	Oppvakt
85 – 115	68.3	Middels læringspotensial	Middels
70 – 85	13.6	Under middels læringspotensial	Lærevansker
< 70	2.2	Lavt læringspotensial	Mentalt retardert

Det forventes at barn med et høyt eller over middels lærepotensial selvstendig vil kunne dra fordel av undervisningen som blir tilbudt. Barn med et lavt eller under middels lærepotensial danner en risikogruppe. De vil muligens trenge individuell veiledning for å kunne dra fordel

av undervisningen. Barn med et middels lærepotensial, vil i prinsippet kunne følge basisundervisningen, så lenge det på en tilstrekkelig måte tas hensyn til barnets spesifikke egenskaper. (Min oversettelse av LEM 1991:71).

Elever som får en høy LP-skåre viser større progresjon i skolelæringen enn de elevene som får en lav LP-skåre. En elev med høy LP-skåre kan i mange tilfeller skåre lavt på tradisjonelle IQ tester. Det vil si at LP-skåre er et bedre estimat på skolesuksess enn IQ skåren.

Testen består av til sammen 6 deltester 3 av testene er nonverbale og 3 av testene er verbale. De verbale deltestene bruker non-ord, det vil si oppdiktete ord uten mening. Deltestene som LEM består av er nøye utvalgt. 3 av deltestene skal måle induktiv resonering, fordi dette alltid har spilt en viktig rolle i intelligens tester. De andre deltestene måler verbale evner og arbeidsminnets kapasitet. De utvalgte deltestene er aldersrelaterte, sensitive til individuelle forskjeller og korrelerer med elevens skoleprestasjoner.

Testen er lett å administrere og det er beregnet at den ikke skal bruke mer tid enn eksisterende evnetester. Det er utført mange studier som viser at LEM er en valid og reliabel test for majoritets- og minoritets elever i Nederland.

Tabell 3: Sammenligning av WISC-III, Leiter-R og LEM.

<b>WISC-III</b>	<b>Leiter-R</b>	<b>LEM</b>
Verbal evnetest	Non verbal evnetest	Non verbal og verbal evnetest
Statisk test	Statisk test	Dynamisk test
Alder: 6 - 16 år	Alder: 2 år og 0 mnd - 20 år og 11mnd.	Alder: 5 - 8 år.

Testen er ikke egnet for de som ikke behersker svensk. Testen er beregnet for de som kan landets majoritets språk.	Testen er egnet for minoriteter.  Er spesielt beregnet for de som ikke har engelsk som førstespråk	Testen er laget for tyrkiske, marokkanske og nederlandske barn bosatt i Nederland.  Egnet for etniske minoriteter i andre land.
Normert for svenske barn	Normert for amerikanske barn.	Normert for marokanske og nedelandske barn.
13 deltester	20 deltester	6 deltester
Stopper når 3 feil på rad	Stopper når 3 eller 5 kumulative feil	Stopper når alle oppgavene er utført.
Tidsbegrensning. Overskrides tiden er det 0 poeng.	De fleste oppgavene er uten tidsbegrensning.	Tidsbegrensninger er ikke brukt.
Fare for læringseffekt ved re-testing.	Re-testing er greit	Re-testing er greit.
Skal kun brukes av de med kunnskap om testen og psykologiske forhold.	Omtaler brukerens kvalifikasjoner og etiske betraktninger	Testadministrator må kunne teste med andre tester.
Middels administrering	Innviklet administrering	Lett administrering
Oppfyller psykometriske krav.	Oppfyller psykometriske krav.	Oppfyller psykometriske krav.

## 1.7 Sakkyndig utredning

### 1.7.1 Det er viktig å skille mellom testing og utredning.

Utredning er mer enn testing, testing er bare en del av utredningen. En utredning kan bestå av testing, intervju med foreldre eller lærere, observasjon og andre undersøkelser som, pedagogisk psykologisk tjeneste (PPT) eller som andre har utført. Med andre ord kan man si at en utredning er innsamling av data. Det er en kritisk fase i starten av utredningen. (Lidz 2003). Med kritisk fase menes; det er nå føringen for den videre utredning, data innsamlingen, og valg av metoder for utredningen legges. Hvilke metoder og strategi som velges for utredningen er med å påvirke hva man greier å avdekke av barnets sider og hvilke valg man tar med tanke på tiltak for barnet. Det er viktig at de som utreder er fleksible slik at de ikke velger samme utredningsstrategi uansett hvem de utreder. Her er det et paradoks da det kan være lurt å holde seg til et visst sett med tester man føler seg trygge på og kan. Det er en fare ved å være for fleksible og velge tester man kjenner dårlig og dermed ikke kan foreta en god tolkning av testresultatene. (Lidz 2003).

### 1.7.2 Forholdet mellom tilpasset opplæring og spesialundervisning

Tilpasset opplæring er et overordnet prinsipp i læreplanverket (L97) for grunnskolen. I Opplæringsloven heter det i §1-2 ”...opplæringa skal tilpassast evnene og føresetnadene hjå den enkelte eleven, ...”. Denne tilpasningen skal skje innenfor skolens rammer. Når skolen

---

ikke greier å tilrettelegge for eleven med tilpasset opplæring innenfor skolens rammer sendes det en tilmelding til PPT. PPT foretar en sakkyndig vurdering for å se hvilke behov eleven har og om eleven har behov for spesialundervisning. Utdannings- og forskningsdepartementet har utarbeidet ”Veiledning for spesialundervisning i grunnskolen og videregående opplæring” (2004) der står det;

”Spesialundervisning er en form for tilpasset opplæring, men ikke all tilpasset opplæring er spesialundervisning. Spesialundervisning innebærer at det på bakgrunn av enkeltvedtak settes i gang særskilte tiltak for en som ikke får tilfredsstillende utbytte av opplæringen, tiltak som bare noen har krav på.” (Utdannings- og forskningsdepartementet 2004:19).

Enkeltvedtaket det refereres til i veiledningen er blant annet basert på en sakkyndig vurdering. Opplæringsloven sier at kommunene skal ha en PPT som utarbeider sakkyndig vurdering som kan konkludere i et enkeltvedtak om behov for spesialundervisning. I opplæringsloven heter som følger;

§5-6 Pedagogisk-psykologisk teneste.

Kvar kommune og kvar fylkeskommune skal ha ei pedagogisk-psykologisk teneste... Tenesta skal hjelpe skolen i arbeidet med kompetanseutvikling og organisasjonsutvikling for å leggje opplæringa betre til rette for elevar med særlege behov. Den pedagogisk-psykologiske tenesta skal sørge for at det blir utarbeidd sakkunnig vurdering der lova krev det...

Her sier opplæringsloven at det er PPT som har ansvaret for å utarbeide sakkyndig vurdering. En sakkyndig vurdering kan konkludere med et vedtak om spesialundervisning. Det vil si at det er PPT sitt mandat å vurdere behovet for spesialundervisning.

### 1.7.3 Når blir språket et problem for språklige minoriteter og hva skjer når problemet oppstår?

Bruken av spesialundervisning for språklige minoriteter har vært kritisert. (Pihl 2000, Hessels 2000, Lidz og Penà 1996, Samuda og Kong 1989). På bakgrunn av denne kritikken ønsker jeg å få frem hva lovverket sier om spesialundervisning og hvem som har ansvaret for henvisningen.

Språklige minoriteter som ikke har tilstrekkelig ferdigheter i norsk til å følge den vanlige undervisningen skal få tilbud om særskilt norskopplæring. Opplæringsloven sier som følger:

”§2-8 Særskild språkopplæring for elevar frå språklege minoritetar.

Elevar i grunnskolen med anna morsmål enn norsk og samisk har rett til særskild norskopplæring til dei har tilstrekkeleg dugleik i norsk til å følge den vanlege opplæringa i skolen. Om nødvendig har slike elevar også rett til morsmålsopplæring, tospråkleg fagopplæring eller begge delar.

Morsmåloplæringa kan leggjast til annan skole enn den eleven til vanleg går ved.

Når morsmålsopplæring og tospråkleg fagopplæring ikkje kan givast av eigna undervisningspersonale, skal kommunen så langt det er mogleg leggje til rette for anna opplæring tilpassa føresetnadene til elevane.”

Som det går frem av §2-8 har elever fra språklige minoriteter krav på særskilt norskopplæring til de har tilstrekkelige norskferdigheter til at de kan følge vanlig undervisning. Det vil si at språket er et problem når elever fra språklige minoriteter ikke kan følge vanlig norsk undervisning. Hvem har ansvaret for å vurdere om elevene har tilstrekkelige norskferdigheter til å ha tilstrekkelig utbytte av undervisningen? I et rundskriv fra utdannings- og forskningsdepartementet står det bland annet:

---

”Morsmålsopplæring må komme i tillegg til det vanlige timetallet i grunnskolen..... det er kommunenes ansvar å kartlegge elevenes behov med sikte på å gi den enkelte et tilbud som er mest mulig i samsvar med elevenes forutsetninger. Departementet har utarbeidet kartleggingsmateriell som støtte i skolens arbeid med å avdekke elever for særskilt norskopplæring, morsmålsopplæring og tospråklig fagopplæring. Dette kan fåes ved å kontakte læringscenteret.” (Rundskriv F-008-03:1).

I dette rundskrivet kommer det frem at det er kommunene som har ansvar for å kartlegge elevenes behov for særskilt norskopplæring og skolen som skal avdekke dette behovet.

Hva skjer når elever fra språklige minoriteter ikke har tilstrekkelig utbytte av særskilt norskopplæring? I de tilfellene der skolen vurderer elevens utbytte av undervisningen ikke er tilstrekkelig, sender skolen en tilmelding til PPT.

Skolen sender tilmeldingen når de har mistanke om at eleven har lærevansker i tillegg til at de er svake i norsk. Når PPT mottar en tilmelding fra skolen foretar de en sakkyndigvurdering. Den sakkyndige vurderingen skal vise elevenes særlige opplæringsbehov og kan resultere i en anbefaling om spesialundervisning.

#### Opplæringsloven §5-3 Sakkunnig vurdering

Før kommunene eller fylkeskommunen gjer vedtak om spesialundervisning etter §5-1 eller 5-2, eller vedtak om spesialpedagogisk hjelp etter §5-7, skal det ligge føre ei sakkunnig vurdering av dei særlege behova til eleven. Vurderinga skal vise om eleven har behov for spesialundervisning, og kva for opplæringstilbod som bør givast.

Det vil si at en sakkyndig vurdering kan konkludere med at eleven har behov for spesialundervisning. Hvilke elever har krav på spesialundervisning?

”§ 5-1 Rett til spesialundervisning.

Elevar som ikkje har eller kan få tilfredsstillande utbytte av det ordinære opplæringstilbodet, har rett til spesialundervisning...”.

Dette tolker jeg dit hen at elever fra språklige minoriteter som ikke har tilstrekkelige ferdigheter i norsk heller ikke har et tilfredsstillende utbytte av opplæringstilbudet, som en følge av dette har krav på spesialundervisning.

Det har vært mye kritikk av språklige minoriteter som har vært henvist til spesialundervisning da spesialundervisning har vært forbundet med lærevansker. Departementet kom med et rundskriv for å rette opp i denne uklarheten om bruken av spesialundervisning og språklige minoriteter.

”Elever med rett til særskilt språkopplæring etter opplæringsloven §2-8, kan også få spesialundervisning etter opplæringsloven § 5-1. Dette blir en konkret vurdering som foretas i det enkelte tilfelle.” (Rundskriv F-008-04)

Det bør sies at dette rundskrivet kom etter at bland annet Pihl (2002) fremsatte sin kritikk om bruken av spesialundervisning for språklige minoriteter.

Hvis PPT ikke finner grunnlag for spesialundervisning i henhold til §5-1, bør PPT gi råd om hvordan skolen best mulig kan tilrettelegge for at eleven kan dra nytte av skolens undervisningstilbud. Se opplæringslovens §2-8, ”Særskilt språkopplæring”. Her mener jeg PPT burde kunne gi råd og veiledning til skolen.

#### **1.7.4 Sakkyndig vurdering**

En sakkyndig vurdering er en grundig faglig vurdering av elevens behov. I et sakkyndig utredningsarbeid er det to sentrale momenter utreder må ta stilling til (Undervisnings- og forsknings departementet 2004). De sentrale momentene er; 1. elevens utbytte av det

---

ordinære opplæringsstilbudet, 2. elevens lærevansker og andre særlige forhold som er viktig for opplæringen. (2004:38).

Den sakkyndig vurdering har en utredningsdel og en tilrådningsdel. (Utdannings- og forskningsdepartementet 2004). Det er i utredningsdelen testene er å finne. Konklusjonene som kommer frem i tilrådningsdelen trekkes blant annet på bakgrunn av tester og annen dokumentasjon som er fremskaffet. PPT er ansvarlig for at disse sakkyndige vurderingene blir utarbeidet. (Opplæringsloven § 5-6). Utredningsansvaret til PPT er ikke ensbetydende med at det er PPT som skal utføre alle undersøkelsene. I Håndbok for PP-tjenesten(2001) står det hvilke ansvar PPT har.

*”Loven må forstås dit hen at PP-tjenesten er ansvarlig for:*

*\*at det blir gjennomført en sakkyndig vurdering.*

*\*de konklusjoner som trekkes, det vil si tilråkningene.”* (Håndbok for PP-tjenesten 2001:35).

Skolen har utført vurderinger av eleven før eleven tilmeldes til PPT, disse opplysningene følger med i tilmeldingen. I utredningsarbeidet kan PPT ta i bruk flere kilder for å få et helhetlig bilde av elevens opplæringsbehov: samtale med elev/foreldrene, utdypende opplysninger og vurderinger fra skolen, egne undersøkelser (testing, observasjon, intervju etc.),rapporter og utredninger fra andre sakkyndige, samtaler med andre berørte. Dette kan være med å gi PPT informasjon og dokumentasjon til å foreta en sakkyndig vurdering av elevens behov.

I ”Veiledning for spesialundervisning” (2004) kapittel 4 Sakkyndig vurdering, heter det:

*”Det finnes en rekke kartleggingsprøver og diagnostiske prøver som er standardisert og gir indikasjon på elevens nivå i forhold til et landsgjennomsnitt. Ulike prøver og tester kan være ett av flere hjelpemidler i vurderingen av om eleven har behov for spesialundervisning.”* (2004:37).

I sitatet over vises det til bruken av tester som et hjelpemiddel i vurdering av elevens behov i en sakkyndig utredning. Slik jeg tolker dette sitatet bygger mye av elevens opplæringsbehov på tester. Det vil si at testene har en sentral plass i en sakkyndig utredning. Standardiserte

tester har lenge vært et godt redskap for å se hvilke sjanser barnet har for å lykkes eller mislykkes akademisk. (Lidz 2003) Men en test som kan forutse barnets fremtid overser samtidig barnets mulighet til fleksibilitet. Lidz sier dette treffende; ”We can no longer be satisfied with such successes. We can no longer be content to predict that a child will fail”. (Lidz 2003:112). Vi kan ikke lenger kan være fornøyd med å predikere at en elev skal mislykkes. Fleksibiliteten må ikke glemmes i lys av prediksjonen. Mye kan skje med en elev i løpet av skolegangen.

### **1.7.5 Elevens fremtidsutsikter**

Resultatet av en utredning munner ut i en antagelse om hva som er barnets problemområde, i dette ligger det også en fremtidsutsikt. (Lidz 2003). Hvilke fremtidsutsikter er det for dette barnet? Sakkyndig vurdering kan dermed være med å forme elevens fremtidsutsikt, derfor er det viktig at vurderingen ikke inneholder feilkonklusjoner. Prediktiv validitet er derfor viktig i en sakkyndig vurdering. Testes en språklig minoritetselev med en test som ikke er egnet for elevens erfaringsbakgrunn med påfølgende tolkning ut i fra testresultat, kan dette få katastrofale følger for elevens fremtidsutsikter og utviklingsmuligheter. Her mener jeg vi klart kan se at dynamiske tester kan være med som alternativer eller supplement til de standardiserte testene. Dynamisk testing er relativt ny, men har fått mye oppmerksomhet, spesielt når det gjelder det kulturelle mangfoldet som er i dagens skole. (Lidz 2003). Dynamiske tester måler elevens læringspotensial og ikke det eleven har lært til nå, slik statiske tester gjør. Videre redegjørelse av dynamiske tester finnes i kapittel 4.

---

### 1.7.6 Kompetansekrav til de ansatte i PPT.

Som vist over har PPT et stort ansvar for å vite hva de gjør, fordi deres konklusjoner kan være med å forme fremtidsutsikter og utviklingsmuligheter til elever. I tidligere forskrift for PPT i grunnskolen var det stilt formelle krav til hvem som kunne ansettes i PPT. (Håndbok for PP-tjenesten 2001). Dagens lovgivning har ikke et slikt formelt krav til kompetanse. Her mener jeg det er en svakhet i lovverket. Bakgrunnen for fjerning av det formelle kravet til hvem som kan ansettes i PPT, er å øke kommunenes frihet og unngå detaljregulering. (Håndbok for PP-tjenesten 2001). Jeg mener dette er en missforstått bruk av handlingsfriheten og en form for ansvarsfraskrivelse fra de sentrale myndigheter. At sentrale myndigheter stiller formelle krav til de sakkyndige i PPT betrakter jeg ikke som detaljregulering, men mer som en kvalitetssikring. Det er viktig å sikre at de sakkyndige har tilstrekkelig kompetanse. Hvordan sikrer man i dag at PPT har den formelle kompetanse? Jeg ser det som viktig at alle som ansettes i PPT som sakkyndig rådgivere har formell kompetanse, slik at man forebygger feilvurderinger av elever. Se tidligere avsnitt 1.3.4 testkompetanse.

Håndboken for PP-tjenesten (2001) viser til tre nivå som kan beskrive kompetansen i PPT.

Det er handlings-, planleggings- og refleksjonsnivå. Dette virker veldig fint, men er det virkelig bevissthet rundt dette ute i felten? Det er et relativt stort press på å få saker gjennom i systemet, så blir det da tid for refleksjoner? Refleksjoner kan ikke regnes ut og vises frem i tabeller som effektivitet eller produkt.

Som profesjonell yrkesutøver står en i et maktforhold i forhold til brukeren. Bevissthet om det maktforholdet mener jeg er viktig. Ethiske betraktninger er av stor betydning både for utøver og bruker. (kapitel 3.1. Hvordan testene bør brukes).

## 1.8 Kilderefleksjoner

I en slik oppgave hvor mesteparten av kildene består av litteraturundersøkelser, bruker man flere skriftelige kilder. Noen kilder er primære andre sekundære, førstehånds, andrehånds, normative eller beskrivende kilder. Skriftlige kilder kaller Kjeldstadli (2000) for meddelende kilder, det er kilder som påstår noe om verden. Det er slike meddelende kilder jeg har brukt i denne oppgaven. Meddelende kilder kan være normativ, beskrivende eller en blanding av begge. Normative kilder sier noe om hvordan noe bør være og beskrivende kilder sier noe om hva som har vært eller kommer til å skje.

Et rundskriv og en lov er normative kilder, men en undersøkelsesrapport som for eksempel Samped er beskrivende kilder, de har funn som er beskrivende. Artikkene jeg har brukt i denne oppgaven er beskrivende kilder. Boken til Sternberg og Grigorenko (2000) er også et eksempel på en beskrivende kilde. Denne beskrivende kilden er også en primær kilde når jeg viser til det Sternberg og Grigerenko har skrevet i sin egen bok.. Refererer jeg derimot til andre forskere som Sternberg og Grigprenko har skrevet om, blir det en sekundær kilde. Det kan være fare for tolkningsfeil ved å bruke sekundære kilder, derfor har jeg forsøkt så langt det har latt seg gjøre å lete opp de primære kildene.

## 1.9 Oppgavens oppbygning

For å svare på oppgavens tre problemstillinger fant jeg det hensiktsmessig og ryddigst å dele dem i hvert sitt kapittel. Jeg har derfor laget en oppdeling på tre hovedkapitler som tar for seg hver sin problemstilling, kapitlene 2, 3 og 4. I disse kapitlene blir problemstillingene belyst med blant annet relevant forskning og drøftning av andres og mine synspunkter.

Kapittel 2 tar for seg problemstilling 1. ”Hvilke problemer innebærer dagens praksis med bruk av statistisk testing av språklige minoriteter? Her viser jeg hva fagfolk selv og andre sier om dagens praksis med bruk av statiske tester på elever fra språklige minoriteter. Jeg viser også dagens praksis ved hjelp av egne analyser av 6 sakkyndige vurderinger.

Kapittel 3 handler om problemstilling 2. ”Hvordan kan man eliminere forekomsten av feilkonklusjoner i testing av elever fra språklige minoriteter?” Her fokuserer jeg på hvordan testresultatene blir tolket og alternative tolkningsmetoder. Jeg ser også på hva testmanualene til WISC-III, Leiter-R og LEM sier om tolkning. Til slutt tar jeg for meg hva jeg mener skal til for å bli god til å tolke testresultat?

Kapittel 4 er om problemstilling 3 ”Kan dynamiske tester redusere forekomsten av feilkonklusjoner og øke reliabel informasjon i utredningen av språklige minoriteter?”. Dette kapitlet er en gjennomgang av hva en dynamisk test er og forskjellige typer av dynamiske tester. Kapitlets siste del har med empiri fra artikler om reduksjon av feilkonklusjoner ved testing av elever fra språklige minoriteter.

I Kapittel 5 avslutter jeg oppgaven med en konklusjon på mine tre problemstillinger.

## 1.10 Avslutning

Med elever fra språklige minoriteter mener jeg de som har et annet morsmål enn norsk.

Individuelle variasjoner, tospråklighetstyper og språkforståelsen er tre fenomen som viser mangfoldet i gruppen språklige minoriteter.

Jeg har undersøkt utbredelsen av språklige minoriteter i den norske grunnskolen. Av de elevene som henvises til PPT i Oslo er 35 % av elevene minoritetsspråklige. I Oslo får 10.5

% av de minoritetsspråklige elevene spesialundervisning. Dette mener jeg er en betydelig andel av den totale elevmassen.

Tester er en av metodene som brukes i en sakkyndig vurdering. Når en elev ikke har et tilfredsstillende utbytte av skolen sendes det en tilmelding til PPT. Det er PPT som er ansvarlig for at det utarbeides en sakkyndig vurdering og vurderer om det er behov for spesialundervisning. Jeg mener det er veldig viktig å sette seg inn i hva testmanualen sier om testen. Tester har vært kritisert fordi testingen ikke har fulgt testspesifikasjonene. Jeg valgte å presentere testene WISC-III, Leiter-R og LEM for å vise hva testmanualene sier om testene. For å sikre seg at tester brukes riktig er det viktig å kjenne testene og vite hvorfor man skal teste med den testen man har valgt.

Det er motbevist at språklige minoriteter er dumme eller at tospråklighet hemmer den kognitive utviklingen. Er det noe med dagens praksis som gjør at elever fra språklige minoriteter fremdeles er overrepresentert i spesialundervisningen? For å få svar på dette spørsmålet undersøker jeg hva de sakkyndige og fagfolk sier om dagens praksis ved testing av elever fra språklige minoriteter. Dette tar jeg opp i neste kapittel.



## **2. KAPITTEL 2 PROBLEMSTILLING NR. 1.: HVILKE PROBLEMER INNEBÆRER DAGENS PRAKSIS MED BRUK AV STATISKE TESTER PÅ ELEVER FRA SPRÅKLIGE MINORTITETER?**

### **INNLEDNING**

For å belyse problemstilling nummer 1 vil jeg vise dagens praksis fra forskjellige vinkler. Jeg starter med min analyse av 6 sakkyndige vurderinger for elever fra språklige minoriteter. Her har jeg sett på om elevens tospråklighet er vurdert, hvilke tester er benyttet og hvordan testene er tolket.

Den andre vinklingen jeg ser på er hva tre norske fagfolk sier om sakkyndige vurderinger og dagens praksis i PPT. Som representanter for norske fagfolk har jeg valgt Hansen, Birkemo og Phil.

I den tredje vinklingen viser jeg til en survey fra Nord Colorado der de sakkyndige har blitt spurt om egen praksis. De sakkyndige har svart på hvilke tester de bruker ved utredning av elever fra språklige minoriteter.

Den fjerde og siste vinklingen jeg har med, viser hva som er foreldrenes reaksjoner når elevene ikke blir utredet, men plassert i NOA (norsk som andre språk) på grunn av sin språklige bakgrunn.

Hvilke problemer er det med statistisk testing av språklige minoriteter? Kunnskap om tospråklig utvikling er av betydning for hvordan tester bør brukes og tolkes. Uten kunnskap om tospråklig utvikling er det fare for feilslutninger som kan føre til sosial reproduksjon og stemping av elever som mentalt retarderte. Løsningen er ikke å fjerne testene og de sakkyndige vurderingene, det vil bare gjøre vondt verre. Dagens praksis ved bruk av statistiske tester på elever fra språklige minoriteter er ikke god nok, noe bør gjøres.

## 2.1 Analyse av sakkyndige vurderinger

I problemstilling 1. spør jeg "Hvilke problemer er det med dagens praksis med bruk av statiske tester av språklige minoriteter?" Her hevder jeg at det er problemer med dagens praksis" Jeg har derfor med 6 sakkyndige vurderinger av elever fra språklige minoriteter, for å vise noe av dagens praksis. Er det virkelig noe problem med dagens praksis som jeg hevder i problemstilling nummer 1? Mine antagelser om dagens praksis før analysen er, WISC-III brukes ved utredninger av elever fra språklige minoriteter, Leier-R brukes lite og dynamiske tester brukes ikke.

Analysen har med 6 sakkyndig vurderinger fra et tilfeldig valgt PPT-kontoret i Oslo. Jeg ba det utvalgte PPT-kontoret om å plukke ut de siste 6 sakkyndige vurderingene av minoritetsspråklig elever fra småskoletrinnet som var utredet av forskjellige PP- rådgiver. Vurderingene er anonymisert, navn, skole, og morsmål er fjernet.

Det er en overvekt av gutter i de sakkyndig vurderingene, 4 gutter og 2 jenter, og den ene jenta har et syndrom. Syndromet fører til at eleven har generelle lærevansker, men tospråkligheten er like mye tilstedet hos denne eleven som hos andre elever fra språklige minoriteter.

I analysen av de 6 sakkyndige vurderingene har jeg ønsket å finne svar på følgende 5 spørsmål:

1. Har de sakkyndige vurderingene vektlagt elevens minoritetsspråklige bakgrunn?
2. Hvilke type data er innhentet og er datainnhenting preget av en helhetlig - økologisk tenkning?
3. Hvilke tester er brukt i vurderingen?
4. Hvordan er testene tolket?
5. Har det vært bruk av dynamiske tester?

Disse 5 spørsmålene mener jeg kan være med å gi svar på om det er problemer ved dagens praksis. Spørsmålene gir også svar på om det er hold i mine antagelser om dagens praksis.

Tabell 4. Analyse av 6 sakkyndige vurderinger.

Kilder metoder	Elev 1	Elev 2	Elev 3	Elev 4	Elev 5	Elev 6
		Svak	Henvist av foresatte.		65 Svært sen faglig utvikling i norsk og matte.	
<b>Henvisnings Grunn</b>	Utrede læreforutsetninger,  Uro og samspilsvansker	konsentrasjon og leseferdigheter. Vansker med felles beskjeder.	Forsinket språkutvikling og konsentrasjonsvansker  Tidligere utredet av barnelege.	Lese og skrive vansker.	Utredning av elevens læreforutsetninger.  Bruker briller men klør i øynene.	Syndrom , ny utredning.
<b>Anamnese</b>	Ja	Nei	Ja	Nei	Ja	Ja
<b>Intervju av lærer</b>	Kontaktlærer	Ja	Ja	Ja.	Ja	-----
<b>Intervju av foreldre</b>	Ja	Ja	Ja	Ja	Ja	Ja
<b>Intervju av andre</b>	Nei	Eleven	Nei	Eleven	Nei	Eleven
<b>Morsmålet til eleven</b>	Tospråklig behersker morsmålet best	Tospråklig behersker morsmålet best.	Tospråklig svak i begge språk.	Tospråklig	Tospråklig  Snakker morsmålet	Tospråklig  Kan et par ord.
<b>Observasjon av eleven</b>	Klasserom	Klasserom	Klasserom  Skolegården.	Nei	Lesegruppe	Klasserom
<b>Tester som er brukt</b>	WISC-III  Berry VMI  Raven	WISC-III  Kartlegger basisferdigheter I lesing, skrijving og matte.	Wipsi-r.  Leiter-R.  Vineland.  Medisink utredning	WISC-III  VMI  Oppmerksomhetsprøve.d2  Carlsten  STAS	WISC-III  Raven  VMI  Leseoppgaver.	VMI – eleven er ikke testbar. Øvelse gir lærings-effekt.
<b>Hvordan er WISC-III tolket?</b>	Analyse og tolkning av deltestene	Analyse og tolkning av deltestene	Analyse og tolkning av deltestene	Analys av hoveddelene.  ACID-profil.	Analyse av noen av del testene som var vanskeligst	-----  -----
<b>Tolkninger</b>	Brå!	Ok	Ok	Ok	Ok	-----
<b>Hensyn til tospråklig.</b>	Ja,	Ja.	Ja	Nei	Ja	Nei
	Ikke §5.1  Trenger §2.8	Ikke §5.1  BUP-traumer.	Tildeles §5.1  IOP	Tildeles § 5.1	Tildeles §5.1	Tildeles §5.1

Kommentarer til tabellen:

Elev 1. Ser at språkvanskene kan skyldes andre ting en tospråkligheten.

Elev 2. Årsaken til vansken kan skyldes ubearbeidete traumer i barndommen.

Elev 3. Mulig generelle lærevansker, bør utredes videre av habiliteringsteam.

Elev 4. Misforhold mellom elevens intellektuelle forutsetninger og lese- skriveferdigheter, tilsier at eleven har spesifikke lesevansker. Her er det ikke tatt i betraktning at eleven er tospråklig.

Elev 5. Brille og kløen i øynene til eleven burde vært utredet. Jeg ville tatt en videre utredning/ henvisning da morsmåslærer sier at eleven trenger mer forklaringer enn andre.

Elev 6. ”Oppgavene måtte gjentas så mange ganger at det ble til en trenings – og lærings situasjon.”

### **2.1.1 Har de sakkyndige vurderingene vektlagt elevens minoritetsspråklige bakgrunn?**

Alle de 6 sakkyndige vurderingene har nevnt at eleven er fra et annet land og at eleven har et annet morsmål en norsk. Elevenes tospråklighet og hvordan elevene behersker morsmålet og norsk er beskrevet i 4 av de 6 sakkyndige vurderingene. I de 2 andre sakkyndige utredningene er ikke elevens minoritets språklige bakgrunn tatt opp i tolkningen av de innhentede dataene, dette gjelder elev nummer 4 og 6. Når det gjelder elev nummer 4 er det heller ikke skrevet om eleven behersker morsmålet eller norsk best eller like godt. Elev nummer 4 ble henvist til PPT for lese- og skrivevansker og har etter utredning fått tildelt spesialundervisning etter § 5.1. i opplæringsloven. I elevens sakkyndige utredning heter det: ” Misforhold mellom elevens intellektuelle forutsetninger og lese- skriveferdigheter, tilsier at eleven har spesifikke lesevansker.” (Sakkyndig vurdering for elev nummer 4). Hos elev nummer 6 står det skrevet at eleven ikke behersker noen av språkene, men kan kommunisere

med tegn til tale. Det vil si at det kun er i en av de 6 sakkyndige vurderingene at elevens tospråklighet ikke er vektlagt, i de resterende 5 sakkyndige vurderingene er elevens tospråklighet vektlagt.

### **2.1.2 Hvilke type data er innhentet og er data innhenting preget av en helhetlig - økologisk tenkning?**

Alle de 6 sakkyndige vurderingene har en helhetlig data innhenting som jeg mener er preget av økologisk tenkning. (Kap. 3 Tolkningstyper, Økologisk tenkning.) Det er innhentet data fra lærer og foresatte i alle de 6 tilfellene. Ellers er data hentet inn fra samtaler med 3 av elevene og det har vært klasseromsobservasjon av 5 elever. Alle elevene har blitt testet med flere typer tester.

### **2.1.3 Hvilke tester er brukt i vurderingen?**

Det er brukt flere typer tester i de 6 sakkyndige vurderingene, men jeg kommenterer kun bruken av de testene som er omtalt i denne oppgaven. (Kap. 3) Jeg har valgt å ikke kommentere de andre testene som er brukt i de 6 sakkyndige vurderingene fordi det er hovedgruppene av tester, statiske og dynamiske tester, tolkningen og bruken av tester på elever fra språklige minoriteter som er oppgavens problemstillinger.

Alle de 6 sakkyndige vurderingene har brukt de statiske evnetestene WISC-R og WIPSY, med unntak av eleven med syndrom, elev nummer 6. Hos elev nummer 6 konkluderer den sakkyndige at eleven ikke er testbar fordi eleven trenger så mye øving for å forstå oppgaven at det er fare for en læringseffekt. I tilfellet med elev nummer 6 mener jeg at en dynamisk test ville endt med at eleven var testbar, fordi dynamiske tester bygger på at eleven skal lære seg og øve på testen. (Kap. 4. Dynamiske tester.)

Det er kun en av elevene som er testet med den non-verbale evnetesten Leiter-R.

Leiter-R er spesielt utviklet for denne elevgruppen. (Kap. 3 Hva sier Leiter-R manualen om testing). Jeg mener derfor at det er en svakhet i utredningene at Leiter-R ikke er brukt mer i utredningene av elever fra språklige minoriteter.

#### **2.1.4 Hvordan er testene tolket?**

Generelt er alle tolkningene gode og reflekterte. Når det gjelder tolkningene av WISC-III og WIPSY-R i de sakkyndige vurderingene har de en god analyse av deltestene. Unntaket er elev nummer 4 som tolker forskjellen mellom testens to hoveddeler, verbale- og nonverbale del, som en ACID-profil (Kaufman 1994), uten å kommentere at det er en minoritets språklig elev. Jeg kan ikke se at det er tatt i bruk ulike tolkningstyper (Kapittel 3 Tolkningstyper) som jeg mener kan være nyttig ved utredning av elever fra språklige minoriteter.

#### **2.1.5 Har det vært bruk av dynamiske tester?**

Det er ikke brukt dynamiske tester i de 6 sakkyndige vurderingene.

---

### 2.1.6 Min oppsummering av de 6 sakkyndige vurderingene.

Elevens minoritetsspråklige bakgrunn nevnes i alle de 6 sakkyndige vurderingene og dette tas med i de fleste tolkningene. De 6 sakkyndige vurderingene har variert og helhetlig datagrunnlag for sine vurderinger og tolkninger. Elever fra språklige minoriteter testes med statiske evnetester som i følge testens manual ikke er normert for denne gruppen. Ingen av de sakkyndige vurderingene nevner at testen ikke er brukt i henhold til manualens spesifikasjoner. Jeg ser i utgangspunktet ikke noe galt i å bruke WISC-III og WIPSY-R på elever fra språklige minoriteter, men mener at man bør vise at testen ikke er utført i henhold til testmanualens spesifikasjoner. Når det avvikes fra test manualens spesifikasjoner er tolkningen av testens resultater avgjørende for hvor valide konklusjonene er. Pihl (2005) kritiserer at WISC-R er brukt på elever fra språklige minoriteter som testen ikke er normert for. I henhold til Kaufmans bok *Intelligent testing med WISC-III* (1994) er det tolkningen av testresultatene som har avgjørende betydning og ikke testresultatene alene. Tolkingen av testresultatene tar ikke Pihl (2005) opp i sin kritikk.

Leiter-R som er spesielt utviklet for elever fra språklige minoriteter er kun brukt på en av elevene. Dette ser jeg på som et forbedrings potensialet i utredninger av elever fra språklige minoriteter. (Kap. 3 Hva sier Leiter-R manualen om testing).

Alle de 6 sakkyndige vurderingene har utfyllende konklusjoner og flere forslag til tiltak. I det store og hele mener jeg at det er 6 gode sakkyndige utredninger. Jeg kan ikke se at disse 6 sakkyndige utredningene stemmer med det bilde Pihl (2002 a, b, c) har gitt i sinne beskrivelser av sakkyndige vurderinger av elever fra språklige minoriteter.

Jeg sier meg enig med noe av Pihls kritikk, for det er forbedringspotensialet i sakkyndige utredninger av elever fra språklige minoriteter, men det er ikke så svart som Pihl (2002) beskriver. Det er kun i den sakkyndige vurderingen av elev nummer 4 jeg mener at konklusjonen ikke holder mål fordi det ikke er tatt hensyn til elevens tospråklighet. Hos elev nummer 6 mener jeg eleven ville vært testbar med en dynamisk test. Det har ikke vært brukt dynamiske tester i noen av de 6 sakkyndige vurderingene.

Pihl (2005) berører et viktig poeng når hun kritiserer den utstrakte bruken av intelligenstester i utredningsarbeidet, men hvilke alternativer finnes? Hadde dynamiske tester vært mer tilgjengelige, mener jeg dette hadde vært et godt supplement i tillegg til bruken av statiske evnetester.

Analysen av de 6 sakkyndige vurderingene gir støtte for mine tre antagelser. WISC-III brukes ved utredning av elever fra språklige minoriteter, Leiter-R brukes lite og dynamiske tester brukes ikke.

## 2.2 Fagpersoner uttaler seg om dagens praksis

Fagpersoner i Norge ser at det er behov for å gjøre noe med utredningspraksis for elever fra språklige minoriteter. Den første fagpersonen jeg vil nevne er Hansen (2000) som skriver i sin artikkel "Hva innebærer dynamisk testing" at det er missnøye med dagens testpraksis. Hansen kommer her med forslag til hvordan PPT's utredningspraksis kan forbedres, ved å bruke dynamiske tester som et supplement til tradisjonelle statiske tester. Den andre fagpersonen er Birkemo (1996) som skriver om "...begrensninger ved den testingen som tradisjonelt anvendes i PP-tjenesten." (Birkemo 1996:21). Birkemo ser på mulighetene som ligger i å bruke en dynamisk tilnærming til testing. Både Hansen og Birkemo har jobbet i PPT og kjenner derfor godt til utredningspraksis. I sine artikler skriver de om problemer knyttet til dagens praksis med bruk av statiske tester ved utredning av elever fra språklige minoriteter. (Hansen 2000. Birkemo 1997).

Den tredje norske fagpersonen jeg nevner er Pihl (2002 a, b, c. 2005), med sin kritikk av dagens testpraksis av elever fra språklige minoriteter i PPT. Pihl (2002 a, b, c. 2005) har utført den største analysen av sakkyndige utredninger på elever fra språklige minoriteter i Norge. Analysen til Pihl består av 125 sakkyndige vurderinger fra to PP-kontor i Oslo, fra

---

1990-tallet. Phil ser på bakgrunnen for at minoritets elever henvises til spesialundervisning. Hovedfokus har Phil på hvordan PPT utreder minoritets elever. "Det viser seg at dette får avgjørende betydning for kategoriseringen av elevenes vansker og for opplæringen av elevene." (2005:18). Utredningene viser ofte til testresultatene og fremstiller dem som sannheter. (Pihl 2005). Det ligger en makt i "sannhetene" som de sakkyndige vurderingene produserer sier Pihl. Pihl kritiserer det hun kaller en

fordekt intelligenstesting av minoritets elever, fordi intelligenstestene kalles evnetester i de sakkyndige vurderingene.

Grunnlaget den sakkyndige vurderingen er basert på er avgjørende for hvilke "sannheter" pedagogisk-psykologisk tjeneste serverer til eleven, foreldrene og skolen. Eleven og foreldrenes kompetanse er som oftest forskjellig fra den sakkyndige fagpersonen. Denne kompetanseforskjellen mener jeg sakkyndig bør være seg bevist. Jeg mener man må være tro mot den makten som ligger i en sakkyndig vurdering, ved å ha et bevisst forhold til hvordan tester brukes og hvordan de tolkes. (kap.3. tolkningens historie). Anbefalinger som kommer frem i sakkyndige vurderinger bygger blant annet på testgrunnlag som gir "sannheter" om eleven og gir føringer for elevens fremtid.

Det er viktige problemer Pihl tar opp i sine analyser, men det er også de som er kritiske til noen av måtene Pihl har analysert og konkludert med i sine analyser av sakkyndig vurderinger. I medlemsbladet til NFPP (Norsk forening for pedagogisk psykologi) hadde de en av Pihls artikler på trykk. "Det er viktig at det forskes på PPT sitt arbeide slik at det fremkommer ny kunnskap som kan best mulig ivareta brukernes behov, d.v.s. de språklige minoriteter i denne sammenheng". (NFPP 2002:3). Samtidig er de kritiske til Pihl. "For det første er hennes artikkel lite nyansert bl.a. i den forstand at hennes konklusjon er helt i tråd med hennes syn i utgangspunktet. For det andre kan en virkelig spørre seg om hennes artikler er representative for PPT i Oslo" (NFPP 2002:3).

## 2.3 Hva sier de sakkyndige om egen praksis?

Hva er praksis hos de sakkyndige som utreder og tester minoritetsspråklige barn?

Er det noe problem ved utredningspraksis hos de sakkyndige? Eller er det bare noe vi tror siden det er mange minoritetsspråklige som henvises til spesialundervisning. (Pihl 2002. Hessels 1996. Lidz og Penà 1996) Empiri for de sakkyndiges utredningspraksis er blant annet å finne i en survey utført i Nord Colorado "Assessment and intervention practices with second-language learners among school psychologists." (McCloskey og Athanasiou 2000). Artikkelen omhandler en survey utført ved universitetet i Nord Colorado. Det er en ti punkts survey som undersøker skolepsykologers bruk av tester, prosedyrer og intervensjonspraksis ved utredning av tospråklige elever. (Betegnelsen skolepsykolog i artikkelen tilsvarer sakkyndige i PPT) Jeg har valgt å bruke deler av denne artikkelen som dokumentasjon på testpraksis blant sakkyndige, da jeg ikke har kommet over lignende norske survey. Det nærmeste jeg har kommet informasjonen i en slik survey er Pihl sine artikler (2002 og 2003). Pihls artikler analyserer innholdet og konklusjonene i sakkyndig vurderinger av minoritetsspråklige elever. I surveyen fra Nord Colorado får vi høre hva "skolepsykologene" selv sier at de gjør i sine utredninger. Det hadde vært interessant å få utført en lignende survey i Norge. McCloskey og Athanasiou (2000) sier i artikkelen at de har to mål med surveyen. Det første målet er å undersøke bruken av tradisjonelle og ikke tradisjonelle tester og intervensjoner ved utredning av tospråklige. Jeg vil understreke at McCloskey og Athanasiou ikke har utdypet hvilke type tospråklige de mener i sin survey. De sier selv i artikkelens siste del at de burde ha vært tydeligere på hva de mente med tospråklige. Men de har en god beskrivelse og drøfting om bevissthet rundt lingvistiske fenomen. Med det andre målet ønsker forfatterne å dokumentere, via surveyen, hvilke typer tester og intervensjoner som brukes i dag ved utredning av tospråklige. McCloskey og Athanasiou (2000) fant at flertallet av "skolepsykologene" brukte de tradisjonelle og utdaterte testene.

Tabell 5. Metoder de sakkyndige sier de bruker i utredning av elever fra språklige minoriteter. Det er kun deler av tabellen til McCloskey og Athanasiou (2000: 217) som er gjengitt her.

<b>Standardiserte tester</b>	<b>Ikke standardiserte prosedyrer</b>	<b>Alternative prosedyrer</b>
WISC 57%	Klasseromsobservasjon 79%	CBM 51%
TONI-3 43%	Lærerbedømmelse 77%	Dynamiske tester 29%
	Lære-foreldre intervju 55%	

I tabellen kan man lese at det ved utredning av elever fra språklige minoriteter er det 57 % av skolepsykologene som bruker WISC og kun 29 % som bruker dynamiske tester. Dette viser at det oftere blir brukt standardiserte tester ved utredning av elever fra språklige minoriteter enn det brukes dynamiske tester. Derimot er det de ikke-normstandardiserte prosedyrene som brukes oftest ved utredning av elever fra språklige minoriteter. De ikke-normstandardiserte prosedyrene er klasseromsobservasjon 79 %, lærervurdering 77 % og intervju med lærer / foreldre 55 %, dette er prosedyrer som vektlegger et helhetlig perspektiv av barnet. Ved å studere barnet i dens naturlige miljø samt høre lærerens vurdering av eleven får man et mer helhetlig bilde av barnet. (Lidz 2000. Rogoff 1990.)

Surveyen fra Nord Colorado viser at de statiske testene brukes oftere enn de dynamiske testene.

Misnøyen med standard tester er spesielt når det gjelder utredning av språklige minoriteter da det ikke finnes noen passende kognitive tester for disse elevene. (McCloskey og Athanasiou 2000).

## 2.4 Stereotyp behandling av elever fra språklige minoriteter.Hva sier foreldre om dagens praksis.

En stereotyp oppfatning er at ”Alle elever med mørk hud er tospråklige og har dermed vansker med det norske språket” I Utdanning nr 13 – 2004 stod det en artikkel om foreldre som måtte kjempet for å få barna ut av Norsk som andrespråk (NOA). Foreldre må da være fornøyde med at barna får et tilbud om NOA. Når barna får tilbud om NOA får de en mulighet til å lære norsk bedre. Ved å lære norsk bedre får barna forutsetninger som kreves for å følge den ordinære skolehverdagen. Men i artikkelen i Utdanning er det tre foreldre som forteller om hvordan de måtte slåss med skolen for å få barna sine ut av NOA. For det først ble de bare plassert i NOA uten at foreldrene hadde blitt informert. For det andre hadde disse barna hovedsakelig blitt plassert i norsk som andrespråk fordi foreldre ikke var norskfødte. For det tredje lå det ikke skikkelig faglig begrunnelse bak plasseringen av elevene i NOA. Foreldrene sier det ikke er NOA vi protesterer mot, men den manglende vurdering av behovene til den enkelte elev. Den ene av de tre foreldrene kunne fortelle at datteren som var født i Norge ble plassert i norsk som andrespråk fordi hun hadde problemer med norsk skriftlig. Siden datteren greide seg fint i norsk muntlig fikk foreldrene mistanke om at datteren muligens hadde dysleksi. De ba om en utredning av datteren hos en logoped. Foreldrene fikk bekreftet sin mistanke, datteren hadde dysleksi. Det blir også hevdet at barn med lese- skrivevansker, ADHD og andre problemer også plasseres i NOA, selv om det er andre tiltak disse barna har behov for. Er du mørkhudet, har utenlandsk navn eller foreldrene dine er utenlandske er det fare for at dine lærevansker blir identifisert som språkforskjeller og ikke språkvansker. Dette mener jeg er en farlig slutning å trekke. Det må ligge en skikkelig faglig og sikker begrunnelse bak plassering av minoritets elever. Jeg skriver ”sikker” fordi det er i utredningen det kan forekomme usikre slutninger på grunn av uskikkede tester. Når det gjelder plassering av elever i NOA er det skolen som i hovedsak tar denne beslutningen. En sakkyndig utredning kommer først inn når skolen eller foreldre ber om det fordi eleven har skolevansker. Artikkelen viser at man ikke må bli ”fargeblind” og at det trengs skikkelig faglig begrunnelse for plassering av elever i NOA.

---

## 2.5 Hvilke problemer er det med statistisk testing av elever fra språklige minoriteter?

"Low-income, minority children typically score poorly on traditional standardized tests, making valid identification of language disorders difficult. As a result, there is both considerable under- and overrepresentation of minority children in special education programs." (Ukrainetz m.fl. 2000:142).

I dette sitatet vises det til at elever fra språklige minoriteter skårer svakt på statistiske tester og at de både er over og under representert i spesialundervisningen. Dette viser til at det er problemer med statistisk testing av elever fra språklige minoriteter. Hvorfor skårer elever fra språklige minoriteter lavt på statistiske tester?

### 2.5.1 Elever fra språklige minoriteter stemples som mentalt retarderte.

Verbale ferdigheter spiller ofte en stor rolle i forståelsen av de instruksjonene som blir gitt i testen. Før man bruker tester for å måle elevenes generelle evner må man finne ut om testen måler det samme i de forskjellige gruppene. For eksempel kan testen måle hukommelse hos majoritetsbarn mens testen måler språkforståelse hos minoritetsbarn, da har testen en bias mot minoritetsbarna. (Hessels 1996). Testene måler ikke bare hvor godt barnet kan språket, men også hvor godt de kjenner og er del av den kulturen de bor i. (Hessels 2000). Det vil si at standard intelligens tester er kultur og språkavhengige.

"I Oslo intelligens testes og velges minoritets elever til spesialundervisning på grunnlag av WISC-R, som er standardisert for elever med norsk språklig og kulturell bakgrunn." (Phil 2005:107). Her sier Phil at den utstrakte bruken av intelligens testing av minoritets elever

fører til en konstruksjon av mentalt retardasjon. Flere av utredningene vurderer heller ikke behovet for tospråklig opplæring. ”Når testen anvendes på minoritets eleven som har et annet morsmål, og som knapt har begynt å lære norsk, representerer det et grovt brudd på testforutsetningene.” (2002:94) Dette har Phil helt rett i, men jeg mener at Pihl overser et viktig aspekt i bruken av intelligens testene, nemlig hvordan testene tolkes. Det er viktig at det fremkommer i den sakkyndige vurderingen hvordan testen har blitt brukt og tolkningen av resultatene.

## 2.6 Problemene løses ikke ved å fjerne testene eller sakkyndige vurderinger.

I USA har det vært rettsavgjørelser som har vært både for og mot bruken av intelligens tester på minoritetspråklige elever. (Sattler 2001). “..the courts may not be the best place to resolve complex issues concerning the fairness of tests and appropriateness of educational procedures.” (Sattler 2001:666). Løsningen er ikke en rettsavgjørelse eller å slutte med tester. Når det gjelder kritikken av WISC-III sier Kaufman ”And if we abandon IQ tests altogether, while waiting for the promised Holy Grail from respected cognitive scientists and ”wannabe” test developers such as Robert Sternberg or Howard Gardner, then we face a Black Hole in the interim.” (1994:25). Dette er en relativt sarkastisk kritikk fra Kaufman, men jeg er enig med at det ikke gjør saken bedre å slutte med testingen. Jeg mener at testene må sees på som et verktøy som er nyttig i kyndige hender. Til sist vil jeg avslutte dette kapitlet med et klokt sitat.

“Etter mitt syn er det ikke tilrådelig å fjerne kravet om at det må foreligge en sakkyndig vurdering hvis en elev fjernes fra ordinær undervisning og undervises etter en individuell læreplan.....Kritikken av sakkyndige utredninger av minoritets elever bør ikke besvares med å fjerne sakkyndig vurdering som forutsetning for vedtak om

---

spesialundervisning - det vil ytterligere svekke elevenes rettssikkerhet".(Pihl 2005:205)

## 2.7. Avslutning

Med elever fra språklige minoriteter mener jeg de som har et annet morsmål enn norsk.

Individuelle variasjoner, tospråklighetstyper og språkforståelsen er tre fenomen som viser mangfoldet i gruppen språklige minoriteter.

Jeg har undersøkt utbredelsen av språklige minoriteter i den norske grunnskolen. Av de elevene som henvises til PPT i Oslo er 35 % av elevene minoritetsspråklige. I Oslo får 10.5 % av de minoritetsspråklige elevene spesialundervisning. Dette mener jeg er en betydelig andel av den totale elevmassen.

Tester er en av metodene som brukes i en sakkyndig vurdering. Når en elev ikke har et tilfredsstillende utbytte av skolen sendes det et en tilmelding til PPT. Det er PPT som er ansvarlig for at det utarbeides en sakkyndig vurdering og vurderer om det er behov for spesialundervisning. Jeg mener det er veldig viktig å sette seg inn i hva testmanualen sier om testen. Tester har vært kritisert fordi testingen ikke har fulgt testspesifikasjonene. Jeg valgte å presentere testene WISC-III, Leiter-R og LEM for å vise hva testmanualene sier om testene. For å sikre seg at tester brukes riktig er det viktig å kjenne testene og vite hvorfor man skal teste med den testen man har valgt.

Det er motbevist at språklige minoriteter er dumme eller at tospråklighet hemmer den kognitive utviklingen. Er det noe med dagens praksis som gjør at elever fra språklige minoriteter fremdeles er overrepresentert i spesialundervisningen? For å få svar på dette spørsmålet undersøker jeg hva de sakkyndige og fagfolk sier om dagens praksis ved testing av elever fra språklige minoriteter. Dette tar jeg opp i neste kapittel.

### **3. KAPITTEL 3. PROBLEMSTILLING NR. 2: HVORDAN KAN MAN ELIMINERE FOREKOMSTEN AV FEILKONKLUSJONER I TESTING AV ELEVER FRA SPRÅKLIGE MINORITETER?**

#### **INNLEDNING**

I kapittel 2 tok jeg opp problemer ved dagens praksis ved utredning av elever fra språklige minoriteter. Problemer jeg viser til var feilkonklusjoner som kan føre til feilplassering og stempling av elever fra språklige minoriteter, fordi man ikke finner hva som er grunnen til elevens lærevansker. Å eliminere forekomsten av feilkonklusjoner i testing av elever fra språklige minoriteter er lite sannsynlig, men det er fullt mulig å ha det som et mål å arbeide mot. I dette kapitlet har jeg valgt ut noen emner som jeg mener er viktig for å redusere faren for feilkonklusjoner.

Jeg støtter meg til etiske prinsipper når jeg skriver hvordan tesene bør brukes. Testene bør brukes av kyndige fagfolk som er tydelige på hvordan testresultatene kan benyttes. I kapitlets hoveddel tar jeg for meg tolkning av testresultat. I tolkningsdelen er det en gjennomgang av hva testmanualene til WISC-III, Leiter-R og LEM sier om tolkning av testenenes resultat. Det finnes forskjellige tolkningstyper som tilpassning av testen, økologisktenkning, ipsative tolkninger, kritisk tenkning og intelligent testing med WISC-III. Jeg avslutter hoveddelen om tolkning med hvordan tolke i et flerkulturelt samfunn og hvordan bli en god tolker.

Opplæring og faglig oppdatering er viktig for å kunne være en kyndig tester. Kulturreduerte tester er viktig for ikke kun å måle språkkompetansen hos elever fra språklige minoriteter, men også få tak i elevens kognitive funksjoner. Kapitlets siste avsnitt viser til at dynamiske tester er en kulturreduert test som kan benyttes for å redusere forekomsten av feilkonklusjoner i testing av elever fra språklige minoriteter.

---

## 3.1 Hvordan tesene bør brukes

I de etiske prinsippene til NFPP står det tydelig hvordan PP-rådgiveren skal anvende og forholde seg til tester.

### **ETISKE PRINSIPPER NFPP**

#### § 3.5. Tester og utredningsmateriale.

3.5.1. PP-rådgivere administrerer tester i tråd med testens instruksjoner og veiledning for å bevare testens validitet. Dersom det forekommer avvik i testadministrasjonen, prosedyre eller validitet, skal det fremkomme i utredningsrapporten med tolkning av mulige konsekvenser.

3.5.2. PP-rådgiver tolker tester i lys av deres anvendelsesområde, begrensninger, normeringer, reliabilitet, validitet og andre etablerte standarder.

3.5.4. PP-rådgiver verner mot feiltolkning og missbruk av utredningsmaterieell og tar avstand fra at utredningsmaterieell brukes av ukyndige eller ukvalifisert personale.

Jeg støtter meg på disse etiske prinsippene i mine påstander videre i dette avsnittet. Jeg mener at en test kan brukes på andre måter enn det testens manual sier. Brukes testen på andre måter enn det som er testens spesifikasjoner må det fremkomme i testrapporten eller den sakkyndige vurderingen, hvorfor man har valgt å bruke testen og hvordan testen er brukt.

Testeren må være klar på at resultatene ikke kan tolkes i henhold til manualen fordi testen ikke er utført etter spesifikasjonene. Et eksempel kan vær at man tester en minoritetsspråklig elev med WISC-III for å finne ut hvordan eleven skårer i forhold til sine jevnaldrene. Dette kan sees på som en screening av eleven, men det er viktig å presisere at testen ikke må tolkes i henhold til testmanualen, da den ikke har fulgt testmanualens spesifikasjoner.

Tolkningen av testresultatene kan gjøres på forskjellige måter. Ved å bruke forskjellige tolkningsmetoder, mener jeg at test resultatene kan bli nyttigere enn ved tradisjonell tolkning. Eksempelvis ved å bruke ipsativ tolkning for å se hva som er elevens sterke og svake områder, eller vekstkåren i Leiter-R som viser de små variasjonene.

## 3.2 Tolkning

### 3.2.1 Hvordan tolke testresultater?

Hvordan tolke testresultater, er lite omtalt i litteraturen jeg har lest, i forhold til testenets psykometriske egenskaper.

Slik jeg ser det har hovedfokus vært på testens oppbygging og psykometriske egenskaper, mens tolkningen av testen har vært mindre vektlagt. Etter å ha lest bøkene "Intelligent testing with the WISC-III" (Kaufman 1994) og "Test interpretation and diversity" (Sandoval 1998) ser jeg betydningen av å ha et bevist forhold til og kompetanse på tolkning av testresultatene, fordi testresultatene alene sier lite om hva barnet har problemer med. Testmanualene er også tydelige på at resultatene fra en test alene ikke er nok for å diagnostisere en elev. ICD-10, som er et internasjonalt diagnosesystem, er også helt klar på at det må innhentes data fra flere kilder for å stille en diagnose.

Avsnittet starter med tolkningens historie som beskrives i fire bølger. Videre tar jeg for meg hvordan testmanualene til WISC-III, Leiter-R og LEM forholder seg til tolkning, samt se på testene i forhold til tolkningens fjerde bølge. Det finnes flere måter å tolke tester på, dette tar jeg opp under avsnittet om tolkningstyper som beskriver psykometriske forhold, økologisk tenkning, ipsativ tolkning, kritisk tenkning og intelligent testing med WISC-III. Det nest siste avsnittet tar for seg tolkning i et flerkulturelt samfunn, som er et mer og mer aktuelt tema i

---

dagens samfunn. Jeg avslutter med et avsnitt om hvordan bli en god testtolker og mine tanker om hva som kreves.

### **3.2.2 Tolkningens historie.**

Jeg mener at testingen og tolkningens historie har vært farget av et naturvitenskapelig syn. Den tradisjonelle testtenkningen har ofte vært preget av en årsak og virkning tenkning. (Robinson-Zanartu og Sloan Aganza 2000). Min erfaring fra praksisperioden i forbindelse med min utdanning og samtaler jeg har hatt med rådgiver i PPT har vist at det i dag ikke er en snever årsak og virkning tenkning. De 6 sakkyndige vurderingene jeg har analysert i kapittel.2 viser at det tas i bruk flere metoder for å innhente data om eleven. Konklusjonene i de 6 sakkyndige vurderingene mener jeg ikke har preg av en årsak og virkning tankegang. Det er et begrenset bakgrunnsmateriale jeg har for å trekke slutning om dagens praksis, men jeg mener det gir en indikasjon om at det tenkes bredere i dag.

I den testlitteraturen jeg har lest, har jeg funnet lite om tolkningens historikk. Hvordan har testresultat vært tolket opp gjennom tidene? Flanagan, McGrew, Ortiz (1999) viser til en firedeling av tolkningens historie. De kaller det for tolkningens fire bølger. De firebølgene er som følger:

#### 1. Kvantifisering av et generelt intelligensnivå.

Intelligenstesting ble sett på som en objektiv metode for å differensiere folk etter intelligensnivå og sette dem i forskjellige grupper. Gruppene kunne være idiot, imbesil, galing, etc.

Den skåren du oppnådde på testen, resulterte i et generelt intelligensnivå. Det vil si at testresultatet fra en test resulterte i en diagnose.

#### 2. Klinisk profilanalyse.

Her prøvde man å forstå individet ut i fra *mer* enn et generelt intelligensnivå. Det generelle intelligensnivået ble delt opp i subtester som igjen delte opp intellektet i flere enheter. Subtest profilene, diskrepans mellom verbal- og utføringsdelen og noen ganger også individets respons i testsituasjonen, kunne gi en diagnostisk og forklarende mening.

### 3. Psykometrisk profilanalyse.

Denne tolkningstilnærmingen ble muliggjort ved dataens inntog. Psykometrisk profilanalyse bygger på faktoranalyser og kan vise 75 forskjellige profiler. (nevropsykologisk, personlighet, lærevansker, ungdomskriminalitet, etc.) Problemet med psykometrisk profilanalyse er at den ikke har vist å ha empirisk støtte.

### 4. Teorianvendelse.

Behovet for å integrere teori og forskning ved tolkning av testresultater. Intelligenstester har manglet teorigrunnlag som kunne veilede praktikerne i sin tolkning. Denne nye tolkningstilnærmingen skifter fokus fra målingen av et bestemt testbatteri, til en teoridreven metode for organisering av testing og tolkning.

De fire tolkningsbølgene mener jeg gir en god illustrasjon på hvordan tester har vært tolket og hvordan de bør tolkes i fremtiden. Den første bølgen vil jeg hevde var med å plassere språklige minoriteter som mentalt retarderte. Den fjerde bølge har derimot med seg teori og forskning, som jeg tolker er med å åpne opp for økologisk tenkning. Jeg mener at de tre første bølgene hadde lite rom for kulturelle variasjoner og språklige minoriteter.

### **3.2.3 Hva sier WISC-III manualen om tolkning?**

Testmanualen til WISC-III sier ikke mye om tolkning. Det som står om tolkning er hva hver av deltestene kan indikere. Manualen har en tabell for hva som er signifikante skiller. Det vil si hvor stort sprik må det være på resultatene av deltestene for at skillet regnes som

---

signifikant. For videre tolkning viser manualen til Kaufman's bok "Intelligent testing with the WISC-III". I denne boken beskrives "Kaufmans sju steg" som er en metodisk gjennomgang av test resultatene. Her kan man undersøke om barnets testresultat viser til en jevn begavelse hos barnet eller en ujevnhet i barnets begavelse. Kaufman's sju steg for tolkning av et barns testresultat er også gjengitt i testmanualen. Ellers er manualen fri for diagnostisering og skåringskategorier.

Videre sier manualen at testadministrator må ta hensyn til barnets historie i sin tolkning av resultatene. Den som tester må ha kunnskap om reliabilitet og kunne bedømme om måleresultatene er reliable. Hvis det er store avvik i deltestene bør testadministrator kunne avgjøre om det er målefeil. Testadministrator må kunne sette testresultatene inn i en større psykologisk tolkningsramme. Testadministrator bør derfor ha kunnskap om kognitiv utviklingspsykologisk teori og kunnskap om utviklingsavvik. Det stilles høye krav til kompetansen hos testadministrator. "Psykologer är den enda yrkesgrupp som besitter den kompetens som är nödvändig för att WISC-III ska kunna användas på ett professionellt och omdömesigillt sätt och WISC-III är således ett verktyg för psykologer". (104). Jeg forstår testmanualen slik at testresultatene fra WISC-R skal tolkes av kompetente personer som metodisk gjennomgår testresultatene og ser disse i et helhetlig og økologisk perspektiv.

Flanagan, McGrew, Ortiz (1999) hovedmål med boken "The Wechsler intelligence scales. Gf-Gc theory" er å modernisere tolkningen av Wechslers intelligens skala og å føre Wechsler skalaene over til den fjerde bølge, ved hjelp av en metode som veileder tolkningen. De sier videre at Wechsler trilogiens nye revisjoner for det meste har vært av kosmetisk art. Testens mangel på teori har ikke gjort det mulig å føre Wechsler skalaen inn i den fjerde bølge hevder Flanagan, McGrew og Ortiz (1999).

Jeg mener WISC-III i dag har tilgang til teori og forskningsmaterialet som kan veilede praktikerne i sin tolkning, og føre testen over i den fjerde bølge. Dette kan gjøres ved å integrere teori fra Kaufman(1994), ha et økologisk perspektiv og være oppdatert på forskningsfeltet.

### 3.2.4 Hva sier LEITER-R manualen om tolkning?

Manualen har et eget kapittel som heter tolkning. Den viser til en hierarkisk metode for tolkning. Manualen beskriver også hvordan man først må skaffe seg bakgrunnsinformasjon om barnet, kultur, barnets fysikk etc. før man starter testing og tolkning. Leier-R har fem forskjellige skårings nivå/ måter:

Non-verbal IQ skåre

Vekstskåre

Composite score

Subtest scaled scores

The AM battery special diagnostic score.

Man velger en eller to av skårings måtene, avhengig av hva hensikten med testingen er. Hvorfor skal man teste, hva er målet med testingen? Svaret på dette avgjør hvilke tolknings måter man velger. *Skal du slå i en spiker velger du en hammer og ikke en sag.*

Det er nok den normrefererende IQ skår som er mest kjent og mest brukt av de fleste. Forfatterne av Leiter-R uttrykker et sterkt ønske om at den kriterierefererende vekstskåren burde blir tatt mer i bruk som et supplement til IQ. Dette gjelder spesielt for barn som henvises til spesialundervisning. Ved å bruke vekstskåren ser man de små endringene. Det spesielle med vekstskalaen er at den gir en fingradert informasjon om hvor langt barnet har kommet i sin utvikling på de egenskapene som Leiter-R måler. Standardiserte norm - refererende skalaer har fått mye kritikk for at de ikke gir detaljert informasjon om de små endringene i hvert enkelt individ. Leiter-R sin kriterierefererende vekstskåre er en metode som gir detaljert informasjon om individets vekst.

Kriteriereferansen til Leiter-R kan sammenlignes med høydemåling. Et barn som er lavere en jevnaldrene vil alltid være lavere enn aldersgjennomsnittet. Måles barnet i forhold til vekstskalaen kan man si noe om barnet har en normal vekstutvikling, selv om barnet alltid vil være lavere en gjennomsnitte. Hvor på metermålet er barnets høyde? Eller hvor på

---

vekstskalaen har barnet utvikling nådd. Barnet fremskritt kan måles i forhold til egenvekstprogresjon, sannsynligvis har barnet en normal vekst i forhold til seg selv.

Vekstskåren sier noe om barnets vekstpotensial. Vekstskåren er en lineær skala, fra 400 til 580, som viser oppgavens vanskelighetsgrad. Barnets IQ poeng pr oppgave kan regnes om til vekstskåre som viser til hvilke oppgaver barnet vil skåre 50 % riktig på. Det er anbefalt å planlegge intervensjoner for barnet på +20 vekstskåre, for da har barnet noe å strekke seg til og det er innen barnets mulige utviklings område. Det er 90 % sannsynlig at barnet vil skåre riktig på oppgaver som ligger 20 poeng under vekstskåren. Ligger derimot oppgaven 20 poeng over vekstskåren er det kun 10 % sjans på at barnet skårer korrekt. Oppgaver som ligger mellom 0 og 20 poeng under kalles; ”instruksjons område” oppgavene er ikke for vanskelig, men barnet trenger mer trening for at det skal mestre oppgaven. Dette mener jeg har likhetstrekk med sonen for den nærmeste utvikling. (Kapittel 4.1 Teorigrunnlaget i dynamiske tester.).

Vekstskåren er et alternativ til statisk testing som kun måler barnet i forhold til hvordan en bestemt populasjon skårer. Grunnlaget for vekstskårene er fra alle skårene og ikke bare normgruppen. Vekstskåren gir mulighet til å finne egenarten ved hvert barn, dette danner et bedre grunnlag for planlegging av tiltak som kan hjelpe barnet. Dette passer bra for barn som fungerer på et lavt evnenivå og eller barn som forventes å bli re-testet på Leiter-R.

” In other words, the growth scale provides an ”anchor scale” for referencing all subtest and composite score from the Leiter-R.” (Leitr-R:94).

Jeg mener det er likhetstrekk med vekstskåren og Kaufmans ipsative tolkning. (beskrives senere i dette kapittelet). Likhetstrekkene er hvordan barnets sterke og svake sider kommer frem. I testmanualen til Leiter-R står det om vekstskåren: ”This knowledge should assist in planning intervention programs and determining strengths and weaknesses.” (Leiter-R :105). Dette mener jeg er akkurat slik som Kafman(1994) beskriver fordelene ved ipsative tolkninger. Han viser til at ipsative tolkninger sier noe om hva som er barnets sterke og svake sider.

Leiter-R mener jeg passer inn i den fjerde bølge fordi testen bygger på en teori. Teorien om krystalisert og flytende intelligens. ”The Leiter-R was designed to measure a hierarchical ”g” model that recognizes the many faceted nature of intelligence, combined whit the practical need for a global score.” (102). Leiter-R bygger på samme teori som Flanagan, McGrew, Ortiz (1999) ønsker å bruke på WISC-III.

### **3.2.5 Hva sier LEM om tolkning?**

Jeg har fått til gang til den Nederlandsk versjon av LEM testen. Test manualen har et halvt kapittel om tolkning av testresultat. Det første testmanualen tar for seg er totalskåren, det vil si LP-skårenes hovedkategorier. Totalskåren viser til om barnet har et høyt, over middels, middels, undermiddels eller lavt læringspotensial. Barnets LP-skåre sier noe om hvilke tilrettelegging det trenger for å kunne få utviklet sitt potensial. (Kap1.6.3 LEM).

Testmanualen har også tabeller som viser hva de forskjellige deltestene måler.

LEM testen bygger på et teorigrunnlag om læringspotensial (Kapittel 4.2), jeg mener derfor den dermed går over til den fjerde bølge.

Testen bygger på en teori, men det kan ikke garantere at resultatene tolkes i henhold til den fjerde bølge. Det er den menneskelige faktoren som spiller den avgjørende rollen.

### **3.2.6 Tolkningstyper.**

Tester som er *standardiserte* har til hensikt å kontrollere testforholdene slik at de som testes får så lik behandling som mulig. Modifikasjoner av standardiseringen må tas i betraktning

ved testing av elever som ikke passer inn i normeringsgruppen. Dette gjelder spesielt for hvordan endringene i testens standardisering kan virke inn på skåringen og validiteten i disse skårene. (Geisinger 1998).

En og samme test kan tolkes på forskjellige måter. Den tolkningsmåten som velges avhenger av hva som er testingens hensikt og hvem som skal testes.

### **Psykometriske forhold ved tilpassning av tester.**

"Test users and test interpreters must become extremely careful in their reading of test manuals if they are to use tests with and to interpret test scores of members of diverse groups. (Geisinger 1998:19). Her finner jeg endelig stoff som jeg har tenkt mye på når det gjelder bruk av tester. Her fremheves det viktigheten av å lese testmanualen nøye og sjekke om den eleven som skal testes er representativ for normeringsgruppen. Ved å lese testmanualen nøye finner man ut om testen er egnet for den eleven som skal testes. Testmanualene pleier også å omtale hvordan testen er for spesielle grupper. (Geisinger 1998). Både WISC-III og Leiter-R omtaler særskilte grupper. WISC-III har et kapittel som viser til studier av særskilte grupper og hvordan disse har skåret på testen. I Leiter-Rs testmanual står det at testen passer for spesielle grupper. Hva som legges i begrepet 'special population' avhenger av hvem man snakker med. Leiter-R har en definisjon av begrepet, det er følgende: " ..any person between the ages of 2 and 20 years, who you for documented reasons, may not obtain an accurate evaluation from a traditional, verbally-loaded cognitive measure, is included in the term "Exceptional Population" as used here." (Leiter-R 1997:76).

Når testmanualene viser til spesielle grupper sier de også noe om skårenes reliabilitet og validitet for disse gruppene. Hvis det viser seg at testen ikke er egnet for den eleven som skal testes bør det undersøkes om det finnes andre tester som er bedre egnet. Velger man likevel å bruke testen, selv om den er uegnet, må man være bevist at skårene kan ha fått svakere validitet og reliabilitet.

"When testing professionals make changes in either testing materials or test administration, they are appreciably changing the nature of an examination. As such, many of the psychometric aspects of the test in its particular use also may become changed. Test users must assess the continued reliability and validity of the examination once it has been changed." (Geisinger 1998:29).

### **Økologisk tenkning.**

Det heter seg at, summen av delene er mer enn helheten. Det vil si at vi ikke bare kan legge sammen resultatene fra tester og summere. Vi må se på interaksjonen mellom delene når vi skal tolke testresultatene. Endringer i en av delene kan føre til at helheten endrer seg. En slik helhetlig tenkning beskrives som `økologisk tenkning`. (Robinson-Zanartu og Sloan Aganza 2000). Begrepet `økologisk tenkning` drar parallell til et økosystem der hver enkelt bestanddel lever i et gjensidig avhengighet forhold til hverandre.

Både WISC-III og Leiter-R manualen skriver at testadministrator må ha et helhetlig blikk og innhente bakgrunns informasjon om den de skal teste.

Innen statisk testtradisjon tenkes det ofte universelt i årsak og virkning. (Robinson-Zanartu og Sloan Aganza 2000). Psykometrien tar ikke for seg konteksten som barnet lever i og hvilke kulturell bakgrunn barnet har. Kulturen har en avgjørende betydning for hvordan barnet tenker og oppfører seg. (Bråten 1998. Rogoff 1990).

For å finne måter å forstå kulturen på, kan man gå til sosialkonstruktivistene og sosialkulturalistene. (Robinson-Zanartu og Sloan Aganza 2000). Det konstruktivistisk perspektivet på læring er at vi konstruerer våre meninger ut fra våre erfaringer. Elevene skaper sin egen kunnskap gjennom deltagelse i meningsfulle oppgaver. Målet er å utstyre elevene med metakognitive verktøy som de kan bruke etter behov. Konstruktivistene tar kun for seg individet. Hva med samfunnet som individet lever i? Det sosialkulturelle perspektivet tar derimot med seg samfunnets og individets gjensidige påvirkning. Det sosialkulturelle perspektivet er en parallell til Rogoff (1990) "guided participation" og

---

Vygotskys intra- og interpsykologisk, først mellom mennesker, så i mennesket. (Bråten 1998, Vygotsky 1978). Sosialkonstruktivisme sier at det ikke finnes et enkelt perspektiv, men flere likhetstrekk. Eksempler på likhetstrekk kan være: Vær kritisk til ”tatt for gitt kunnskap”, ha fokus på at verdensforståelsen farges av kulturen og historien, ”sannheter” er sosialt konstruert og husk at kunnskap og sosiale handlinger henger sammen. (Robinson-Zanartu og Sloan Aganza 2000).

Comas-Diaz og Ramos Grenier (1998) argumentasjon understreker viktigheten av kulturell sensitivitet, bedre tilpassning av testenes innhold til målgruppen og den kulturelle konteksten de befinner seg i. Det vil si at kultur, tilpassning og kontekst er viktige faktorer i utredningen og testingen av elever fra språklige minoriteter. Med andre ord må det innhentes omfattende informasjon om elevens bakgrunnshistorie. Hva er elevens opprinnelsesland og hvilke kjennetegn har kulturen der? Hvorfor flyttet denne eleven fra hjemlandet sitt? Hvilke trosretning har elevens familie og eleven selv? Hvor lenge har eleven bodd i landet? Dette er spørsmål jeg mener må stilles når en minoritetsspråklig elev skal utredes. En omfattende undersøkelse av elevens historie mener jeg må til for å få et helhetlig bilde av eleven, fordi våre erfaringer er med å forme oss. (Vygotsky 1978).

Det jeg har skrevet om i dette avsnittet mener jeg tydelig viser hvor viktig det er å utføre en grundig undersøkelse av elevens bakgrunn for å få et helhetlig bilde av eleven. Dette helhetlige bilde av eleven er en del av grunnlagsmaterialet som legges til grunn for den sakkyndige vurderingen.

## Ipsativ kartlegging og tolkning.

En av hovedforskjellene mellom statiske og dynamiske tester er hva barnets testresultat, skåren, måles i forhold til. Statiske tester har som oftest et normeringsgrunnlag som består av en populasjon, og denne populasjonen måles barnet i forhold til. Det vil si hvordan skårer barnet i forhold til gjennomsnittet for normeringsgrunnlaget. Dynamiske tester derimot måler barnet i forhold til egne prestasjoner, de måler barnets læringspotensial. (Lidz 2000). Normeringsgrunnlaget i de statiske testene er ikke representativt for mange av barna fra språklige minoriteter. Slik jeg forstår det fremstilles denne forskjellen mellom statiske og dynamiske tester som en av de avgjørende grunnene for ikke å bruke statiske tester på språklige minoriteter. Dette mener jeg er en lite nyansert måte å skille to testtyper på, fordi statiske tester kan også brukes til å måle skåren i forhold til barnets egne prestasjoner. Å måle barnets skårer i forhold til barnets egne prestasjoner kalles for ipsative. Hva betyr ipsative?

### **Ipsative assessment:**(ipsativ kartlegging)

Type of assessment whereby the norm against which assessment is measured is based on prior performance of the person being assessed - the present performance is assessed against performance in the past. In athletics, "personal best" is an example of ipsative assessment.

(<http://www.swap.ac.uk/learning/glossary.asp?initial=>)

Slik jeg forstår forklaringen av 'ipsative' er at barnets skåre måles i forhold til tidligere skåre, det vil si at man ser etter om det er en endring i skåren. Dynamiske tester måler også om det er endringer i skårene i forhold til tidligere prestasjoner. Ipsative tolkninger og dynamiske tester virker veldig like, men det er en avgjørende forskjell. Dynamiske tester har en eller annen form for bevist trening innebygd i testen. Treningen som er innebygd i dynamiske tester er et verktøy som er med å finne utviklingspotensialet til barnet. Slik jeg forstår forklaringen på "ipsative assessment" er dette en tolkningsmetode for å se om barnet har endret seg.

Tidsaspektet er til stede i både "ipsative assessment" og flere av de dynamiske testene, for det er før og nå skårer, eller post og pre-test, som sammenlignes. Kaufman (1994) derimot snakker om ipsative tilnærming til tolkningen av delskårene i WISC-III. Delskårene i WISC-III har ikke tidsaspektet før og nå skåre, fordi delskårene måles i en og samme test. Trening er heller ikke innebygd i delskårene. Det ipsative som Kaufman fremmer i forhold til WISC-

---

III er at barnets skårer skal måles i forhold til barnet eget gjennomsnitt. Det vil si at skåren ikke skal måles i forhold til normerings gjennomsnitte som er 10, men det skal regnes ut hva som er barnets eget gjennomsnitt. Barnets sterke og svake sider kommer tydeligere frem når man ser på skårene ipsativt. (Kaufman 1994). Kaufman er en forkjemper for ipsative tolkninger.

Det har vært mye kritikk av ipsative tolkninger fordi de ikke har psykometriske kvaliteter som reliabilitet og prediktiv verdi, men ipsativ tolkning kan ikke settes inn i statistiske termer. (Kaufman 1994). Hvorfor ipsative tolkninger ikke kan settes inn i statistiske termer kan blant annet forklares på følgende måte;

"An individual's rating of the relative importance of an item compared with other items. Ipsative scores do not allow for comparisons among people; e.g., Lloyd likes football better than basketball, while Bonnie likes basketball better than football. Does Bonnie like basketball better than Lloyd likes basketball? You do not have enough information to answer that question. Hence, when using intentions or preferences to forecast, ipsative scores can be misleading and difficult to interpret. Guard against this problem by finding other ways for framing questions."  
(<http://morris.wharton.upenn.edu/forecast/data/definitions/ipsative%20score.html>)

## **Kritisk tenkning ved tolkning av tester**

Regner man seg selv for å ha et åpent og fordomsfritt sinn bør man kanskje tro om igjen. Kognitive fenomen kan påvirke oss i hverdagen. (Sandoval 1998).

De avgjørelsene og bedømmelsene vi tar, kan være fulle av stereotype forestillinger og skjema som våre tanker sporer inn i. "It is well accepted that human decision making is influenced by strategies and heuristics that are prone to error...The intent is to avoid the biases in human cognition that lead to poor decision making." (Sandoval 1998:31).

Bestemmelser som tas på bakgrunn av velutviklede og valide tester er overlegne de bestemmelser som baserer seg på intuisjonen til de sakkyndige. (Sandoval 1998). Påstanden til Sandoval mener jeg kan begrunnes i at tester fri for kognitive fenomen som kan føre til bias. Det vil si at tester kan være til god hjelp. Når man har med grupper som ikke er representative for flertallet eller normeringsgruppen som testen baserer sine skåringer på, er det viktig å finne gode alternativer. Da må testresultatene tolkes slik at man unngår de kognitive strategier som kan føre til bias. Selv ikke den mest erfarne sakkyndige er fri for feilbedømmelser hevder Sandoval (1998). Det er flere kilder som kan føre til feilbedømmelser, dette er noe vi ofte ikke er klar over. Sandoval sier at vi lett feilbedømmer hvor representativt tilfellet vi bruker for å trekke en slutning er og hva størrelsen på utvalget er. Det er viktig at vi i slike tilfeller sjekker med statistikken, fordi mennesket har en tendens til å bekrefte eksisterende kognitive skjema. Vi ser korrelasjoner som ikke eksistere, fordi de passer i våre kognitive skjema. De korrelasjonene som ikke passer våre kognitive skjema overser vi. Det siste vi har sett eller hørt er det mest tilgjengelige og dermed er det med å påvirke vår avgjørelser. (Sandoval 1998).

*Sandoval (1998) har laget en liste på 6 punkter som er hans anbefalinger å bruke ved kritisk tenkning, når man tester individer fra ulike grupper. Denne listen skal være en hjelp til at vi ikke benytter oss av kognitive strategier som kan føre til bias.*

1. *Identifiser dine forhåndsantagelser.* Identifiser de forestillinger og synspunkter du har om forskjellige grupper. Vær klar over at du har disse forestillingen. Stol på teoriene dine men vær klar over at de kan føre til selektive persepsjon, kognitive begrensninger og feilkilder.
2. *Utvikle komplekse skjema eller antagelser om grupper.* Hensikten med trening i kultursensitivitet er å forhindre stereotyp tenkning, slik at man får mer komplekse skjema om de forskjellige gruppene.
3. *Prøv bevist å finne unntakene som motbeviser reglene.* Våre kognitive skjema lurer oss til å overse det som ikke passer inn. Se om det er noe i konteksten som kan være årsaken til variasjoner og ikke alltid attribuere feilene til noe i individet.

---

4. *Motstå å komme til en rask konklusjon.* Vent til alle fakta er på bordet. Reflekter over hva som kan skyldes hva og ikke aksepter det første som faller deg inn.

5. *Søk råd* hos andre som har mer kjennskap til den gruppen du har undersøkt. Har andre mer kunnskap om den elevgruppen du undersøker, bør du søke råd hos en erfaren kollega. Det er viktig å diskutere de avgjørelser man har kommet fram til med kollegaer.

6. *Ikke stol på hukommelsen.* Skriv ned den informasjonen som skal brukes i forbindelse med en senere tolkning. Vi har lett å ha endret på ting i hukommelsen.

Til slutt er det viktig å være bevist at testdata kun er en liten del av eleven og at mange faktorer kan påvirke elevens liv.

De som skal tolke et testresultat må bruke hodet og være seg bevist sine kognitive fallgruver slik at eleven får en rettferdig og sikker vurdering. Jeg avslutter dette avsnittet med et sitat fra Sandoval "In working with clients from diverse backgrounds, all of psychologists' critical problem solving skills will be needed. There are no cookbooks or simple actuarial tools to fall back on. Clients deserve the clearest thinking."(1998:47).

## **Intelligent testing med WISC-III**

WISC-III er som tidligere nevnt en av de mest brukte tester ved utredning av elever. Jeg har derfor valgt å ta med et eget avsnitt fra Kafmans bok "Intelligent Testing with the WISC-III." (1994). Denne boken tar for seg forskjellige tolkningsmåter av WISC-III.

Etter hvert som jeg leste boken ble jeg mer og mer overbevist om at WISC-III er et godt verktøy, eller er Kaufman bare en dyktig skribent? Kaufman dreper all kritikk av WISC-III med at den som bruker testen må vite hva som er testens begrensninger og muligheter. Det avgjørende er testadministrators evne og kunnskap om tolkning av tester. Testen blir ikke

bedre enn testadministrator. Jeg må si meg enig med Kaufman (1994) i at det er ikke selve testen det kan være noe galt med, men mest sannsynlig testadministrators bruk av testen.

Videre sier Kaufman (1994) at hvis alle som testet med WISC-III legger tolkningens 5 premisser til grunn for sin tolkning ville det ta livet av mesteparten av kritikken mot WISC-III. Det er også nødvendig å lese testemanual for å vite hva testen er beregnet for.

#### WISC-III tolkningens fem premisser.

1. Noen av deltestene måler *hva som er lært til nå*.

Testen måler hva som er lært til nå, det er derfor ikke et spørsmål om arv eller miljø. Det som er lært til nå er lært i en kulturell setting, det vil si at læring er farget av kulturen. Man kan dermed ikke hevde at testing av hva man har lært til nå er kulturfritt.

2. Deltestene i WISC -III er *eksempler på adferd*. Testen sier ikke noe om hvordan adferden er under normale forhold. WISC-III alene sier svært lite om et barns evner og muligheter. Testen bør kombineres med andre tester og tolkes i henhold til annen data. Man skal aldri stille en diagnose på bakgrunn av en enkelt test. Den totale IQ skåren bør ikke brukes isolert.

3. WISC-III tester mentale funksjoner under *kontrollerte eksperimentelle forhold*.

Den standardiserte administrasjons prosedyre for WISC-III sikrer en objektiv evaluering av barna, men den savner den dype forståelsen av kognitive prosesser som kan oppnås gjennom andre tester som for eksempel Feuersteins test – trene – re-teste dynamiske test. Testadferden er et viktig kriterium for den senere tolkning av barnets sterke og svake sider..

4. WISC-III oppnår *optimal tolkning når informasjons - prosesserings modell* tas i bruk.

Informasjons - prosesserings modellen viser hvilke prosesser som er involvert i læringsprosessen. Modellen består av fire prosesser, input - integrering - lagring - output. Identifiseres det prosesser som barnet har vansker med, hjelper det oss i å finne hvilke funksjoner barnet sliter med. Den nye strukturen i WISC-III har fire faktorer, disse fire faktorene korresponderer med informasjons - prosesserings modellens fire prosesser.

*Input: Oppmerksomhet.*

*Integrering:* Verbal forståelse og perseptuell organisering

*Lagring:* Verbal forståelse

*Output:* Hurtighet

5. Hypoteser som lages på grunnlag av WISC-III profiler bør være støttet av data fra flere kilder. *En hypotese trenger data fra flere kilder*, for som Kaufman sier ”, test scores can mislead just as easily as they can lead.” (1994:13). Det er viktig å innhente bakgrunnsinformasjon om barnet, klinisk observasjon av adferden og andre test resultat. Data og informasjon fra andre kilder kan verifisere hypotesen fra WISC-III profilen.

Intelligenstester måler evner og har vist seg å være en god prediksjon for skoleprestasjoner. Får en elev et lavt resultat på en intelligens test skal det ikke tolkes som en fremtidsdom, men at her bør det settes i gang tiltak som kan hjelpe eleven. Kaufman sier at WISC-III bør sees på som en ”helping agent”.( 1994:7) og ikke en skjebne.

### **3.2.7 Tolkning i et flerkulturelt samfunn**

I dette avsnitte baserer jeg meg på boken, "Test interpretation and diversity" (1998). Boken er skrevet på oppdrag fra American Psychological Association (APA) Committee on Psychological tests and assessment.

Sandoval (1998) viser til at det finnes etiske retningslinjer som sier at man skal være bevisst de kulturelle forholdene og at det trengs tilpassninger i administreringen og tolkningen av tester for språklige minoriteter. Disse retningslinjene sier ikke noe om hva man kan gjøre for å utbedre administreringen og tolkningen av tester. "As useful as standards and ethical

principles are, they are clearly not enough to prevent destructive practices." (Sandoval 1998:6). Sandoval refererer til eksempler på feil diagnostisert. På bakgrunn av feil diagnostisering hevder Sandoval (1998) at det er behov for flere retningslinjer ved tolkning av tester. Hvis jeg har forstått Sandoval (1998) rett, er boken ment å gi retningslinjer for tolkning av tester.

De fleste tester brukes for å ta en eller annen form for avgjørelse. Derfor må noe av det som guider testtolkningen være *hvilke konsekvenser* har testresultatet for personen. (Sandoval 1998).

Hvis eleven ikke passer til normeringsgruppe kan det være fristende å ikke bruke testen. Faren ved å ikke bruke tester er at eleven kan fratas tilbud og muligheter. En av bokens antagelser er at personer kan få en mer rettferdig behandling ved klok og forsiktig bruk av uegnede tester enn å ikke bruke tester i det hele tatt. "How can tests be interpreted when they are used in less-than-optimal situations?" (Sandoval 1998:7).

Normer i seg selv er ikke noe problem. Problemet er hvordan man skal tolke skåren til en elev som er forskjellig fra normeringsgruppen testen er utviklet fra. (Sandoval, J. 1998).

"Although this problem in test interpretation is particularly acute for linguistic minorities,....., and from distinct subcultures also is fraught with difficulty. The technical standards for educational and psychological tests address the need for developers, test publishers, and test users to exercise caution when testing nonnorm group populations, but little guidance is actually provided to psychologists in applying standards or ethical principles." ( Sandoval 1998:5).

Sitatet viser at den tekniske standarden på testene krever at testbrukeren må være forsiktig når de tester elever som ikke tilhører normeringsgruppen. Det er første gang jeg har sett at noen har skrevet direkte at testbrukeren må være forsiktig med testenes tekniske standard. Jeg forstår det slik at testens tekniske standard lett kan føre til at man glemmer å sjekke om eleven og normeringsgruppen er kompatible. Hvis dette ikke sjekkes er det fare for at eleven

---

blir sammenlignet med en gruppe som er forskjellig fra elev selv. Resultatet av en slik sammenligning kan føre til at eleven blir feilplassert og feilaktig stemplet.

For å illustrere hva feil bruk av normer kan føre til har jeg laget et eksempel fra den medisinske verden. Tenk deg at det er en norm for lungekapasiteten til toppidrettsutøvere. Hvis denne normen brukes på pensjonister er det stor sannsynlighet for at de fleste pensjonistene ville blitt akutt innlagt på sykehus på grunn av faretruende svak lungekapasitet. Dette eksemplet mener jeg viser hva som kan skje når det brukes en norm som ikke passer til den personen som skal undersøkes. Disse pensjonistene ville med størst sannsynlighet blitt medisineret og behandlet på grunnlag av sin ekstremt svake lungekapasitet. Hadde pensjonistene derimot vært målt i forhold til sin egen normerings gruppe ville det vist seg at de ikke trengte behandling. Slik mener jeg det også kan være med minoritets språklige elever som måles i forhold til en normeringsgruppe som ikke passer dem. Den minoritetsspråklige eleven kan på bakgrunn av feil norm bli behandlet med uegnede tiltak.

Fokuset bør være på individet og ikke gruppen, fordi det kan være store variasjoner i gruppen. Samtidig er det fare for stereotyping på gruppenivå, derfor ønsker de å fokusere på individet. Hvert individ er unikt, så derfor trengs det individuell tolkning og behandling. Det finnes noen sosiokulturelle og psykologiske faktorer som ofte blir assosiert med grupper, og som virker inn på testprestasjonene. (Sandoval 1998) Disse faktorene er tospråklighet, fattigdom, hvor lenge har de vært i landet og funksjonshemninger. Dette er faktorer som testadministrator bør ta med i tolkningsprosessen. Sandoval (1998) anbefaler ikke en test fremfor en annen. I stedet fokuserer Sandoval på tolkningen og at denne prosessen bygger på psykologisk forskning og teori. Dette mener jeg passer med den fjerde bølge som ble nevnt unger tolkningens historikk.

Tolkningen må reflektere en kritisk tenkning i stede for kokebok tenkning. Testadministrator bør ha en forsker - praktiker perspektiv. Det vil si at testadministrator ser hvert individ som unikt slik en forsker ser hvert problem som unikt og anvender vitenskapelige metoder.

### **3.2.8 Hvordan bli en god testtolker**

I dette avsnittet, hvordan bli en god test tolker, er det snakk om folk som er ferdig utdannet og i arbeid med tolkning av tester.

Det er viktig at den som tester er fleksibel og ikke holder seg til en testretning. Lange ventelister og mange forskjellige barn som skal utredes er hverdagen til utrederne. Det er ofte ikke tid til å bruke mange tester eller dynamiske tester som modellene til Feuerstein eller Budoff. (Kaufman 1994)

Kaufman (1994) mener at arbeidsmengden og press om produksjon ikke skal få skylden for dårlige tolkninger av testene eller annen missbruk. Uavhengig av arbeidsmengde er det mye en samvittighetsfull testadministrator kan gjøre for å få til en intelligent tolkning. (Kaufman 1994) *Kaufman har en liste på tolv punkt* som viser hva testadministrator kan gjøre. Følges disse tolv punktene vil man bli en intelligent tester til tross for stort arbeidspress, rigid skolepolitikk eller offentlige pålegg. (Kaufman 1994).

Tolv punkter som er ment å skulle lette intelligent tolkning er følgende:

1. Informer skolen og ledelsen samt dine kollegaer om hva som er intelligenstestenes begrensning og riktige bruk.
2. Behandle oppgavene i testen som eksempler på adferd som er målt under kontrollerte eksperimentelle forhold. Testen måler hva barnet har lært til nå.
3. Tolk de innsamlede data etter informasjons - prosesserings modell sammen med data fra andre kilder.
4. Ikke overvurder IQ skåren eller tro et det er barnets medfødte potensial.
5. Lær en tolkningsmetode så den kommer godt under huden. Metode bør fremme fleksibilitet, gi en systematisk modell for å ta fatt i profilene og oppmuntrer til å gå utover skårene for å forstå barnet.
6. Bli kjent med nabolagets kultur og språkbruk. Hvordan er skolebarnas hverdag? Med andre ord utvid dine skjema.

- 
7. Bli kjent med lærerne ved å observere dem i klasserommet og samarbeid med dem når det gjelder oppfølging og modifier anbefalingene i rapportene.
  8. Vær bevist potensialet til eksisterende og nye tester.
  9. Vær bevist på begrensningene i standardiserte tester og testers evne til å tolke individer som de ikke kjenner den kulturelle bakgrunn til. *I dette og foregående punkt mener jeg det er en klar hentydning til at man må sette seg godt inn i testens manualer.*
  10. Hold deg oppdatert på forskning og teorier i psykologi og relaterte områder da de kan gi deg ny innsikt i tolkning av profiler. *Dette punktet mener jeg er med å føre tolkingen over i den fjerdebølge ved at man tar i bruk nyere forskning og teori.*
  11. Bli kjent med curriculum baserte tester, dynamiske tester og eller andre alternativer til standardiserte intelligens tester. Lær hvordan disse alternative testene kan gi supplerende data til WISC-III og andre normerte instrument.
  12. Lær deg hvilke læreplan og undervisningstilnærminger som brukes på skolen. Hold deg oppdatert på nytt materiale, prosedyrer og bruk denne kunnskapen når du anbefaler tiltak som baserer seg på utredningsresultat.

Følges disse tolv punktene hevder Kaufman (1994) at man skal bli bedre til å tolke tester. Jeg er enig med Kaufman (1994) at dette er tolv viktige punkter for å bli god til å tolke test resultat.

Det jeg savner i Kaufmans liste er noen ord om kritisk tenkning og selvrefleksjon. Vår medfødte egenskap til kognitiv bias i det vi ser bør være et sentralt tema når man skal tolke tester. *Mitt forslag til hvordan bli en god test tolker er: Sandovals seks punkt ved Kritisk tenkning og Kaufmans tolv punkter for en intelligent tolkning.* Har man Sandovals og Kaufmans lister under huden når man tolker tester, mener jeg man har gode forutsetninger for å bli en god testtolker.

Jeg mener at det tar tid å sette seg inn i og gjennomføre punktene til Sandoval og Kaufman. Hadde det vært avsatt tid for et kompetanseløft i PPT kunne dette vært en ting å gripe fatt i. UFDs (Utdannings og forskningsdepartementet) strategi for kompetanseutvikling i

grunnopplæringen har ikke nevnt PPT i sine planer. (Tinnesand 2005). "Det bør kartlegges om PP-tjenesten har behov for kompetanse- og organisasjonsutvikling for å kunne fylle rollen som veilednings- og støtte apparat for skolene på en tilfredsstillende måte."(Tinnesand 2005:56). Jeg sier meg enig med at PPT burde vært nevnt i UFD's strategi planer om kunnskaps løfte i skolen, fordi PPT er et viktig støtteapparat for skolene. Det er viktig at dette støtteapparatet er kompetent.

### 3.3 Opplæring og faglig oppdatering

Opplæringen i utdannelsen er grunnlaget for kompetansen til de sakkyndige i PPT, men det er like viktig å holde seg faglig oppdatert. Det kan bli problematisk med kompetansen til de sakkyndige skriver NFPP i en artikkel som heter "Kvalitetsreformen og utdanning i PP-rådgivning ved de pedagogiske institusjoner".(NFPP 2002). I artikkelen er de skeptiske til om kvaliteten på studiet holdes oppe når studietiden reduseres. De sier de er av den oppfatning at et en toårig mastergrad alene ikke vil være tilstrekkelig for å kunne romme den teoretiske kunnskap, pensum og praksis som er nødvendig for å bli profesjonsutøver i pp-rådgivning. (NFPP 2002.) Jeg sier meg enig med NFPP sin bekymring om at redusert utdannelsesetid kan redusere kompetansen til PP-rådgivere. Desto viktigere er det da å holde seg faglig oppdatert for å få og ha den nødvendige kompetanse.

McCloskey og Athanasiou (2000) har undersøkt i sin survey fra Nord Colorado hvordan skolepsykologene har fått kjennskap til ikke tradisjonelle tester. Der svarer sakkyndige at de har fått kjennskap til ikke tradisjonelle tester fra følgende steder: 60 % innservice, 59% trening på jobben, 58% av andre i faget, 56% artikler i tidsskrift og 38% kurs. (McCloskey og Athanasiou 2000). Det var 46 % som først fikk høre om nasjonalt utarbeide retningslinjer gjennom denne surveyen. Det vil si at dette ikke er rette veien for å få frem informasjon til skolepsykologene.

Videre spurte McCloskey og Athanasiou (2000) hvorfor ikke-tradisjonelle tester var lite brukt? Svarene i surveyen var; manglende trening eller informasjon 36 %, hadde ikke tilgang på testmateriale 30 % og politiske barrierer 17 %. Surveyen som var utført ved universitetet i Nord Colorado viste at det ikke var motstand mot alternative tester som var hovedårsaken

til at de var lite brukt ved utredning av minoritets elever. McCloskey og Athanasiou spurte også om hvor trygge de var ved testing av minoritets elever? På en skala fra 1 til 5, der 1 var usikker og 5 var sikker, var svargjennomsnittet 2,6. Forskerne som utførte denne surveyen betraktet trivselsgjennomsnittet som lavt. 24 % svarte at de var sikre, 76 % mindre sikre og 16 % usikre ved testing av minoritets språklige elever. McCloskey og Athanasiou konkluderte med at det er viktig å fortsette med utviklingen av nye tester og gjøre alternative teknikker tilgjengelig for profesjonen.

Hvis vi gjennomførte en lignende survey i Norge som den McCloskey og Athanasiou gjennomførte i Nord Colorado hadde det vært lettere å vite hvor tiltak burde vært iverksatt.

Sakkyndige rettferdiggjør sin bruk av statiske tester med at de ikke har andre alternativ og at testene er nyttige i en viss grad, deler av testene kan brukes. (Hessels 2000).

Det finnes mange tilbud i markedet som gir faglig oppdatering. Thorsov kompetansesenter tilbyr årlige kurs i administrasjon, testing og tolkning av Leiter-R for ansatte i PPT. Det har vært stor pågang på disse Leiter-R kursene. (ref. tlf. samtale, 11.10.05. med Karen Svennhovd ved Thorsov kompetanse senter). Et av kursene som Thorsov kompetansesenter tilbyr heter: "Utredning av minoritetsspråklige med særskilte behov". I følge kursansvarlig Espen Egeberg, omhandler kurset blant annet: Bakgrunn for språkutvikling hos barn og forståelse av tospråklighet. Gjennomgang av Leiter-R og WISC-III samt dynamisk tilnærming. Problematisering av testbruken på minoritetsspråklige elever.

*Man bør selv gå aktivt inn for å holde seg oppdatert.*

### 3.4 Kulturreduerte tester

Kulturreduerte tester er tester som skal kunne brukes på elever fra forskjellige kulturer. Det vil si at testene ikke favoriserer noen kulturer fremfor andre kulturer.

Begrepene kultur og kulturelle forskjeller, forandrer innhold etter konteksten de omtales i. (Frisby 1998). Det er ikke enighet bland sosial- og adferdsforskere hva begrepet kultur skal inneholde. "At present, there is no consensus about a universal definition of culture." (Frisby 1998:69). Begrepene er derfor ofte brukt feil i litteraturen hevder Frisby (1998). Selv om det ikke er enighet om definisjonen på begrepet 'kultur' så er det mange som skriver om hvordan kulturen påvirker oss.

For å finne ut hvorfor kulturen spiller en viktig rolle ved testing går jeg til Vygotsky (1978).

Vygotskys begrep om høyere psykologiske prosesser har en sosial og historisk kvalitet.

"An interpersonal process is transformed into an intrapersonal one. Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first between people(interpsychological), and then inside the child (intrapsychological). This applies aqually to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions orginates as actual relations between human individuals. (Vygotsky 1978:57).

Kulturen går under huden på oss og blir en del av oss. Først er kulturen mellom mennesker, interpsykologisk, så i mennesket, intrapsykologisk. Det vil si at kulturen blir internalisering i oss mennesker. Høyere psykologiske prosesser er kulturelle og kognitive redskaper.(Bråten 1998). Identiteten er knyttet til kulturen skriver Robinson-Zañartu og Sloan Aganza (2000). Det selvbilde eleven har er farge av kulturens historie, er det en erobret kultur eller en erobrer kultur? Ser eleven seg selv som en taper eller vinner. Om eleven ser seg selv som en taper eller vinner virker inn på hvordan eleven lærer. Tror eleven at den er en taper er det ikke rart om eleven også blir en taper i skolen. (Robinson-Zañartu og Sloan Aganza 2000).

Vygotsky og Robinson-Zañartu og Sloan Aganza har vist eksempler på hvordan kulturen er en viktig del av mennesket. Med andre ord virker kulturen inn på de forutsetninger man har for å bli testet, fordi kulturen er en integrert del av eleven.

De non-verbale testene regnes ofte som kulturreduerte tester, men det er flere ulemper ved å bruke non-verbale tester på elever fra språklige minoriteter. En ulempe er at verbale tester predikerer verbale egenskaper bedre enn nonverbale tester. Vår evne til å bruke språket er viktig i skolen, jobb og den sosial kompetansen i majoritetskulturen.(Sandoval1998). En annen ulempe ved å bruke non-verbale tester er at de ikke er så kulturreduerte som man først trodde. ”Et økende omfang av forskning peker faktisk i den retning av at ikke-språklige tester i stedet kan være mer kulturelt ladet enn språklige tester”. (Hansen 2000:29). Jeg har derfor valgt å ikke vektlegge nonverbale tester som en mulighet for å redusere feilkonklusjoner ved testing av elever fra språklige minoriteter.

### 3.5 Benytte dynamiske tester

"Dynamic assessment is applicable as a window on learning, regardless of cultural community." (Ukrainetz m.fl. 2000:143). Det vil si at dynamiske tester ikke er kulturavhengige, fordi de tester hva barnet kan lære og ikke hva det allerede har lært. Tester man eleven for hva det allerede har lært kommer elevens kultur med i testresultatet fordi kulturen er en del av eleven.

"These dynamic testing conditions are believed to provide examinees with a testing environment that is more similar to the manner in which thinking and learning typically occur. Therefore, predictors derived in this manner should be more valid indices of important criterion measure." (Dillon 1998:8). Oppsummert hevder jeg at hvis elever fra språklige

minoriteter testes med dynamiske tester reduseres faren for feilkonklusjoner. Dette vil bli utdypet i kapittel 4 som beskriver dynamiske tester og forskning.

## 3.6 Avslutning

Jeg har sammenfattet i sju punkter hvordan man kan redusere feilkonklusjoner i testing av elever fra språklige minoriteter.

1. Bruk testene på en faglig forsvarlig måte. Avvik fra manualen mener jeg kan gjøres så lenge man er tydelig på hvordan testen er brukt og hva som er grunnlaget for denne bruken.
2. Bli en god tolker. Følg Sandovals 6 punkter og Kaufmans 12 punkter.
3. Vit hva du gjør og kjenn din egen kompetanse.
4. Kompetanse på tospråklighet og normalutvikling, husk at elever fra språklige minoriteter også er normale barn.
5. Jobbe i team og kontakte andre fagpersoner eller kompetansesentra.
6. Holde seg faglig oppdatert med bland annet kurs og artikler.
7. Benytt dynamiske tester.

Man er selv ansvarlig for å holde seg faglig oppdatert.

Husk at det er et barn og barnets utviklingsmuligheter du har med å gjøre. Den sakkyndige og barnet er i et asymmetrisk forhold..

Dynamiske tester var en av metodene jeg nevnte i dette kapitlet for å redusere faren for feilkonklusjoner ved testing av elever fra språklige minoriteter. Hva som kjennetegner dynamiske tester, hva de måler og hvorfor jeg anbefaler dem ved testing av elever fra språklige minoriteter blir behandlet i neste kapittel, kapittel 4.

## **4. KAPITTEL 4. PROBLEMSTILLING NR.3: KAN DYNAMISKE TESTER REDUSERE FOREKOMSTEN AV FEILKONKLUSJONER OG ØKE RELIABEL INFORMASJON I UTREDNINGEN AV ELEVER FRA SPRÅKLIGE MINORITETER?**

### **INNLEDNING**

For å belyse denne problemstillingen starter jeg med hva som er det felles teorigrunnlag i dynamiske tester og bakgrunnen for at dynamiske tester ble utviklet. Hva er en dynamisk test? Det finnes mange varianter av dynamiske tester som dekker ulike områder. Jeg har derfor valgt å vise hovedgrupper av dynamiske tester og testformat for å gi et innblikk i mangfoldet. Hovedgrupper av dynamiske tester er metakognitive intervensjoner, restrukturere testen, læring med testen, trene en enkelt kognitiv funksjon. Jeg redegjør for spesifikke tester innen hver av hovedgruppene. Dynamiske testers format viser til hvordan selve testen gjennomføres. Testformatet er enten som pretest-trening-posttest eller hint og ledetråder under selve testingen.

Når bør man ta i bruk dynamiske tester? I de fleste utredninger kan dynamiske tester tas i bruk, men de er spesifikt anbefalt å bruke ved testing av språklige minoriteter. Anbefalingen bygger på at dynamiske tester får tak i læringsprosesser og ikke kun hva eleven har lært til nå. Har dynamiske tester virkelig en bedre effekt enn statiske tester? Evalueringen av dynamiske tester viser til analyser av publiserte studier som viser at de har bedre effekt på flere områder. Elevens læringspotensial sier noe om hvilke intervensjoner som er mest hensiktsmessig for eleven, dette beskriver jeg som den praktiske verdien av dynamiske tester.

Jeg har vektlagt å bruke empiri fra artikler som viser til reduksjon av feilkonklusjoner og økning i reliabel informasjon ved bruk av dynamiske tester på minoritetsspråklige. Artikkelen viser at det er en positiv effekt av å bruke dynamiske tester, men hvorfor er da dynamiske tester lite brukt.

---

Pihl (2004) fokuserte på problemområder ved utredning av elever fra språklige minoriteter som i seg selv er viktig, men det er ikke forslag til tiltak og forbedringer hos henne. Dette kapitlet om dynamiske tester er mitt hovedforslag til forbedring av dagens praksis ved testing av språklige minoriteter.

I forbindelse med arbeidet til denne oppgaven har jeg blant annet vært i kontakt med PP-rådgivere, lærere og medstudenter, her har jeg erfart at det er viktig å skille mellom dynamiske tester og dynamisk testing. I disse samtalene om dynamiske tester har det vist seg at noen av dem jeg har møtt ikke har oppfattet skillet mellom dynamiske *tester* og dynamisk *testing*. Jeg har ofte hørt "vi driver jo med dynamisk testing", når jeg snakker om dynamiske tester. Viser til kapittel 1.2.2 "test og testing" som viser hva som skiller test og testing. I dette kapitlet er det *dynamiske tester* jeg skal ta for meg.

## 4.1 Teorigrunnlaget i dynamiske tester.

Utviklingen av dynamiske tester stammer fra Vygotsky og Feuerstein. (Lidz og Eliott 2000. Sternberg og Grigorenko 2002). Det er Vygotskys teorier som er selve teorigrunnlaget i dynamiske tester. Vygotskys teori "Sonen for den nærmeste utvikling" viser til hvilke læringspotensial barnet har. (Lidz 2000, Seng, Pou og Tan 2003, Tzurriel 2003). Det vil si at dynamiske tester måler barnets læringspotensial. Hvert barn har sitt særegne læringspotensial. Vygotsky(1978) mente det var et behov for både å teste hva barna alt kunne, men også hva de kan gjøre med litt hjelp. Hva barna kan gjøre med litt hjelp viser til hvilket læringspotensial det har. Dette viste Vygotsky (1978) i sin teori om sonen for den nærmeste utvikling. Vygotsky beskriver sonen for den nærmeste utvikling som følger: "It is the distance between the actual developmental level as determined by independent problem

solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers." (Vygotsky 1978:86).

Feuerstein er en av de første som utviklet metoder for å måle sonen for den nærmeste utvikling. (Lidz og Eliott 2000). De høyere psykologiske prosessene som Vygotsky beskriver (Kap.3 forklarer Vygotskys internalisering av begrep), blir til ved mediert læring. (Bråten 1996. Seng, Pou og Tan 2003). Det er disse høyere psykologiske prosessene som danner grunnlaget for Feuersteins mediert læring. (Seng, Pou og Tan 2003). Det skal sies at ikke alle dynamiske tester bruker mediert læring. Hessels var tydelig på at LEM ikke bruker mediering. (Kongressen 2004). Jeg kan si meg enig med Hessels at LEM ikke følger det som Feuerstein legger til grunn for hva som er mediert læring. Tar man derimot tak i det som Vygotsky beskriver i sin teori om hvordan de høyere psykologiske prosesser blir til, mener jeg at alle dynamiske tester har i seg mediering. Dette hevder jeg fordi mediering er noe som foregår mellom individ, og det er akkurat det som skjer når eleven får trening i testens oppgaver. Det foregår alltid en eller annen type for kommunikasjon mellom testadministrator og eleven i alle dynamiske tester. Hadde det ikke foregått kommunikasjon hadde det ikke vært mulig å finne ut hva eleven kan gjøre med litt hjelp.

## 4.2 Hvorfor kom dynamiske tester.

Jeg ser to grunner for at dynamiske tester ble utviklet, men begge grunnene hadde sin opprinnelse i at de eksisterende testene ikke strakk til. Den ene grunne baserer seg på antagelsen om at mentalt retarderte var mer opplæringsdyktige enn det vanlige tester viste. (Sternberg og Grigorenko 2002). Den andre grunnen var problemer ved testing av immigranter og barn med annen kulturbakgrunn. (Seng, Pou og Tan 2003). Med andre ord kom dynamiske tester som en motreaksjon til de eksisterende statiske testene.

---

Målpopulasjonen for dynamiske tester mente Budoff var ".....the target population is a broadly defined population of disadvantaged low-IQ students, wich includes underachievers, children with learning disabilities, immigrants, and others." (Sternberg og Grigorenko 2002:73). Det vil si at dynamiske tester kom som et alternativ til statiske teste for elever som fikk lav IQ skåre på statiske tester. Det ble trodd at disse underryterne hadde et større potensial en det IQ skåren indikerte. (Sternberg og Grigorenko 2002).

Feuerstein og kollegaer møtte på problemer der de ved hjelp av statiske tester ikke fikk tak i det de mente var barnets potensial. (Seng, Pou og Tan 2003). Etter Holocaust kom det mange Jødiske barn til Israel. Feuerstein og kollegaer jobbet med å integrere disse foreldreløse og sterkt traumatiserte barna i Israel. De Jødiske barna kom fra land som hadde annen kultur enn Israel. Feuerstein erfarte at de eksisterende testene kun så på hva barnet ikke greide å gjøre og ikke hva barna kunne lære. Disse testene var til liten hjelp på de utfordringene Feuerstein stod ovenfor. Feuerstein var kreativ og forut for sin tid da han utviklet en teori om mediert læring. (Seng, Pou og Tan 2003) Teorien bygger på at kognitive strukturer er modifiserbare. Denne tanken om at sinnet er modifiserbart er kjernen i Feuersteins Structural Cognitive modifiability, SCM. Modifiseringen kan gjøres med mediert læringserfaring som består av *intensjon - mening - overføring*. Intensjon står for lærerens beviste intensjon om det som gjøres. Meningen må medieres, fordi meningen er en hovedkomponent i motivasjon. Overføring viser til at kunnskapen kan brukes i andre settinger. (Seng, Pou og Tan 2003). Ved bruk av dynamiske tester får man en rikere og nyttigere historie med kognitive og konseptuelle komponenter. (Lidz 2000).

### 4.3 Hva er dynamiske tester?

Dynamiske tester kan ikke defineres på en enkel måte, fordi det er så mange typer av dynamiske tester. Dynamiske tester bør heller beskrives etter de felles egenskapene testene bygger på. (Lidz 2000).

Dynamiske tester måler læringspotensialet. Det vil si at barnet måles i forhold til seg selv ved dynamisk testing, mens det innen tradisjonell statisk testing måles i forhold til en normeringsgruppe.

Testsituasjonen i dynamiske tester er en gjensidig dialog mellom testadministrator og barnet. (Lidz 2000) Statisk testing har et fast mønster som skal følges, alle som testes skal behandles på samme måte.

Som sagt tidligere finnes det mange ulike typer dynamiske tester. Men de forskjellige dynamiske testene har en del felles kjennetegn. Jeg har valgt å bruke Tzurriel (2003) sine fire punkter for hovedforskjellene mellom dynamisk- og statisk testing. Disse fire punktene mener jeg gir en god oversikt over hva som kjennetegner dynamiske tester.

1. Dynamiske tester undersøker hva som er elevens læringspotensial, mens statisk testing finner frem til hva eleven kan i dag. Hva eleven kan i dag, sammenstilles med hva han/hun har lært til nå. Det vil si at dynamiske tester fokuserer på prosessen og statisk testing fokuserer på sluttproduktet. Dynamiske tester kan sies å ha fokus på både prosess og produkt, men at de har et dynamisk syn på `produktet` det vil si det er i bevegelse.

2. Testadministrator har forskjellige roller i de to testtradisjonene. Ved dynamiske tester blander testadministrator seg mer inn i testpersonens oppgaveløsningsprosess. Ser testadministrator at testpersonen har vansker med å løse oppgavene alene kan testadministrator komme med hint og stikkord som letter testpersonens oppgaveløsning.

---

Testpersonen får også øve seg på testoppgaver før selve testen begynner. Statisk testing har derimot helt klare linjer for hva testadministrator skal si og gjøre. Testadministrator skal gjøre eksakt det samme med hver enkelt testperson som testes statisk.

3. Oppgavetyperne i dynamiske tester bygger på hverandre. Å bygge på hverandre vil si, har du lært en lett oppgave kan du trekke med deg den kunnskapen til neste oppgave. Statisk testing bygger ikke på samme læringsprinsipp selv om oppgavene blir mer og mer kompliserte.

4. Det er det kvalitative aspektet som vektlegges ved tolkning av testresultatene i dynamiske tester. Hvordan lærte og hvilke nytte hadde testpersonen av læringen og hintene som ble gitt under oppgaveløsningen. Testpersonen måles i forhold til seg selv og sine egne prestasjoner. Statistiske testresultat baserer seg i hovedsak på kvantitative aspekt. Kvantitative aspekt som hvor mange poeng oppnådde testpersonen og hvordan skåret testpersonen i forhold til jevnaldrende.

Jeg presiserer at dynamiske tester har mer i seg enn statiske tester som ”diagnostiserer” også stopp. Dynamiske tester bygger på kartlegging – tiltak – tolkning – kartlegging. Noen ganger settes det i gang tiltak innen den ordinære rammen, mens i andre tilfeller må man aktivisere spesialundervisning etter § 5.1.

Dynamisk tilnærming til testing fremskaffer minst tre typer informasjon. (Lidz og Penà 1996). For det første får man innsikt i barnet som en lærende. Hvordan nærmer barnet seg oppgaven? Er det mønstre i feilene som gjøres? Hvordan er selvkorreksjonen? Den andre typen informasjon man får tak i er hvor modifiserbart barnet er. Hvor mye greier barnet å nyttegjøre seg av intervensjonene? Hvilke innsats yter barnet? Det tredje Lidz og Penà (1996) nevner er hvilke effekt har intervensjonen på barnet? Siden intervensjonene er en del

av testen får man sjekket hvordan intervensjonen fungere på barnet. Har barnet nytte av intervensjonen som brukes i testen?

Dynamiske tester vektlegger ”hvordan” og ikke bare ”hva” som er vanlig for de tradisjonelle testene. Dynamiske tester legger stor vekt på metakognitive prosesser som problemdefinisjon, avgjørelser, hvordan strategier anvendes, hvordan oppmerksomheten holdes, hvordan barnet reagerer på tilbakemeldinger. (Lidz og Penà 1996). Dynamiske tester gir mer informasjon om hva som kan være passende intervensjoner for eleven. Det vil si at dynamiske tester fremskaffer informasjon som de statiske testene ikke får tak i.

Kozulin og Garb (2002) har laget en tredeling av dynamiske testers praktiske verdi.

1. Skårer eleven gjennomsnittlig på pretesten, men viser et lite læringspotensial bør eleven få ekstra trening i informasjons - prosesseringsstrategier og lære dem å lære.
2. Skårer eleven gjennomsnittlig på pretesten og viser et høyt læringspotensial bør eleven få større utfordringer og muligheter til mer selvstendige oppgaver.
3. Skårer eleven lavt på pretesten og viser et lavt læringspotensial bør eleven få betydelig støtte for generell læring og problemløsnings ferdigheter.

Denne tredelingen mener jeg gir gode føringer for hvilke tiltak som vil være best for eleven.

Jeg har nå nevnt hva som kjennetegner dynamiske tester. Hvilke dynamiske tester som finnes og hva de er ment å skulle måle, tar jeg opp i neste avsnitt, hovedgrupper av dynamiske tester.

---

## 4.4 Hovedgrupper av dynamiske tester og testformat.

Det finnes mange metoder og prosedyrer for dynamiske tester. (Lidz og Elliott 2000). Det vil si at dynamiske tester ikke er en ensartet gruppe med tester. Videre sier Lidz og Elliott (2000) at det ikke er begrensninger på hvilke områder hvor dynamiske tester kan brukes. Selv om det ikke er begrensninger på hvilke områder man kan bruke dynamiske tester, så er det ikke utviklet tester for alle områder. Jeg synes det har vært vanskelig å få oversikt over hvilke områder dynamiske tester dekker i dag. I den litteraturen jeg har lest i forbindelse med denne oppgaven har de dynamiske testene som oftest vært fremstilt hver for seg. Lidz og Elliott (2000) har en oversiktlig tabell i sin bok på side 8 og 9, over tester som testutviklere selv har skrevet om i boken.

Jeg har funnet to steder hvor dynamiske tester har vært gruppert, men det har vært forskjellig grunnlag for grupperingene. Jeg har valgt å vise begge grupperingsmåter, fordi de hver på sin måte gir et tydeligere bilde av dynamiske testers omfang.

Jeg prøvde å sammenligne hvordan de to grupperingsmåtene hadde fordelt representantene for gruppene, men det gikk bare til en viss grad fordi de har forskjellig grunnlag for grupperingene.

Først viser jeg hvordan Lidz ([www.dynamicassessment.com](http://www.dynamicassessment.com)) har gruppert de dynamiske testene. Den andre grupperingsmåten er det Sternberg og Grigorenko (2002) som har beskrevet. Jeg tar så en gjennomgang av relevante dynamiske tester for hver av gruppene.

På nettsiden til Lidz ([www.dynamicassessment.com](http://www.dynamicassessment.com)) fant jeg en gruppering av dynamiske tester som var basert på intervensjonenes standardiseringsgrad og innholdet i testen. Denne grupperingen mener jeg viser hvilke prosedyrer som brukes innen dynamiske tester og hva som er målet for testen. Jeg ser denne grupperingen som en forklaring på hvilke prosedyrer dynamiske tester bruker.

Lidz har på sin nettside delt dynamiske tester inn i det hun kaller fire basismodeller.

1. *Åpen slutt* (open-ended). Her brukes fellesbetegnelsen problemløsning på oppgavenes art. Prinsipper og strategier for problemløsning og hvordan fremme selvstendig problemløsning. Feuerstein er brukt som eksempel her.

2. Her brukes fellesbetegnelsen problemløsning, men *intervensjonene er standardisert*. Disse testene forsøker å redusere kulturell bias. Budoff og Guthke er nevnt som representanter for denne gruppen.

3. *Graderte hint* gis ved feil svar på oppgavene inntil rett svar blir gitt. Alle går gjennom samme prosedyre, men antallet hint avhenger av hvor mange som må til for at oppgaven løses. Her er Campione og Brown nevnt som representanter.

4. *Curriculum-baserte* tester bruker innholde fra den lærenes undervisningsmateriale. (Her har jeg valgt å bruke det engelske begrepet `curriculum` da det norske begrepet `læreplan` ikke er en dekkende oversettelse for begrepets innhold.). Intervensjonene baserer seg på "beste praksis" for undervisning. Representanter er Lidz og Jepsen.

Sternberg og Grigorenko (2002) har delt dynamiske tester inn i fire hovedklasser, disse vil jeg nå gå grundigere inn på. Denne inndelingen av dynamiske tester i hovedklasser, ga meg en bedre forståelse av hvilke informasjon testene fremskaffer og hvilke område de er beregnet for. De fire hovedklassene er; *Metakognitive intervensjon, Læring med testen, Restrukturere testsituasjonen og Trene en enkelt kognitiv funksjon*. (Sternberg og Grigorenko 2002:47). Disse fire hovedklassene viser at dynamiske tester dekker et stort område, alt fra generelle kognitive funksjoner til spesifikke funksjoner. Jeg skal nå nevne representanter og deres tester for hver av de fire hovedklassene med dynamiske tester.

#### 4.4.1 Metakognitive intervensjoner

Sternberg og Grigorenko (2002) nevner bland annet Feuerstein som representant for gruppen *metakognitive intervensjoner*. Målet i denne gruppen av tester er å lære barnet begrep og prinsipper som kan generaliseres. Det vil si å utvikle uavhengig problemløsning hos barnet.

Feuerstein og kollegaer utviklet en test som heter "Learning Potential Assessment Device (LAPD)". (Sternberg og Grigorenko 2002:49). LAPD bygger på Feuersteins teorier om mediert læringserfaring (MLE) og Structural Cognitive modifiability (SCM). Det kan trekkes likhetstegn mellom MLE og Vygotskys internalisering av begrep. (Sternberg og Gigorenko 2002). LAPD har verbale og nonverbale oppgaver for spesifikke ferdigheter. Ferdighetene LAPD tar for seg er analogisk og numerisk resonering, kategorisering og hukommelsesstrategier. Hele testbatteriet består av 15 instrumenter. Mesteparten av testen baserer seg på standardiserte psykometriske instrumenter som blant annet Ravens matriser. Andre deler av testen er spesifikt utviklet for LAPD eller modifiseringer av eldre tester. Det er kun utvalgte deler av testbatteriet som brukes avhengig av elevens profil. Hvis hele testen skal brukes ville det tatt ca.10 timer. LAPD har vært kritisert blant annet i Lidz (1991) hvor Budoff kritiserer manglende standardiseringen i LAPD.

Feuerstein har brukt uttrykket "retarded performer rather than retarded person". (Stenberg og Gigorenko 2002:49). Det vil si at det er ikke personen som er retardert, men personens utførelse. Får barnet MLE kan det utvikle sine kognitive ferdigheter.

#### 4.4.2 Lære med testen

Denne gruppen med tester baserer seg på at barnet skal lære med testen.

Guthke og, Campino og Brown har utviklet tester som baserer seg på at barnet skal *lære med testen*. (Sternberg og Grigorenko 2002).

Guthke har utviklet en serie med tester med feste i pretest - trening - posttest formatet. Jeg har valgt ut en test som passer for oppgavens målgruppe. Testen er ikke spesifikk for minoritetsspråklige, men den blir sett på som et nyttig redskap ved bestemmelse om eleven skal ha spesialundervisning. (Lidz og Elliott 2000).

Leipzig Learning Test (LLT) er utviklet av tyskerne Guthke og Beckmann. Dette er en normert test for barn i alderen 6 - 8 år. LLT skal måle induktiv resonering og bruker standardiserte hint. Testen er ment å gi en indikasjon på hvordan barnet drar nytte av assistanse.

Hensikten med denne type test er å se hvordan barnet lærer, derfor er læring en del av testen. (Guthke og Beckmann 2000).

"Wundt's discipline Meumann (1922) was probably the first scientist who explicitly insisted that intelligence tests had to be administered repeatedly." (Guthke og Beckmann 2000:18).

Sternberg og Grigorenko (2002) har ikke nevnt Lidz sin curriculum-baserte test, ACFS: The application of cognitive functions scale. Dette er en test som ikke er beregnet for språklige minoriteter, men jeg mener at testen kan være aktuell å bruke på minoritetsspråklige elever for å finne hvilke intervensjoner som passer for eleven. Jeg har derfor tatt med Lidz sin test ACFS her.

Utgangspunktet for ACFS er spørsmålet "Hva forventes det at barn trenger for å lykkes i en typisk akademisk og sosial setting, en skole?". Denne kognitive funksjonsskalaen er utarbeidet av Lidz og er ment å gi svar på det ovenfor nevnte spørsmålet. (Lidz 2000).

ACFS er en curriculum (læreplan) - basert dynamisk test for barn i 3 – 5 års alder og retarderte elever som tilsvarer samme mentale alderen. Testen består av pretest – posttest. Testen er i hovedsak verbal og tar ca 2 timer å gjennomføre. Testen består av 6 deler. ACFS

---

tester evner og egenskaper eleven trenger for å ha utbytte av skolen, tester læreprosesser som er typisk for det amerikanske preschool programmet. (Lidz 2000).

Det er viktig å måle hvilket utbytte barnet har av skolen. For å finne hvilke utbytte barnet har av skolen, måles de prosesser barnet tar i bruk ved læring og hvilke de har vansker med å bruke. Samtidig finner man ut hvilke intervensjoner som passer for barnet.

ACFS er ikke testen man bør velge for høyt fungerende barn eller barn uten språk. Testen egner seg ved prehensivning og ved anbefalinger av tiltak i klasserommet. Spesielt for barn med forskjellige typer lærevansker. Testen egner seg ikke når barnet med stor sannsynlighet trenger spesialundervisning. (Lidz 2000). *Hvordan kan man på forhånd vite om eleven trenger enkle tiltak i klasserommet eller om det er behov for spesialundervisning?*

Testens teoretiske rammeverk bygger på Vygotskys teori om sonen for den nærmeste utvikling og Feuersteins medierte læring. ACFS måler ikke IQ, men barnets modifiserbarhet. Høy IQ er ikke ensbetydende med høy modifiserbarhet. Testen prøver å finne ut hvordan barnet anvender sine kognitive og metakognitive funksjoner når det løser oppgaver. (Lidz 2000)

Lidz sier at hun liker å bruke ACFS i slutten av en utredning for å få maksimalt utbytte. Det er sjelden hun bruker kun en dynamisk test. Hun bruker også statiske normerte tester som er et middel for å beregne barnets risiko og nåværende nivå for fungering, sonen for nåværende utvikling. (Lidz 2000).

#### **4.4.3 Restrukturere testsituasjonen**

Budoff og, Carlson og Wiedl's dynamiske tester bygger på å *restrukturere testsituasjonen*. Det vil si at de har tatt utgangspunkt i eksisterende tester og omarbeidet disse.

Budoff er den mest omtalte, derfor har jeg valgt hans tilnærming. Budoff utviklet ikke noen egne tester. Han brukte eksisterende statistiske tester og restrukturerte testsituasjonen slik at de målte læringspotensialet. (Sternberg og Grigorenko 2002). Budoff baserer sitt arbeid på antagelsen om at noen av de opplæringssvake barna kan lære mer enn det som kommer frem av de konvensjonelle testene. Hypotesen hans var, hvis de opplæringssvake barna fikk mulighet til å lære hvordan testenes oppgaver skulle løses ville noen av disse barna øke sine prestasjoner. Det var en forutsetning at opplæringen i oppgaveløsning skulle være organisert på en standardisert måte slik at testene opprettholdt sine psykometriske egenskaper.

Målpopulasjonen for Budoff var elever med lav IQ, underyttere, barn med lærevansker, immigranter og andre. (Stenberg og Grigoenko 2002).

Sternberg og Grigorenko (2002) skriver at Budoff kritiserte Feuersteins arbeider for manglende standardisering. Videre skriver de at det i Budoffs publiserte materiale er lite om treningsprosedyrene. Det vil si at testenes treningsprosedyre ikke var så standardiserte som Budoff ønsket at de skulle være. Videre er det ikke normer for læringspotensialet i Budoff's tilnærming. Budoff viser kun til høyt og lavt læringspotensial. (Sternberg og Grigerenko 2002).

Birkemo (1996) sier det samme som Budoff, at tradisjonelle tester ikke får tak i elevens mulighet til å lære. Birkemo ser en mulighet for PPT i Budoffs tilnærming til testing. "Den psykometriske tilnærmingen en finner hos Budoff er forholdsvis enkel, og skulle være hensiktsmessig å anvende i praktisk diagnostikk. Hvis en også kan utvikle normer for læringspotensialet hos ulike aldersgrupper, skulle en her kunne få et redskap som kan fungere som et supplement til tradisjonelle testmetoder som anvendes." (Birkemo 1996:26). I sitatet sier Birkemo at Budoff har en tilnærming som kunne vært anvendelig for PPTs utredningsarbeide, men at man også bør utvikle normer for læringspotensialet hos ulike aldersgrupper.

---

#### 4.4.4 Trening av en enkelt kognitiv funksjon

Den siste gruppen av dynamiske tester er *Trening av en enkelt kognitiv funksjon*, som har Swanson som sin representant. (Sternberg og Grigorenko 2002).

S-CPT: Swanson Cognitive Processing Test.

Det som gjør S-CPT dynamisk er at den forsøker å teste barnets grenser, "test-the-limits" (Swanson 2000:73) ved å gi standardiserte hint. Testen er standardisert og normert for U.S. befolkning, men den er ikke representativ for etniske minoritetsgrupper sier Swanson (2000). S-CPT sitt syn på lærevansker har sine teoretiske røtter i informasjons–prosesserings-teorien sier forskerne Sternberg og Grigorenko (2002). Testen baserer seg ellers på arbeidene til Feuerstein og Brown og Campione, både når det gjelder det metodiske og terminologien sier forskerne.

Testens målpopulasjon er personer i alderen 5 år til voksne som har lærevansker. (Swanson 2000). S-CPT er utviklet for å evaluere prosesserings - komponentenes evne under standardiserte dynamiske testbetingelser. Testen forsøker å fremskaffe et omtrentlig innhold av individets prosesseringspotensial. Det er dette som er det nye med denne testen, at den undersøker prosesseringspotensial i arbeidsminne. Det vil si at testen måler kapasiteten i individets arbeidsminne. Informasjonen om individets arbeidsminnekapasitet er med å gi en tydeligere identifikasjon av lærevanskene.

Prosesseringspotensial er definert som forskjellen mellom individets nåværende utførelse (utførelse som blir målt på statiske tester) og i hvilken grad individet greier å dra nytte av de hint som testeren gir.

S-CPT består av 11 deltester som kan tas hver for seg eller samlet. Tas hele testen er den estimerte tidsbruken 3 timer. Testens oppgaver består av forskjellige typer stimuli og prosesserings situasjoner. Det er oppgaver som er verbale, non verbale, visuelle og auditive presentert. Utførelsene av oppgavene kan enten være retrospektive eller `prospektive`. (Swanson 2000). Det vil si at oppgavene tar i bruk både fortid og fremtids perspektiv i prosesseringen. Resultatene fra S-CPT beregnes i syv forskjellige sammensatte skårer.

Forskerne hevder at barns lærevansker attribueres til svakheter i arbeidsminnet, men høy kapasitet i arbeidsminnet forbindes med akademisk suksess. De konkluderer med at S-CPT er en psykometrisk robust test.

Fordelen med denne testen er at den omfatter mange typer stimuli, det vil si at testen får undersøkt mange sider av elevens kognitive funksjoner. Testen er også normert og standardisert og tar for seg arbeidsminnets kapasitet, som regnes for å være det mest sentrale elementet for at en elev skal lære i skolesammenheng.

Hesselss var missfornøyd med at Sternberg ikke hadde nevnt LEM eller Lidz sine curriculum baserte tester. (Kongressen 2004). Jeg mener LEM tilhører metakognitive intervensjoner og Lidz sin curriculum baserte test tilhører hovedgruppen læring med testen.

Forskerne sier i at de har med et utvalg av representanter for hver av hovedgruppene. (2000:47) Jeg undrer meg over hvilket grunnlag de har brukt for å velge ut tester i sin bok. Lidz og Elliott's (2001) bok som tar for seg de mest brukte dynamiske tester har med et eget kapittel om LEM. Forøvrig er Lidz den forfatteren som har skrevet mest og om dynamiske tester.

#### **4.4.5 Testenes format**

Dynamiske tester kan i hovedsak gjennomføres på to forskjellige måter. Måten testen gjennomføres på kalles testens format. Sternberg og Grigorenko (2002) bruker metaforene `sandwich` og `kake` for å beskrive formatene som brukes ved dynamisk testing. Metaforen `Sandwich` står for pretest - trening - posttest. Differansen mellom resultatet på pre- og posttest viser hvilke nytte barnet har hatt av treningen. Har barnet hatt stor nytte av terningen betyr det at barnet skårer betydelig høyere på posttesten enn pretesten. Det vil si at barnet har et høyt lærings potensial. Metaforen `kake` henviser til hvor mange hint eller ledetråder barnet må ha for å løse oppgaven? Antall hint som må gis sier noe om hvor lett eller vanskelig barnet har for å nyttegjøre seg hjelpen. Et barn som må ha mange hint for å løse en

---

oppgave regnes for å ha et lavt læringspotensial. Begge formatene bygger på antagelsen om mengden med hjelp barnet trenger i testen viser om det har høyt eller lavt læringspotensial.

## 4.5 Når bør man bruke dynamisk testing?

Før jeg sier noe om når dynamiske tester bør tas i bruk vil jeg presisere at dynamiske tester sees på som en komplementering til de statiske testene og ikke et alternativ. (Seng, Pou og Tan 2003, Hansen 2000). Lidz (2000) presiserer at det ikke er spesifikke områder for anvendelse av dynamiske tester, men at de kan anvendes i de fleste områder. Det vil si at dynamiske tester kan brukes i de fleste områder og som komplementering av tradisjonelle tester.

Under hvilke betingelser er dynamisk testing *spesielt egnet*? Klinisk erfaring har vist at dynamisk testing er nyttig å bruke ved følgende seks punkter. (Seng, Pou og Tan 2003);

1. Lavt skåre på standardisert test samt mistanke om et ikke rettferdig bilde av barnets evner.
2. Mentalt retarderte fordi mange standardiserte tester ikke er beregnet for de som skårer under en viss poengsum.
3. Når det er diskrepans mellom barnets væremåte og utførsel.
4. Barn med lav sosialøkonomisk status, annen kulturbakgrunn eller minoritets språklig bakgrunn.
5. Barn som viser tegn til emosjonelle forstyrrelser, personlighetsforstyrrelser eller lærevansker.
6. Når man vill finne barnets læringspotensial og ikke kun hva barnet alt kan.

De ovenfor nevnte punktene er tilfeller der de tradisjonelle statistiske testene ofte ikke greier å avdekke informasjon eller fremskaffe reliabel informasjon. Det er ved slike tilfeller dynamiske tester kan komme inn å avdekke informasjon og fremskaffe reliabel informasjon. (Hessels 1996. Lidz & Penà 1996). Punkt 4. nevner bland annet fremmedspråklige barn som er denne oppgavens tema. Jeg forstår punkt 4, til Seng, Pou og Tan (2003) slik at de tradisjonelle testene ikke fremskaffer reliabel informasjon om fremmedspråklige barn. Jeg tolker dette som at dynamisk testing er et nyttig verktøy ved kartlegging av fremmedspråklige barn for å fremskaffe reliabel informasjon.

”A growing need for dynamic assessment of young children has emerged recently because decisions about treatment of specific cognitive functions should be made as early as possible.” (Tzuriel 2003). Det er viktig å sette i gang tiltak så tidlig så mulig slik at eleven ikke får huller i sin kunnskap slik som Sveitserosten (Kapittel 1.2.4) er et bilde på.

## 4.6 Evaluering av dynamiske tester.

Dynamiske tester er ikke en ensartet gruppe med tester. Dette har ført til inkonsistens i forskningsresultat. "It appears that some of the inconsistencies in research findings are due to the multifaceted nature of the goal of dynamic testing. (Sternberg og Grigorenko 2002:117).

Jeg mener at McCloskey og Athanasiou (2000) er svake i sin evaluerende beskrivelse av dynamiske tester. De bruker kun en bok av Lidz fra 1991 som referanse, det har skjedd mye på dette feltet siden 1991. Jeg mener de burde ha tatt seg bryet med å finne noe mer relevant og nyere forskningslitteratur. De hevder, ved å referere til Lidz (1991), at det ikke er vist at

---

dynamisk testing kan gi prediksjoner for akademisk suksess. Hessels (2000) har en kvalitativ analyse av læringspotensialet og IQ skårer. Denne analysen viser at dynamisk testing er bedre predikasjon på skolesuksess enn statisk testing. LEM kan predikere elevens akademiske suksess hevder Hessels (2000). LEM er med å finne elevens læringspotensial. De elevene som har høyt læringspotensial har også større sjanse for å lykkes i skolen. Målet med LEM er at den skal gi et bedre bilde av barnets generelle kognitive evner enn en tradisjonell IQ test.

Den beste evalueringen av dynamiske tester fant jeg i en artikkel av Swanson og Lussier (2001). Swanson og Lussier (2001) oppsummerer og foretar en meta-analyse av 30 publiserte studier om dynamiske tester. Meta-analyse er en statistisk teknikk som gjør det mulig å sammenligne funn fra de publiserte undersøkelsene. Swanson og Lussier (2001) bruker effekt størrelsen (Effect sizes (ES)) som beregning av forskjeller mellom posttestenes gjennomsnitt resultat på henholdsvis statiske og dynamiske tester, eller forskjellen mellom pre og posttest ved dynamisk testing. En lav ES er verdier opptil .20, moderat ES er verdier opptil .60 og høy ES er verdier opp til .80 og høyere. Meta-analysen har undersøkt 170 ES i de 30 publiserte studiene. Det vil si at Swanson og Lussier (2001) har foretatt meta-analysen av flere variabler i hver av studiene. Variabler som det ble utført meta-analyse på var alder, evnegrupper, evneområde, utvalgets størrelse, type test og testdesign.

I analyse av kronologisk alder viste det seg at de yngste, <10år hadde den høyeste ES .89, mens de i alderen 10 - 13 hadde .65 ES og den eldste gruppen, >13 år hadde lavest ES .24. I evnegruppene var det signifikante ES variasjoner. Underyttere hadde den høyeste ES 1.01 og de med lærevansker .56 ES, dette var en lavere ES enn gjennomsnitt til de hørselshemmede som hadde .83 ES og mentalt retarderte .17 ES. Når det gjaldt utvalgets størrelse var det de middels store utvalgene som hadde høyest ES. Studiene ble delt to områder, det var verbale og visuell-spatiale. Begge evneområdene hadde en moderat ES på henholdsvis, .69 ES for verbale og .73 ES for visuell-spatiale. Utvalgets størrelse viste at de minste og de største utvalgene hadde en lavere ES enn de mellomstore utvalgene. Det som gav størst predikasjon på ES var test prosedyre. Testene ble delt inn i følgende tre typer, scaffolding, coaching og strategitrening. Scaffolding fikk .77 ES og strategitrening fikk .66 ES det var høyere enn coaching som fikk .08 ES. Testdesign ble delt i pretest-trening-posttest og hint. Analysen

viste at det var pretest-trening-posttest designet som hadde den høyeste ES verdien .84 i motsetting til hint som hadde en ES verdi på .45.

Etter å ha lest denne artikkelen sitter jeg igjen med følgende oppsummering. Det vil si at unge underyttere i et middels stort utvalg får høyest ES ved bruk av dynamiske tester som bygger på scaffolding eller strategitrening med et pretest – trening - posttest design.

Swanson og Lussier (2001) undersøkte om det var retestings effekten i et pretest - posttest design som var utslagsgivende for en økning i posttesten, det viste seg at dette ikke var tilfelle. De dynamiske testene med et pretest - trening - posttest design hadde en signifikant større økning på posttesten enn et enkelt pretest - posttest design.

Artikkelen til Swanson og Lussier (2001) avslutter med følgende konklusjon: "In conclusion, the current meta-analysis shows that DA procedures improve testing performance over static condition." (Swanson og Lussier 2001:343). Med andre ord gjør man det bedre på dynamiske tester enn på statiske tester. Dette var ikke den samme konklusjonen som McCloskey og Athanasiou (2000) kom frem til. Jeg mener at Swanson og Lussier's analyse er mer troverdig enn McCloskey og Athanasiou, fordi de bygger på et mer omfattende datamateriale.

Det er synd at Swanson og Lussier (2001) ikke har sett på variabelen minoritetsspråklige og ES. Hvilke nytte har minoritetsspråklige av dynamiske tester?

Jeg kan heller ikke finne de publiserte undersøkelsene til Hessels, fra 1996 og 1997 i utvalget til Swanson og Lussier. Undersøkelsene til Hessels (1996, 1997) er ikke med i artikkelen til Swanson og Lussier (2001). Det ville vært svært interessant å vite hvilke ES Hessels undersøkelser hadde fått, fordi jeg mener at dette er en effektiv måte å teste minoritetsspråklige barn.

---

## 4.7 Empiri om reduksjon av feilkonklusjoner ved bruk av dynamiske tester på språklige minoriteter.

Undersøkelser som ser på forskjeller mellom elever med forskjellig bakgrunn, er ikke av nyere dato, dette har vært et aktuelt tema lenge. Den eldste undersøkelsen denne oppgaven bruker er fra 1979. Sewell (1979) undersøkte den prediktive effekten i intelligens tester og læringsoppgaver hos svarte og hvite. (Jeg har valgt å bruke Sewells betegnelse "black and white", selv om dette betraktes som diskriminerende i dag). Hvor sikre var testene til å forutsi svarte og hvite elevers fremtidig skoleprestasjoner? I undersøkelsen brukte de Stanford-Binet IQ test, den viste en forskjell i gjennomsnitt IQ hos svarte og hvite med henholdsvis 87,3 og 114,8.

Undersøkelsens dynamiske test, Paired-associate test, viser en minimal forskjell mellom svarte og hvite på 12,42 og 12,10. (Sewell 1979) Sewell konkluderer med at prosessorienterte tester (Dette var det tidligere navnet på dynamiske tester) viser lovende resultat for å forutsi fremtidige opplærings behov hos svarte elever som kommer fra lav sosioøkonomisk status.

Dillon (1998) har undersøkt hvordan Afrikanske Amerikanere og hvite college studenter i Illinois skåret på statiske og dynamiske tester for induktiv resonering. Undersøkelsen hadde 122 deltagere med høyere grads studenter, fordelingen var 77 hvite og 45 Afrikanske Amerikanske. Den dynamiske testen som ble brukt i undersøkelsen hadde formatet graderte hint. Undersøkelsens resultat viste at de hvite studentene skåret bedre enn de Afrikanske Amerikanske studentene på statiske tester. Når studentene ble testet med dynamiske tester skåret de Afrikanske Amerikanere og hvite studentene likt. Ingen forskjeller ble funnet mellom Afrikanske Amerikanere og de hvite studentene ved dynamisk testing.

Dillon konkluderer med følgende; "Clearly, equivalent inductive reasoning ability is identified across ethnic groups under subject-directed procedures. In short, the data reported in this experiment support the proposal that conventional testing, at least at times, fails to reveal the true abilities of African American examinees." (Dillon 1998:14).

Deltagerne i undersøkelsen til Kozulin og Garb(2002) var 23 akademisk risikostudenter med engelsk som andrespråk, de hadde ikke bestått engelsk eksamen på high school.

Dette var en curriculum basert dynamisk test som hadde et pretest - mediering - posttest format. Undersøkelsen viste at det var flere studenter som hadde like pretest skårer, men hadde forskjellige posttestskårer. Denne forskjellen bekrefter den praktiske verdien i dynamiske tester. Studenter har forskjellig læringspotensialet, men kan skåre likt på standard tester. "The results of the study indicate that dynamic procedure indeed provides information on students' learning potential over and beyond that which is available from the static test." (Kozulin og Garb 2002:112).

Hadde det vært en annen person som hadde utført medieringen kunne resultatene vært annerledes sier Kozulin og Garb(2002) i slutten av sin artikkel. Jeg er enig med dem i at det burde vært utført reliabilitets studier for å vite hvilke rolle dette spiller.

Learning potential test for ethnic minorities (LEM) er en test som er ment å måle de samme evnene hos barn fra forskjellige etniske grupper. (Læringspotensial test for etniske minoriteter).

Hessels har utført sine undersøkelser i Nederland. Han har undersøkt om LEM testen gir samme resultat uavhengig etniske forskjeller. (Hessels 1996). LEM testen har blitt utført på tre grupper med barn. De tre barnegruppene var tyrkiske, marokanske og nederlandske. Barna var i alderen 5 til 8 år. Totalt var det 800 barn med i undersøkelsen, 400 tyrkiske og marokanske, og 400 nederlandske. Hessels ville finne ut om de tre barnegruppene oppnådde sammenlignbare resultater på oppgavene i testen. Om testen måler de samme evnene hos de marokanske, tyrkiske og nederlandske barna.

Hovedkonklusjonen fra undersøkelsen viste at LEM måler de samme evnene i de tre barnegruppene. Det var noen små variasjoner på et par av del testene, men dette mener Hessels skyldes virkelige forskjeller og ikke målefeil. Disse forskjellene gikk i favør til de nederlandske barna, majoritetsbarna. Dette betyr at LEM ikke helt har greid å fjerne forskjellene i testskåren mellom majoritets- og minoritetsbarn, som vanligvis finnes i de tradisjonelle IQ testene. Gruppeforskjellen kan skyldes manglende utviklingsmuligheter for

---

kognitive strukturer og strategier som trengs for å utføre testen. Den manglende utviklingsmuligheten kan man ikke regne med at barnet vil overkomme i den korte tiden det tar å utføre testen. Uansett var gruppeforskjellene i LEM testen signifikant mindre enn på RAKIT, en standard intelligensstest som brukes i Nederland. Hessels konkluderer med at LEM er et brukbart verktøy for testing av generelle kognitive evner hos minoritets barn. (Hessels 1996).

LEM hadde høy korrelasjon med lærerens rangering av skole prestasjoner. Jo høyere LP-skåre, jo mer nytte hadde eleven av klasseroms undervisning.

LEM har vist å ha langtids prediksjon. Det var større korrelasjon mellom LP-skåren og type (secondary education) videre utdanning fire og syv år etter testing enn mellom IQ-skåren og videreutdanning. (Hessels 2000).

Evt. Flere case i Lidz bokside 124-126 eksempler case om LEM

Studier som ser på språkforskjeller og språkvansker.

Lidz og Penà (1996) beskriver i sin artikkel, "Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to latino american preschool children.", hvordan dynamiske tester kan brukes av logopeder for å skille mellom barn som har språkvansker og barn med annet språk. (Lidz og Penà 1996).

Den store mengden av ikke engelskspråklige minoriteter har skapt utfordringer. Teamene som diagnostiserer språkvansker hos minoritetsbarna er advart mot å fortsette overrepresentasjonen av minoriteter i spesialundervisning. Men problemet er at det ikke finnes noen valid tilnærming som skiller mellom barn som virkelig trenger spesialundervisning og barn som bare trenger mer trening i språket som brukes i skolen. (Lidz og Penà 1996).

*Hvordan skille mellom språkvansker som skyldes annet morsmål og språkvansker som skyldes kognitive skader?* Lidz og Penà (1996) viser til studiet utført av Penà, Quinn, & Iglesias i 1992. I dette studiet bruker de en dynamisk modell, test – trening - retest, for å diagnostisere språk- og talevansker hos minoritetsspråklige barn. De viser til en interessant

sammenligning av to barn fra dette studiet. Det er to Puerto Rikanske førskolebarn, Luis (3,11 år) og Vicky (4,1 år). Begge barna skårer under gjennomsnitt på pre-testen, word picture vocabulary test (EOWPVT). Målet med hver økt i testen er å øke barnets bevissthet ved å navngi bilder og objekter med spesielle navn og ikke bare beskrive dem ut fra kategori eller funksjon. Under normale forhold ville begge barna blitt henvist til spesialundervisning fordi de skåret under gjennomsnittet. Men siden testen bruker en dynamisk modell som består av test – trening - retest undersøkes barnas språkvansker grundigere. Det viser seg under medieringen / treningen at barna responderer veldig forskjellig. Det viser seg at Luis forstår testens mål og mening lett og kjapt. Vicky derimot trenger mye mer støtte av en voksen for å klare testen. Testresultatene viser at Luis skårer høyt på modifiserbarhet mens Vicky har vansker med å lære seg nye læringsstrategier. Når begge barna blir retestet med EOWPVT har Vicky en veldig liten forbedring mens Luis har økt sin skåre betraktelig. Ved å bruke denne dynamiske modellen viser det seg at det kun er det ene barnet som trenger spesialundervisning.

Testresultatene vises på to måter. Det ene er hvordan mediasjonen virket og det andre er modifiserbarhet skåren. Den førstnevnte består av 7 deler; oppmerksomhet, diskriminering, planlegging, selvregulering, overføringer, motivasjon og interaksjon. Den sistnevnte består av 3 deler; responsen, eksaminators innsats og overføringer. Testresultatene gir et bilde av hva som fungerer og ikke fungerer ved trening og læring hos hvert enkelt barn.

Artikkelen konkluderer med at barnets respons på en intervensjon-inni-testen synes å gi en mer meningsfull basis for diagnostisering enn en statisk normbasert test. (Lidz og Peña 1996).

Gutiérrez-Clellen og Peña (2001) undersøkte dynamiske testers egnethet for å skille mellom lærevansker som skyldes språkvansker eller språkforskjeller hos minoritetsspråklige elever. De skriver at tradisjonell språktesting fokuserer på barnets nåværende nivå og ser det i forhold til jevnaldrene. Minoritetsspråklige barn har en tendens til å få svake testresultat som man forklarer med språkforskjeller. "On the other hand, CLD children with language impairment may be at risk for under-referral if language difficulties are believed to be language differences." (Gutiérrez-Clellen og Peña 2001:1) CLD er forkortelse for 'children

---

form culturally and linguistically diverse`. Det er derfor viktig å finne en metode som kan undersøke om det er språkvansker eller språkforskjeller den minoritetsspråklige eleven sliter med.

Ikke alle dynamiske tester er like nyttige ved utredning av minoritetsspråklige barn. Gutiérrez-Clellen og Peña (2001) kom frem til at det var de dynamiske testene som hadde et pretest - trening - posttest format som var best egnet for å skille mellom språkvansker og språkforskjeller hos minoritetsspråklige barn.

"...,the test-teach-retest approach appears to be best suited for the identification of language disorders because the examiner makes assumptions about children`s ability based on previous knowledge and experiences with test situations and tasks." (Gutiérrez-Clellen og Peña 2001).

Det som skiller elever med språkvansker og språkforskjeller vises på testresultatene i form av at de med språkvansker viser liten eller ingen nytte av treningen, MLE: mediert lærings erfaring. "When an appropriate MLE is provided, children who are different, but typical, language learners are capable of demonstrating significant changes. On the other hand, children with language impairments may benefit from the MLE sessions but would demonstrate little or no quantitative change." (Gutiérrez-Clellen og Peña 2001:222).

Barnehagebarna som deltok i undersøkelsen til Ukrainetz m.fl. (2000) var stammemedlemmer fra Arapahoe og Shoshone. Disse barna hadde kun lært engelsk, men skulle nå lære stammenes eget språk av eldre stammemedlemmer.

Studiet undersøkte om man ved hjelp av dynamiske tester kunne finne et skille i læringspotensialet mellom barn som var sterke og svake til å lære et nytt språk. Resultatet viste at de barna som var best til å lære det nye språket også hadde høyest læringspotensialet. De barna som var svakest til å lære det nye språket hadde et lavt læringspotensial. Størrelsen på læringspotensialet viser til hvor stort utbytte barnet hadde av treningsperioden. Det vil si hvor stor forskjellen var mellom pretesten og posttesten.

"Dynamic assessment, still in the early stages of development in speech-language pathology, shows strong potential as a language evaluation approach for distinguishing differences from disorders in minority children." (Ukrainetz m.fl. 2000:152).

Viktig å skille mellom språkforskjeller og språkvansker.

Studiet til Peña (2000) undersøker skillet mellom språkvansker og språkforskjeller og predikere barnets endringspotensial ved bruk av dynamiske tester.

"These findings suggest that modifiability is a useful construct in providing a less biased assessment and that the descriptive results obtained from this observation can also be used to link assessment with intervention." (Peña 2000:87)

Deltagerne i de studiene Peña (2000) refererer til var Puerto Rikanske og Afrikanske Amerikanske barn i alderen 2 til 4 år bosatt i Amerika. Studier har vist at det er en sammenheng mellom endringspotensialet og språkevnen hos minoritetsspråklige barn. Barn som er svake til å lære språk har et lavt læringspotensial og barn som har en normal utvikling i tilegnelse av det nye språket har vist en høyere læringspotensial. Det vil si at minoritetsspråklige barn med et lavt læringspotensial kan ha språkvansker og barn med et høyt læringspotensial ikke har språkvansker, men språkforskjeller.

"Children from culturally and linguistically diverse backgrounds, who tend to score low on standardized language tests due to language difference, may be able to demonstrate higher performance under a mediated teaching condition. Performance in a mediated teaching condition may be a better indicator of ability for CLD children because of elimination of assessment bias." (Peña 2000:88).

Psykometrisk bias er begrenset ved dynamiske tester fordi man ikke kun vektlegger resultatet fra en testing, men pretest og posttest. (Peña 2000).

Hva skjer når det brukes tester som er utviklet i vesten på de som ikke bor i vestlig miljø?

Sternberg m.fl (2002) ville undersøke følgende spørsmål: Hva forteller de tradisjonelle evnetestene oss om evnene til svarte Afrikanere som ikke bor i et vestlig miljø? I undersøkelsen brukte de omarbeidete eksisterende tester.

Undersøkelsen hadde med 358 skolebarn i alderen 11 til 13 år fordelt på en eksperimentgruppe og en kontrollgruppe. Eksperimentgruppen fikk pretest - trening - posttest og kontrollgruppen fikk kun pretest og posttest. Pretestene hadde samme type oppgaver som ordinære intelligens tester. I eksperimentet fikk de tilgang til både statiske og dynamiske test skår på grunn av måten de administrerte testene på. Stenberg m.fl.(2002) kunne derfor sammenligne testresultatene fra de dynamiske og statiske testene og finne ut hvor effektive de var for å få tilgang til evnene hos skolebarna fra Tanzania.

"These results suggest that the posttests are better predictors of the reference ability measures than are the pretests." (Sternberg m.fl. 2002:9). Undersøkelsen konkluderte med at posttesten hadde høyere korrelasjon med referanse målingene som var hvordan de gjorde det på skolen enn med pretesten.

"In fact, we found that the children in our study seem to have had important intellectual abilities not measured by static intellectual tests." (Sternberg m.fl. 2002:10).

Undersøkelsen har god dokumentasjon av eksperimentet og de kommenterer begrensninger i undersøkelsens resultat. De stiller seg spørsmålet om en lignende undersøkelse ville fått samme resultat i et annet land. Oppfordrer andre til å undersøke muligheter og begrensninger ved dynamiske tester.

## 4.8 Hvorfor dynamiske tester er lite brukt.

Sternberg (2000) og Elliott (2000) har forklaringer på hvorfor dynamisk testing er lite bruk. Både Stenberg og Elliott har fire forklaringene for hvorfor det har vært lite bruk av dynamisk testig. Forklaringene til Sternberg og Elliott har felles trekk, men er også forskjellige. Jeg har derfor valgt å supplere Sternberg sine forklaringer med Elliotts forklaringer på hvorfor

dynamiske tester er lite brukt. De tre første grunnene for at dynamiske tester er lite brukt er relativt like hos Sternberg og Elliott. Den fjerde og siste grunnen for at dynamiske tester er lite brukt er det som skiller Sternberg og Elliott. Det vil si at det er fem forklaringer for at dynamiske tester er lite brukt.

1. Den første grunnen for at dynamiske tester er lite brukt er fordi psykologer *velger det kjente og trygge*, nemlig de statiske testene som de er vant med å bruke. Problemet som oppstår når man ikke prøver dette nye og ukjente er at man ikke får med seg modernisering og forbedringer som skjer i feltet. Sternberg (2000) hevder at psykologisk testing ikke har endret seg på lenge. "What other field of psychology (or anything else) is still using instrumentation that is little changed after the course of one hundred years?" (Sternberg 2000:xiv). Samfunnet har endret seg mye de siste hundre år så hvorfor har ikke psykologiske tester gjort det?

2. Den andre grunnen er at dynamiske tester er *kompliserte å bruke*. De fleste testene krever spesial opplæring av testadministrator. Det finnes få tilbud om trening i dynamisk testing. (Elliott 2000). I Norge er det i dag ikke tilgang til trening og opplæring i bruk av dynamiske tester.

Noen av de dynamiske testene er standardiserte andre ikke. Psykologer foretrekker tester som er lette å administrere hevder Sternberg (2000). Videre hevder Elliott (2000) at feltet er forvirrende fordi det finnes mange metoder og prosedyrer samt at det ikke er lett å lære en prosedyre gjennom et kort kurs.

3. Tredje grunnen er den *kompliserte psykometri*en i dynamiske tester. Det er mange måter man kan analysere dynamiske tester på, men det er ingen universell måte å gjøre det på. Noen av testene egner seg best for en klinisk analyse og ikke for psykometrisk analyse. (Sternberg 2000). Det vil si at dynamisk testing baserer seg på kvalitative data som ikke passer så godt inn den vestlige verdens tenkemåte. (Elliott 2000). Eksempel på klinisk analyse er tester som tolkes ipsativt, disse baserer seg på kvalitative data og er derfor ikke egnet for tradisjonell psykometrisk analyse. (Kap.3 Tolkning)

4. Den fjerde og siste grunnen Sternberg nevner, er redusert validitet i dynamiske tester i forhold til de vanlige testene. Det er ikke sikkerhet i hvor mye man kan vektlegge dynamiske tester i forhold til de tradisjonelle testene, spesielt med tanke på de tre først

---

nevnte punktene. "although there are demonstrations of incremental validity of the dynamic tests over conventional tests, it is not clear how much one can count on the tests adding to a conventional testing battery, especially when one considers the factors mentioned above." (Sternberg 2000:xv).

5. Den femte grunnen som Elliott nevner for at dynamiske tester er lite brukt er fordi dynamisk tilnærming *krever tid*, noe en utreder ikke har for mye av. Elliott nevner også at noen dynamiske tester kan gjennomføres på kort tid.

*Oppsummeres de fem punktene kan man si at dynamisk testing er lite brukt fordi man liker å bruke det kjente, som ikke krever for mye av oss og som vi kan kvantifisere på kort tid.*

Lidz (2000) viser til sin survey fra 1992 hvor hun undersøkte hvilke kjennskap skolepsykologer i U.S. har til dynamiske tester og om de brukte testene. Bakgrunnen for surveyen var å finne hvordan det hadde gått med tidligere kursdeltagere som hadde lært om dynamiske tester. De fleste svarte at de kjente til dynamiske tester, men at det ikke hadde vært nok tid eller støtte for å sette i gang med dynamiske tester. (Lidz 2000). Jeg mener undersøkelsen til Lidz bekrefter noen av forklaringene til Elliott og Sternberge på hvorfor dynamiske tester er lite brukt.

## 4.9 Avslutning

Dynamiske tester bygger på et felles teorigrunnlag, det er Vygorskys teori "Sonen for den nærmeste utvikling". Testene måler læringspotensialet, i dette ligger det at de kognitive strukturene er modifiserbare. Testadministrator har en aktiv rolle i dialog med eleven. Testens oppgaver bygger på hverandre og det kvalitative aspektet vektlegges i tolkningen av

testresultatene. Dynamiske tester har en praktisk verdi som gir indikasjoner på hvilke intervensjoner som eleven har best nytte av.

Dynamiske tester ble utviklet fordi eksisterende tester ikke strakk til for blant annet mentalt retarderte og immigranter, barn med en annen kulturbakgrunn. Det blir presisert fra flere forfattere at dynamiske tester er et tillegg til dagens statiske tester og ikke en erstatning.

Det finnes mange forskjellige typer dynamiske tester og de kan deles inn i hovedgrupper. Jeg har vist to grupperingsmåter. Den ene grupperer etter hvilke prosedyrer som brukes i testen og hva som er dens mål. Den andre grupperer etter hvilke informasjon testen kan fremskaffe og hvilke område den er beregnet for.

Det er foretatt en meta-analyse av 30 studier som viser at dynamiske tester med et pretest – trening – posttest design hadde en signifikant økning på posttesten i forhold til et enkelt pretest -posttest design. Unge underytere i et middels stort utvalg får høyest ES ved bruk av dynamiske tester som bygger på scaffolding eller strategitrening med et pretest - trening - posttest design.

Dynamiske tester har vist seg å være en god metode for å skille mellom språkforskjeller og språkvansker. Høy LP, men lav IQ kan indikere at elevens vansker er språkforskjeller. Er det en lav LP og en lav IQ kan det være en indikasjon på at eleven har en språkvanske.

Dynamiske tester har eksistert i mange år, men er fremdeles lite brukt. Grunner for at dynamiske tester er lite bruk er at man velger det kjente og trygge som ikke er tidskrevende. Det hevdes at testene er komplisert å bruke og de fleste testene krever opplæring av administrator. Psykometrien i testene er komplisert og det er usikkerhet rundt validiteten.

*Betyr det at hensynet til testadministrator er viktigere enn hensynet til barnet?*

## 5. KAPITTEL 5. KONKLUSJON

Bruk av statistiske tester er dagens praksis ved testing av elever fra språklige minoriteter. Det viste seg at WISC-III ble brukt i de sakkyndige rapportene jeg analyserte. Statistiske tester viser hva eleven har lært til nå innenfor bestemte kunnskaps og ferdighetsområder, men ikke hva som er elevens læringspotensial. Språkforskjeller kan føre til feilkonklusjoner, dette gjelder spesielt ved bruk av statistiske tester, fordi det er språket og ikke ferdigheten og kunnskapen som måles. Flere studier har vist at elever fra språklige minoriteter kan skåre lavt på en statisk test, men høyt på en dynamisk test. Det betyr at statistiske tester kan undervurdere elevens læringspotensial. Undervurderes elevens læringspotensial er det fare for at eleven blir hindret i å utvikle seg i forhold til sitt potensial, noe som igjen kan føre til sosial reproduksjon. Det er viktig med et perspektivskifte fra statisk testing til dynamisk kartlegging, for å hindre at elever stemples og stagnerer.

I min analyse av sakkyndige vurderinger var elevens minoritetsspråklige bakgrunn nevnt og tatt med i tolkningen i 5 av 6 tilfeller. I den ene sakkyndige vurderingen der eleven minoritetsspråklige bakgrunn ikke var tatt med i tolkningen, ble eleven tildelt spesialundervisning. Dette mener jeg er en feilkonklusjon fordi det er fare for at det var elevens språkkompetanse som ble testet og ikke kunnskap og ferdigheter som var testingens hensikt.

Elevenes tospråklige utvikling var ikke nevnt i noen av de sakkyndige vurderingene. Dette ser jeg som bekymringsfullt da tospråklig utvikling virker inn på elevens forståelse av språket og testspråket.

Ingen av de sakkyndige vurderingene hadde brukt dynamiske tester som hevdes å kunne gi indikasjoner på om eleven har en språkvanske eller om det er snakk om språkforskjeller.

For å redusere feilkonklusjoner ved testing av elever fra språklige minoriteter, må noe gjøres. Løsningen er ikke å fjerne testene og de sakkyndige vurderingene. Det vil bare gjøre vondt verre. De fleste sakkyndige vurderingene jeg har analysert, er etter min mening gode selv om ikke alle holdt mål. Dagens praksis ved bruk av statiske tester på elever fra språklige minoriteter er ikke helt tilfredsstillende, fordi språkforskjeller kan dekke over språkvansker og motsatt. Leiter-R som er en nonverbal test spesielt utviklet for denne elevgruppen, ville hjelpe noe, men den måler ikke språket som er en viktig faktor hos elever fra språklige minoriteter.

Jeg har følgende forslag til forbedringer av dagens praksis.

- Bevisstgjøring om testteori, tolkningsmetoder og tospråklig utvikling er av stor betydning for hvordan tester brukes og hvordan de blir tolket. Dette kan være med å redusere faren for feilkonklusjoner. En start kan være å følge Kaufmans 12 punkt og Sandovals 6 punkt som gir retningslinjer for hvordan bli en god testtolker.
- Dynamiske tester ville være et nyttig supplement til statiske tester i utredningen av elever fra språklige minoriteter, fordi dynamiske tester har vist seg å være et godt verktøy ved testing av elever fra språklige minoriteter. Det har vist seg at dynamiske tester kan gi gode indikasjoner på om eleven sliter med språkvansker eller språkforskjeller.
- Utfordringer videre blir å finne gode måter for bevisstgjøring om testteori, tolkningsmetoder og tospråklig utvikling, samt videreformidle nytten ved bruk av dynamiske tester slik at de sakkyndige ser det som et pluss og ikke mer arbeid i utredningen av elever fra språklige minoriteter.

## 6. Kildeliste

Kildeliste Badderley, A. D. 2002. The psychology of memory. Badderley, A. D., Kopelman, M. D. og Wilson, B. A. (red) *The handbook of memory disorders*. West Sussex: John Wiley & Sons Ltd.

Bråten, I. 1998. Om Vygotskys liv og lære. Bråten, I. (red.). *Vygotsky i pedagogikken..* Gjøvik: Cappelen akademiske forlag as.

Birkemo, A. 1996. Dynamisk testing som metodisk tilnærming i pedagogisk-psykologisk utredningsarbeid. *Skolepsykologi* (1996) nr.3 s.21-28.

Cohen, R. J. & Swerlik, M. E. 1999. *Psychological testing and assessment*. An introduction to tests and measurement. California: Mayfield publishing company.

Comas-Diaz, L. og Ramos Grenier, J. 1998. Migration and Acculturation. Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

Cozby, P. C. 1997. *Methods in Behavioral Research*. London: Mayfield Publishing Company.

Cummins, J. 1987. Language development, academic learning, and empowering minority students. Fradd S. H. og Tikunoff W. J.(Red.):*Bilingual education and bilingual special education*. Boston: A college- Hill Publication.

Dillon, R. F. 1998. Dynamic Assessment of inductive Reasoning. A Test of Induced Versus Imposed Mediation. Illinois: Soutern Illinois University: U.S. Department of education *Reports - Reaearch* (143).

Elliott, J. G. 2000. Dynamic assesement in educational contexts: Purpose and promise. Lidz, C. S. & Elliott, J. G.(red). *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Flanagan, D. P. og Kaufman A. S. 2004. *Essentials of WISC-IV Assessment*. New Jersey: John Wiley& sons, Inc.

Flanagan, D. P., McGrew, K. S. og Ortiz, S. O. 1999. *The Wechslers intelligence Scales. Gf-Gc Theory*. Boston: Allyn and Bacon.

Fossestøl, B og Sandvei, M. 1989. *Tanums store rettskrivnings ordbok*. Oslo: Kunnskapsforlaget.

Frisby, C. L.1998. Culture and cultural differences. Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

---

Geisinger, K. F. 1998. Psychometric Issues in Test Interpretation. Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

Gottfredson, L. S. 1997. Mainstream Science on Intelligence: An Editorial With 52 Signatories, History, and Bibliography. Detterman, D. K. (Ed). *Intelligence. A multidisciplinary journal*. Greenwich: Ablex Publishing Corporation.

Gould, S. J. 1997. *The mismeasure of man*. London: Penguin books.

Gundersen, I. 2001. *Elevforutsetninger og et sosialt konstruert språk*. Upublisert semesteroppgave 1.avd. Cand.ed. Pedagogisk forsknings institutt. Universitetet i Oslo.

Guthke, J. og Beckmann, J. F. 2000. The learning test concept and application in practice. Lidz, C. S., & Elliott, J. G. (Red.): *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Gutiérrez-Clellen, V. F. og Peña, E. 2001. Dynamic Assessment of Diverse Children: A Tutorial. Amerika: *Language, Speech, and Hearing Services in Schools*, vol 32, s.212 -224.

Hansen, A. 2000. Hva innebærer dynamisk testing? *Skolepsykologi* nr 1, 11-45.

Hessels, M. G. P. 1996. Ethnic Differences in Learning Potential Test Scores: Research into Item and Test Bias in the Learning Potential Test for Ethnic Minorities. *Journal of Cognitive Edukation*,(1996) 5, 133-153.

Hessels, M. G. P. 1997. Low IQ but high learning potential: why Zeyneb and Moussa do not belong in special education. *Educational and Child Psychology*. (1997), Vol 14 (4). 121-136.

Hessels, M. G. P. 2000. The learning potential test for ethnic minorities (LEM): A assessment of children in kindergarden and the first years of primary school. Lidz, C. S., & Elliott, J. G. (Red.): *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Hessels, M. G. P. & Hessels-Schlatter, C. 2002. Learning potential in immigrant children in two countries: the Swiss-French version of the learning potential test for ethnic minorities. Van der Aalsvoort, G. M. , Resing, W. C. M., Ruijsenaars, A. J. J. M. (Red): *Learning potential assessment and cognitive training*. Amsterdam: Jai. Elsevier Science Ltd.

Jackson, C. 2002. *Psykologisk testing*. København: Psykologisk Forlag A/S.

Herrnstein, R. J. og Murray, C. 1994. *The Bell Curve*. New York: The free press.

Håndbok for PP-tjenesten. 2001. Læringscenteret i Oslo: GAN grafiske as.

ICD-10: <http://www.who.int/classifications/icd/en/>

Jensen, A. J. 1980. *Bias in mental testing*. London: Methuen.

Kaufman A. S. 1994. *Intelligent Testing with the WISC-III*. New York: John Wiley & sons, Inc.

Kjeldstadli, K. 2000. *Fortida er ikke hva den en gang var. En innføring i historiefaget*. Oslo: universitetsforlaget.

Kongressen 2004. *Dynamiske arbeidsmåter i opplæring og kartlegging*. Holmenkollen Park Hotell Rica, 14.-16.oktober 2004. Arrangør: Bredtvet kompetansesenter i samarbeid med Nordisk spesialpedagogisk nettverk.

Kozulin, A. og Garb, E. 2002. Dynamic Assessment of EFL Text Comprehension. *School Psychology International* (2002, vol. 23(1): 112-127.

Kunnskapsforlagets blå ordbøker 1996. *Engelsk - norsk blå ordbok*. Oslo: Aschaug - Gyldendal.

Leiter-R . *Leiter internasjonal performance scale-revised*. 1997. Illinois: Stolting Co.

LEM (Leertest voor Etnische Minderheden) 1991. Lisse: Swets & Zeitlinger B.V.

Lewis, J. 1989. Innovative approaches in assessment. Samuda, R. J. & Kong, S. L. (Red.): *Assessment and placement of minority students*. Toronto: C. J. Hogrefe.

Lidz ([www.dynamicassessment.com](http://www.dynamicassessment.com)) (15.09.2005).

Lidz, C. S. 2003. *Early childhood assessment*. New Jersey: John Wiley & sons, Inc.

Lidz, C. S. & Elliott, J. G. 2000. *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Lidz, C. S. 2000. The application of kognitive functions scale (ACFS): an exsample og curriculumbased dynamic assessment. Lidz, C. S. & Elliott, J. G. (Red.):*Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Lidz, C. S. & Penà, E. D. 1996. Dynamic assessment: The model, its Relevans as a Nonbiased Approach, and its Application to Latino American Presvhool Children. Amerika: *Language, Speech, and Hearing Services in Schools*, 27, 367-72.

Lidz, C. S. 1991. *Practitioner`s guide to dynamic assessment*. New York: The Guilford press.

Læringscenteret 2001. *Håndbok for PP-tjenesten*. Oslo: GAN Grafiske AS

McCloskey, D. & Athanasiou, M. S. 2000. Assessment and intervensjon practices with second-language learners among school psychologists. *Psychology in schools*, 37(3), 209-225.

Nygård, R. 1993. *Aktør eller brikke? -om menneskers selyforståelse*. Oslo: Ad Notam Gyldendal.

---

NFPP (Norsk forening for pedagogisk psykologi). Medlemsblad. September 2002.

NFPP (Norsk forening for pedagogisk psykologi). Etske prinsipper for fagmedlemmer i NFPP og profesjonsområdet pedagogisk-psykologisk rådgivning. Oslo

Peñã, E. 2000. Measurement of Modifiability in Children From Culturally and Linguistically Diverse Backgrounds. *Communication Disorders Quarterly* vol. 21, no. 2. s.87-97.

Peña, E. D. og Gillam, R. B. 2000. Dynamic assessment of children referred for speech and language evaluations. Lidz, C. S. & Elliott, J. G. (Red.). *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Pihl, J. 2005. *Etnisk mangfold i skolen. Det sakkyndige blikket*. Oslo: Universitetsforlaget AS.

Pihl, J. 2002, a. Viten, makt og konstruksjon av mental retardasjon i den flerkulturelle skolen anno 2000. *Norsk tidsskrift for migrasjonsforskning*, 2, 89-108.

Pihl, J. 2002, b. Sakkyndig konstruksjon av avvik og stigma i en interetnisk relasjon. *Sosiologisk tidsskrift*, 10, 91-111.

Pihl, J. 2002, c. Sakkyndige vurdering av minoritets elever. *Tidsskrift for norsk psykologforening*, 39, 604-612.

Robinson-Zanartu, C. A. og Sloan Aganza, J. 2000. Dynamic assessment and sociocultural context: Assessing the whole child. Lidz, C. S. & Elliott, J. G. (Red.). *Dynamic assessment: Prevailing models and applications*. New York: Elsevier Science Inc.

Roid, G. H. Og Miller, L. J. 1997. *Leiter internasjonal performance scale-revised*. Examiners manual. Illinois: Stoltz Co.

Rogoff, B. 1990. *Apprenticeship in thinking. Cognitive development in social context*. New York: Oxford university press.

Rundskriv F-008-03, Tilskudd til opplæring for elever fra språklige minoriteter i grunnskolen, kap 221 post 65. Utdannings og forskningsdepartementet.

Rundskriv F-008-04, Endringer i opplæringsloven og friskoleloven. Utdannings og forskningsdepartementet.

Rust, J. and Golombok, S. 1999. *Modern psychometrics. The Science of Psychological Assessment*. London: Routledge.

Ruud, M. 2004. Ville ha barna ut av norsk som andrespråk. *Utdanning*,13, 22-24.

Samuda, R. J. & Kong, S. L. 1989. *Assessment and placement of minority students*. Toronto: C. J. Hogrefe.

Sandoval, J. 1998. Critical Thinking in Test Interpretation. Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

Sandoval, J. 1998. Testing in a changing world. I Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (Red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

Sandoval, J. og Duràn, R. P. 1998. Language. I Sandoval, J., Frisby, C. L., Geisinger, K. F., Dowd Scheuneman, J. og Ramos Grenier, J. (Red). *Test interpretation and diversity. Achieving Equity in assessment*. Washington, DC: American Psychological Association.

Sattler, J. M. 2001. *Assessment of Children. Cognitive Applications*. Fourth edition. USA: Maple-Vail Book Manufacturing Group.

Schultz, D. P. & Schultz, S. E. 1996. *A history of modern psychology*. Florida: Harcourt Brace & Company.

Seng, Seok-Hoon, A., Pou, Kwee-Hoon, L. og Tan, Oon-Seng. 2003. *Mediated learning Experience with Children*. Singapore: McGraw-Hill Education.

Skutnabb-Kangas, T. 1981. *Bilingualism or not, The education of minorities*. Cornwall: Robert Hartnoll Ltd.

Sternberg, R. J. m.fl. 2002. Assessing intellectual potential in rural Tanzanian school children. *ScienceDirect*. New York: Elsevier Science Inc.

Sternberg, R. J. & Grigorenko, E. L. 2002. *Dynamic Testing. The Nature and Measurement of Learning Potential*. Cambridge: Cambridge university press.

Sternberg, R. J. 2000. Prologue. Dynamic assessment: prevailing models and applications. Lidz, S. C. og Elliott, G. J. (red). *Dynamic assesement: Prevailing models and applications*. New York: Elsevier Science Inc.

Sewell, T. E. 1979. Intelligence and learning tasks as predictors of scholastic achievement in black and white first-grade children. *The journal of school psychology, Inc.*, 17(4),325-332.

Swanson, H. L. og Lussier, C. M. 2001. A Selektive Synthesis of the Experimental Literature on Dynamic Assessment. *Review of Educational Research*. Vol. 71, no. 2, s.321-363.

Swanson, H. L. 2000. Swanson-Cognitive Processering Test: Review and applications. Lidz, S. C. og Elliott, G. J. (red). *Dynamic assesement: Prevailing models and applications*. New York: Elsevier Science Inc.

Tinnesand, T. 2005. Hvor blir det av PP-tjenesten i kunnskapsløftet? *Skole psykologi*. nr.2, 2005, s. 55 - 58.

---

Tzuriel, D. 2003. Foundations of Dynamic Assessment of Young Children. I Seng, Seok-Hoon, A., Pou, Kwee-Hoon, L. og Tan, Oon-Seng. 2003. *Mediated learning Experience with Children*. Singapore: McGraw-Hill Education.

Ukrainetz, T. A., Harpell, S., Walsh, C. og Coyl, C. 2000. A Preliminary Investigation of Dynamic Assessment With Native American Kindergartners. *Language, Speech, and Hearing Services in Schools*. vol. 31, s142-154.

Undervisnings- og forskningsdepartementet. Lov om grunnskolen og den videregående opplæringen (Opplæringsloven).

Utdannings- og forskningsdepartementet. 2004. *Veiledning. Spesialundervisning i grunnskolen og i videregående opplæring*. Oslo: GAN Grafiske AS

Vygotsky, L. S. 1978. *Mind in Society*. The Development of Higher Psychological Processes. London: Harvard University Press.

WISC-III. Wechsler Intelligence Scale for Children – Third Edition. Sverige: Psykologiförlaget AB.

[www.utdanningsetaten.oslo.kommune.no/utdanningsstatistikk/](http://www.utdanningsetaten.oslo.kommune.no/utdanningsstatistikk/) (Antall minoritets-språklige elever i Oslo). (01.10.2004).

[www.wis.no/gsi](http://www.wis.no/gsi) GSI'04 - Grunnskolens informasjonssystem (01.10.2004)

[www.swap.ac.uk/learning/glossary.asp?initial](http://www.swap.ac.uk/learning/glossary.asp?initial) = Ipsative (03.11.2005)

<http://morris.wharton.upenn.edu/forecast/data/definitions/ipsative%20score.html> = Ipsative  
(03.11.2005)

Øzerk, K. 1992. *Om tospråklig utvikling*. Haslum: Oris forlag.

Øzerk, K. 1998. Vygotsky i tospråklighetsforskning. Bråten, I. (red.): *Vygotsky i pedagogikken*. Gjøvik: Cappelen akademiske forlag as.

Øzerk, K. 2003. *Sampedagogikk*. Vallset: Opplandske Bokforlag.