



**Spesialiserte metoder for avdekking av malingering i  
nevropsykologiske undersøkelser; en vurdering og  
sammenligning av symptomvaliditetstester og  
spørreskjema**

*Malingering detection methods in neuropsychological  
examinations; an evaluation and comparison of symptom  
validity tests and inventories*

*Anita Puhr*

*Veileder  
Professor Knut Waterloo*

*Hovedoppgave  
Emnekode PSYKPRO 501  
Profesjonsstudiet i psykologi  
Institutt for Psykologi  
Universitetet i Tromsø  
Høsten 2007*

## Forord

Temaet for denne oppgaven ble valgt av forfatteren. Forfatteren har selv utarbeidet problemstillinger som undersøkes i denne oppgaven, samt skaffet til veie litteraturen som er blitt benyttet.

Kasusbeskrivelsen i denne oppgaven er hentet fra det pågående multisenterprosjektet ”Symptomvaliditet og differensialdiagnostikk i nevropsykologiske undersøkelser” ved Sunnaas sykehus HF, Psykiatrien i Vestfold HF, Rikshospitalet-Radiumhospitalet HF og Ullevål Universitetssykehus HF, hvor forfatteren er prosjektmedarbeider.

Jeg vil takke Knut Waterloo for god veiledning i forbindelse med denne oppgaven. Jeg vil også takke Vegard Klausen for god hjelp og støtte underveis. Sist vil jeg rette en takk til Anne-Kristine Schanke og Stein Andersson for nyttige råd.

Anita Puhr  
Stud. Psychol.

## Sammendrag

Tittel: Spesialiserte metoder for avdekking av malingering i nevropsykologiske undersøkelser; En vurdering og sammenligning av symptomvaliditetstester og spørreskjema.

Forfatter: Anita Puhr, stud.psychol.

Veileder: Knut Waterloo, dr.psychol. e-post: waterloo@psyk.uit.no

Fag/grad: Profesjonsstudiet i psykologi, Universitetet i Tromsø

Semester: Høst, 2007

”Malingering” refererer til et klinisk fenomen hvor en person, motivert av ytre insentiver, bevisst og intensjonelt produserer falske eller grovt overdrevne fysiske eller psykiske symptomer. Dette fenomenet vil kunne forekomme i nevropsykologiske undersøkelser hvor personen kan oppnå økonomiske, trygdemessige eller andre fordeler ved å få påvist en sentralnervøs skade eller dysfunksjon. De siste tiårene har det vokst frem flere ulike metoder som er spesifikt utviklet for å avdekke malingering i nevropsykologiske undersøkelser. I denne oppgaven belyses to av disse metodene; symptomvaliditetstester og spørreskjema. Etter en vurdering og sammenligning av metodene, finner forfatteren at det pr i dag er ingen avdekkingsinstrumenter som alene kan avgjøre hvorvidt malingering har forekommet, blant annet på grunn av metodologiske begrensninger ved studier som søker etter å fastsette metodenes effektivitet. I tillegg til dette, tyder eksisterende forskningsdata på at symptomvaliditetstester er de mest validerte, og derfor de mest pålitelige av de to metodene. Til tross for dette, anses metodene for å være et viktig bidrag i vurderinger om symptomvaliditet og malingering, hvor en kombinasjon av multiple avdekkingsmetoder og informasjon fra multimodale kilder vil kunne øke påliteligheten av utrednings- og klassifikasjonsnivået ved nevropsykologiske undersøkelser.

# INNHALDSFORTEGNELSE

<b>1. INNLEDNING</b> .....	<b>1</b>
1.1 PROBLEMSTILLINGER .....	2
1.2 BEGREPSAVKLARING .....	3
<b>2. SPESIFIKKE AVDEKKINGSMETODER</b> .....	<b>5</b>
2.1 BAKGRUNNEN FOR UTVIKLING AV SPESIFIKKE AVDEKKINGSMETODER.....	5
2.2 SYMPTOMVALIDITETSTESTER .....	8
2.2.1 Tradisjonelle FCT-tester: Tallgjenkjenningprosedyrer.....	11
Portland Digit Recognition Test .....	11
Victoria Symptom Validity Test.....	13
2.2.2 Alternative FCT-tester .....	14
Test of Memory Malingering.....	14
21-Item Test .....	16
2.2.3 Oppsummering.....	16
2.3 SPØRRESKJEMA .....	17
2.3.1 Om spørreskjema .....	18
2.3.2 Utvalgte spørreskjema .....	20
Structured Inventory of Malingered Symptomatology.....	20
Neuropsychological Symptom Inventory.....	22
Memory Complaints Inventory.....	23
Malingering Detection Scale .....	24
2.3.3 Oppsummering.....	25
<b>3. DISKUSJON</b> .....	<b>26</b>
3.1. METODOLOGISKE HENSYN VED STUDIER AV SPESIFIKKE METODER .....	27
3.1.1 Oversette basisrater .....	27
3.1.2 Simulatorer versus simulanter .....	29
3.1.3 Kjente kliniske grupper .....	30
3.2 SAMMENLIGNING AV SYMPTOMVALIDITETSTESTER OG SPØRRESKJEMA .....	32
3.2.1 Klassifisering; Etablering av motivasjon.....	32
3.2.2 Simulering versus symptomaggraving .....	34
3.2.3 Vurdering om påliteligheten av symptomvaliditetstester og spørreskjema.....	34
3.3 HOLDNINGER OG PRAKSIS VED MISTANKE OM MALINGERING .....	39
3.3.1 Mistanke og moralisering .....	39
3.3.2 Valg av metoder .....	41
<b>4. KONKLUSJON</b> .....	<b>42</b>
<b>REFERANSER</b> .....	<b>44</b>

## 1. INNLEDNING

*K. J., en 51 år gammel mann, presenterte flere klager om alvorlige kognitive funksjonstap, som han hevdet var forårsaket av arbeidsforholdene (kontakt med industrielle kjemikalier) i fabrikken han hadde jobbet i over en 15-årsperiode. Han beskrev hvordan livet hans nå var blitt "et ork". Han fortalte hvordan han ble utmattet av hverdagslige aktiviteter han hadde behersket hele livet (for eksempel å tilbringe tid med arbeidskameratene sine), og hvordan han hadde blitt avhengig av kona si med hensyn til å huske enkle ting (for eksempel små beskjeder, klokkeslett og gjøremål). Han fortalte også at som resultat av sin vanskelige livssituasjon, hadde han i økende grad isolert seg og blitt nedtrykt og engstelig. På bakgrunn av erstatningssaken K. J. førte mot arbeidsgiverne sine, samt spørsmål om videre utbetalinger av uføretrygd, ble han henvist til utredning med spørsmål om løsemiddelbetinget encefalopati. Den billeddiagnostiske undersøkelsen som bestod av både magnettomografi (MR) og computertomografi (CT) ga ingen holdepunkter for hjerneskade. Deretter gjennomgikk han en nevropsykologisk undersøkelse som bestod av tester som stilte krav til blant annet finmotorikk, oppmerksomhet, psykomotorisk tempo, hukommelse, visuokonstruktive evner og eksekutive funksjoner. I tillegg ble det administrert symptomsjekklistere for psykopatologi, hukommelsesvansker og trøttbarhet. Ved gjennomgang av resultatene fra undersøkelsen, fant nevropsykologen det svært vanskelig å fortolke funnene. Flere av funnene syntes uforklarlige.*

Til tross for veksten av moderne teknologi med billeddiagnostikk som har vist seg å være uvurderlige i undersøkelse av hjerneskade, og som i dag har tatt over mye av det diagnostiske arbeidet, er det stadig tilfeller hvor selv slike teknikker ikke strekker til. Eksempelvis er det noen typer nevropatologi hvor de nevrologiske symptomene er så subtile at de ikke fanges opp av billeddiagnostiske metoder. Dette gjelder ved visse former for hjerneskader og nevropsykiatriske tilstander, som for eksempel løsemiddelskader, lette hodeskader og begynnende demenstilstander. Billeddiagnostiske metoder vil ved slike omstendigheter være supplerende, men ikke diagnostiske.

Et illustrerende eksempel er pasienter som er henvist til nevropsykologisk utredning med spørsmål om lette hodeskader, hvor pasientene hevder å oppleve svikt i kognitiv funksjon, men hvor det ikke eksisterer noen form for objektive dokumentasjon av nevropatologi (for eksempel funn fra billeddiagnostiske metoder). I disse tilfellene vil den nevropsykologiske undersøkelsen kunne bli den primære kilden til dokumentasjon, ettersom sensitive nevropsykologiske instrumenter har påvist kognitive svekkelser også ved mindre "usynlige" cerebrale skader (Pankratz & Binder, 1997).

I en nevropsykologisk undersøkelse samles informasjon inn i form av testprestasjoner og symptomrapportering. Begge disse formene for informasjon kan påvirkes eller manipuleres av en pasient som ønsker å fremstå som funksjonssvekket. Dersom det eksisterer muligheter for å oppnå ytre gevinster ved å få påvist en sentralnervøs skade eller dysfunksjon (for eksempel utbetaling av trygd eller erstatning, unngåelse av ansvar, plikter eller straff), bør nevropsykologen, i følge Binder (1993), i slike tilfeller vurdere muligheten for at pasienten er motivert til å aggravere, eller simulere, symptomer. Dersom dette er tilfellet, vil hele undersøkelsen trekkes i tvil ettersom funnene anses som invalide. Det er dette kliniske fenomenet som begrepet *malingering* refererer til. I slike sammenhenger kan økonomiske insentiver spille en betydelig rolle i vurderingen av funksjonssvikt/uførhet etter lette holdeskader, noe også meta-analytiske studier demonstrerer (Binder & Rohling, 1996). ”Veileder i klinisk nevropsykologi” utarbeidet av Norsk Nevropsykologisk Forening (NNF, 2002) anbefaler i slike situasjoner bruk av spesifikke metoder eller prosedyrer for å belyse troverdigheten av undersøkelsesresultatene. Spesifikke avdekkingsmetoder for malingering inkluderer symptomvaliditetstester og spørreskjema.

## 1.1 Problemstillinger

Mens malingering som fenomen er relativt enkelt å definere, er det desto mer komplisert å undersøke og avdekke, noe som gjenspeiles i nevropsykologisk litteratur om metoder utviklet for å avdekke fenomenet klinisk. Det finnes flere metoder som er spesifikt utviklet for å bistå i kliniske vurderinger av malingering. Disse benytter ulike fremgangsmåter og avdekkingsstrategier. Hovedproblemstillingen i denne oppgaven omhandler behovet for, og nytteverdien av, spesifikke avdekkingsmetoder og ulike kliniske og metodologiske utfordringer knyttet til disse. Mer spesifikt er jeg interessert i følgende:

- Hvorfor er det nødvendig med spesifikke avdekkingsmetoder?
- Hvordan skiller de spesifikke avdekkingsmetodene symptomvaliditetstester og spørreskjema seg fra hverandre med hensyn til avdekkingsstrategier?
- Hvilke instrumenter brukes i dag?
- Hvilke svakheter er knyttet til studier som søker etter å etablere de to spesifikke metodenes kliniske nytteverdi? Vurdering av ulike design.

- Hvor pålitelige er disse spesifikke metodene og hvilken klinisk nytteverdi har de? Er den ene mer pålitelig enn den andre? Fordeler og ulemper ved symptomvaliditetstester og spørreskjema.
- Hva bør den enkelte kliniker være oppmerksom på ved bruken av disse spesifikke metodene? Etske utfordringer knyttet til bruken av metodene og klassifisering i praksis.

## 1.2 Begrepsavklaring

Nevropsykologiske undersøkelser dreier seg om å måle kognitive funksjoner og å sannsynliggjøre hvorvidt kognitive funksjoner er endret som følge av skade eller sykdom som påvirker hjernefunksjon, samt vurdere omfang og type kognitiv reduksjon (for eksempel hukommelsesvansker, oppmerksomhetsvansker og perseptuelle vansker). De viktigste og mest sentrale behovene innen nevropsykologisk praksis og forskning, er å sikre validitet ved bruk av tester og metoder. Med *validitet* menes at en metode måler det man har til hensikt å måle eller undersøke (Liebert & Liebert, 1995). I forlengelsen av dette refererer *symptomvaliditet* til hvorvidt man måler de (kognitive) symptomene man har til hensikt å måle. *Validering* av en metode refererer til forskning som har som formål å sikre metodens validitet. *Validitetsmål* beskriver metoder som benyttes klinisk for å avdekke hvorvidt funn fra nevropsykologiske tester er *invalid*, det vil si at de ikke gjenspeiler den faktiske tilstanden til en gitt kognitiv funksjon, men andre utenforstående faktorer.

Det kliniske fenomenet *malingering* kan være en av faktorene som gjør funn fra nevropsykologiske undersøkelser invalide. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association [APA], 1994) definerer malingering som ”en intensjonell produksjon av falske eller grovt overdrevne fysiske eller psykiske symptomer, motivert av ytre insentiver”. I Den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer (ICD-10; Verdens Helseorganisasjon [WHO], 2004) er malingeringbegrepet ikke brukt. Her benyttes koden Z76.5 ”Simulant”, hvor simulering defineres som ”bevisst produksjon eller etterligning av somatiske eller psykiske symptomer eller dysfunksjoner som er motivert av ytre belastning eller impulser”. Diagnostisering av malingering vil bli grundigere gjennomgått senere i oppgaven.

Når man skal beskrive et sensitivt område som malingering er, er det viktig å nyansere mellom begrepene som benyttes. For eksempel er det en avgjørende differensialdiagnostisk forskjell mellom en person som simulerer symptomer av indre årsaker, eller årsaker som er

ukjent for klinikeren, og en person som er motivert til å simulere av ytre forhold. I denne oppgaven vil jeg benytte *simulering* om selve *atferden* til person hvor motiverende forhold er ukjente, men hvor man har holdepunkter for at personen aggraverer eller simulerer symptomer. Selv om aggravering og simulering av symptomer er relaterte begreper, er de ikke synonyme. Likevel vil jeg av hensyn flyt i teksten, benytte simulering om begge, og skille på dem der det er nødvendig.

*Suboptimal prestering* benyttes her om tilfeller hvor en pasient av ukjente årsaker intensjonelt produserer prestasjoner på nevropsykologiske tester som er svakere enn pasientens reelle kognitive fungeringsnivå.

I tilfeller hvor ytre motiver for simulering eller aggravering er etablert, vil jeg benytte DSM-kategorien *malingering*. Engelskspråklig malingeringlitteratur benytter termen *malingerer* om en *person* som simulerer eller aggraverer symptomer for å oppnå ytre gevinster. I stedet for malingerer, vil jeg i denne oppgaven benytte ICD-kategorien *simulant*, ettersom denne termen er norsk og derfor egner seg bedre i en norskspråklig oppgave.

Om friske deltakere som av eksperimentelle hensyn er bedt om å simulere kognitive vansker (analoge studier), vil jeg av mangel på en mer egnet norsk term benytte det engelske ordet *simulator*.

Fenomenet malingering har nå blitt introdusert, og de viktigste begrepene er avklart. Jeg vil komme tilbake til diagnostisering forbindet med malingering, men først vil jeg ta for meg metoder som er spesifikt utviklet for å avdekke malingering.



## 2. SPESIFIKKE AVDEKKINGSMETODER

*K. J.s resultater på de nevropsykologiske standardtestene var vanskelige å fortolke. Et flertall av resultatene var så svake at de tilsvarte lavområdet; finmotorikk, oppmerksomhet, psykomotorisk tempo, mental effektivitet, delt oppmerksomhet, visuell hukommelse for kompleks figur og mental fleksibilitet og impulshemming. På Stroop-testen var resultatene på de enkle benevningsprøvene like svake som benevningsprøven som inkluderer interferens. Vanligvis vil man kunne tenke seg at sistnevnte hadde vært noe svakere enn de andre, ettersom denne er en mer kognitivt krevende oppgave. Ettersom resultatene på disse prøvene var så svake at de falt utenfor normeringen, var det imidlertid ikke mulig å skille på dem. Til tross for svært svekkete prestasjoner på en oppmerksomhetsprøve, tilsvarte prestasjoner på en prøve for verbal hukommelse (California Verbal Learning Test) midtre normalområdet. Likevel bestod ikke K. J. en enkel gjenkjenningssprøve som tilhører denne hukommelsesprøven. På en enkel test for visuell hukommelse, produserer K. J. kun 8 av de 15 stimuliene, hvilket anses som dårligere enn ved de mest alvorlige hjerneskader. Var disse resultatene tilstrekkelige for å trekke en slutning om at K. J. simulerte de kognitive vanskene han hadde presentert?*

### 2.1 Bakgrunnen for utvikling av spesifikke avdekkingsmetoder

Til tross for nytteverdien ved standard nevropsykologisk undersøkelse ved generell utredning av kognitiv svikt, vil det faktum at man er avhengig av pasientens samarbeid og motivasjon gjøre undersøkelsen særlig sårbar for malingering (Heubrock & Petermann, 1998). I verste fall kan dette resultere i feildiagnostisering. På grunn av fare for feildiagnostisering, stilles det generelt et krav til nevropsykologiske metoders evner til korrekt klassifisering. Denne evnen bestemmes ut fra hensyn om *sensitivitet* og *spesifisitet* (Etcoff & Kampfer, 1996). I en kontekst hvor for eksempel en avdekkingsmetode for malingering undersøkes, vil sensitivitet bety hvor stor prosentandel av simulantene blir korrekt klassifisert på bakgrunn av resultater på testen som indikerer malingering. En test som ikke er sensitiv vil kunne resultere i at personer som aggraverer eller fabrikerer symptomer ikke vil oppdages (falske negative). Spesifisitet angir hvor stor prosentandel av deltakerne som *ikke* simulerer eller aggraverer symptomer, og som får resultater som indikerer fravær av simulering og malingering (Bianchini, Mathias & Greve, 2001a). Mangler undersøkelsen tilstrekkelig spesifisitet, står man i fare for å feilklassifisere pasienter med reelle hjerneskader som simulanter (falske positive).

I en nevropsykologisk kontekst kan malingering forekomme på to måter: 1) viljestyrt presentasjon av overdrevne symptomer som er vanskelige å definere og måle nøyaktig (for eksempel ved selvrapportering), og 2) viljestyrt produksjon av svake prestasjoner på psykologiske eller nevropsykologiske tester (Iverson & Binder, 2000). Sårbarheten for simulering ved nevropsykologisk testing demonstreres i studier hvor man har undersøkt klinikers evne til å avdekke suboptimal prestering/simulering kun ut fra resultater på standard nevropsykologisk testing. Resultatene fra disse studiene indikerer generelt en lav grad av presisjon i diagnosefastsettelse (Franzen, Iverson & MacCracken, 1990). For eksempel, i en studie hvor 113 tenåringsdeltakere ble bedt om å simulere hjerneskade, var det kun 20 % av klinikerne som mente at resultatene kunne tilskrives noe annet enn kortikal skade (Faust, Hart, Guilmette & Arkes, 1988). Haines og Norris (1995) påpekte en annen svakhet ved nevropsykologiske tester og batterier; de kan være oversensitive, samt at de har lav grad av spesifisitet, noe som kompliserer vurderingsprosessen og øker risikoen for falske positive. van Gorp m.fl. (1999) hevder at selv statistisk analyse av testprofiler, som for eksempel diskriminant funksjonsanalyse, har vist seg å være lite reliable.

Etttersom analyse av nevropsykologiske testresultater alene har vist varierende grader av reliabilitet med hensyn til å avdekke suboptimal prestering/simulering, har man forsøkt å undersøke testprestasjoner ut fra idé om at inkonsistente forhold i en testprotokoll kan avdekke suboptimal prestering (Etcoff & Kampfner, 1996; Bianchini m.fl., 2001a; Iverson & Binder, 2000). Eksempler kan være inkonsistens mellom resultater på tester som har som hensikt å måle lignende underliggende evner, inkonsistens mellom forventede resultater (vurdert ut fra type skade, skadeomfang, demografiske data) og faktiske resultater, eller inkonsistens mellom resultater fra testing og retesting (Iverson & Binder, 2000). Stroop-testen kan være et illustrerende eksempel på en slik "symptomvaliditetsindikator" (Bianchini, m.fl., 2001a). Her vil man ikke forvente bedre prestasjoner på deltesten som inkluderer interferens sammenlignet med prestasjoner på ord- og fargebenevning, ettersom førstnevnte anses å være en mer kognitivt krevende oppgave.

Nevropsykologisk kunnskap om hukommelse og typiske prestasjoner til reelle hjerneskadepasienter på hukommelsesprøver, har vist seg særlig nyttig i avdekking av malingering (Nies & Sweet, 1994). Det forventes for eksempel at gjenkjenning er enklere enn gjenhenting, at primacy-effekter er intakte, at responslatens følger visse mønstre, og at prestasjoner på hukommelsesprøver ikke kan være bedre enn prestasjoner på prøver som stiller krav til oppmerksomhet og konsentrasjon (Haines & Norris, 1995; Nies & Sweet, 1994). California Verbal Learning Test (CVLT) er et eksempel på en hukommelsesprøve som

benytter seg av denne kunnskapen. CVLT består av 5 innlæringsprøver, en interferensprøve, en gjenhentingsprøve etter et kort tidsopphold etter interferensprøven, en gjenhentingsprøve etter et 20 minutters tidsopphold, og til slutt en gjenkjennelsesprøve som skal reflektere motivasjon for å simulere. Gjenkjennelsesprøven har gjennom flere studier demonstrert gode sensitivitets- og spesifisitetsrater (Etcoff & Kampfer, 1996).

Til tross for deres nytteverdi som validitetsmål, anses likevel usannsynlige eller inkonsistente testprofiler (og symptomvaliditetsindikatorer) for å være utilstrekkelige med hensyn til å avdekke personens intensjon om å bedra, og ytterligere dokumentasjon er nødvendig i en kontekst hvor økonomiske eller strafferettslige forhold er til stede (Hiscock & Hiscock, 1989). Ut fra dette behovet, er det utviklet andre prestasjonsbaserte avdekkingsmetoder, ofte kalt *symptomvaliditetstester* (SVT), som refererer til tester som er spesielt konstruert for å oppdage urimelige prestasjoner og for å dra slutninger om intensjon (Iverson & Binder, 2000). Disse metodene er også blitt kalt motivasjonstester ("effort tests"), ettersom de er utviklet for å avdekke hvorvidt en person er motivert til å prestere maksimalt. Symptomvaliditetstester har i de to siste tiårene vært hovedfokus i nevropsykologisk litteratur om malingering, og testantallet har økt i takt med dette.

En alternativ metode til prestasjonsbaserte avdekkingsmetoder, er bruken av *spørreskjema*. Til tross for at denne tankegangen er relativ ny innenfor forskning på malingering, er det i dag flere spørreskjemaer, som i likhet med symptomvaliditetstester, er spesielt utviklet for å oppdage simulering. Denne tilnærmingen er utviklet ut fra praktiske hensyn, som for eksempel behovet for multiple kilder i vurderinger om symptomvaliditet og malingering, og begrenset mengde tid pr undersøkelse.

Mens spørreskjema og symptomvaliditetstester er spesifikt utviklet for å bidra til å avdekke malingering hos pasienter, er rasjonale bak de to metodene noe ulikt. Når det gjelder symptomvaliditetstester, antas det at personen som simulerer vil prestere under evne, det vil si at han/hun aktivt *unngår* å respondere (korrekt) (Bianchini m.fl., 2001a). I kontrast, antas det at respondering på spørreskjema vil gjenspeile at personen *avgir* falske svar som han/hun tror gir inntrykk av kognitiv svikt, det vil si at han/hun aggraverer eller simulerer symptomer (Iverson, 2006). I de følgende avsnitt gjennomgås historikken og avdekkingsstrategier til henholdsvis symptomvaliditetstester og spørreskjema.

## 2.2 Symptomvaliditetstester

*På symptomvaliditetstesten TOMM klarte K. J. 23, 24 og 26 riktige av 50 gjenkjenningsprøver på henholdsvis trial 1, trial 2 og retention trial. Med en cutoff-skåre på 45 riktige, har K. J. ikke bestått denne symptomvaliditetstesten. 23 og 24 riktige svar er videre under det som anses som sjansenivå, og 26 er rett over grensen.*

Symptomvaliditetstester ble opprinnelig utviklet for vurderinger om tvilsomme påstander om sensoriske svikttegn, som for eksempel hørselsvansker, blindhet, fargeblindhet, tunnelsyn, tåkesyn og parestesi (Pankratz, 1979, i Rogers 1997). Pankratz (1983) anses for å være opphavsmannen til begrepet *symptomvaliditetstester*, og for å ha videreført metoden til nevropsykologiske vurderinger om simulering av hukommelsesproblemer (Heubrock & Petermann, 1998). I Pankratz-prosedyren velger klinikerer ut en målstimulus som presenteres for pasienten, etterfulgt av en oppgave hvor pasienten blir bedt om å diskriminere mellom to stimuli (målstimulus og en ny stimulus) og tilkjenne hvilken av de to han eller hun gjenkjenner. Denne teknikken har fått tilnavnet ”forced-choice testing” (FCT).

FCT består av to elementer; 1) En spesifikk evne undersøkes ved hjelp av et høyt antall testledd presentert i et tovalgsformat og 2) en persons prestasjoner sammenholdes med estimer om sannsynligheten for å lykkes ved vilkårlig prestering på testen (for eksempel ved ren gjetting) (Pankratz & Binder, 1997). I et tovalgsformat vil denne sannsynligheten være 50 %. I følge FCT-teori vil prestasjoner som er svakere enn vilkårlig prestering reflektere *intensjonalitet*, ved at det krever at pasienten faktisk har persipert og gjenkjent målstimulus, men er motivert for å prestere svakt og derfor velger aktivt å *unngå* å respondere korrekt (Bianchini m.fl., 2001a; Pankratz & Binder, 1997; Faust & Ackley, 1998). Det er dette som ligger i begrepet under sjansenivå (”below chance”). Intensjonell produksjon av ukorrekte responser vil kunne trekke i tvil prestasjoner på alle andre utførte tester i undersøkelsen, og indikerer at psykometriske abnormaliteter i det enkelte tilfellet er enten ikke-tolkbare, eller indikative for manglende samarbeidsvilje eller malingering, fremfor en reell svekkelse (Bianchini, m.fl., 2001a).

FCT-tester har ofte benyttet gjenkjenningsprøver, på bakgrunn av kunnskap om hukommelse som er blitt tilegnet gjennom empiriske og kliniske undersøkelser (Nies & Sweet, 1994). Denne kunnskapen tilsier at en FCT gjenkjenningsprøve stiller egentlig svært lave krav til hukommelse, slik at en ”normal” deltaker, eller pasient med hjerneskade, kognitiv svekkelse eller psykiatriske problemer, vil kunne oppnå nær perfekte resultater

(Guilmette, Hart, Giuliano & Leninger, 1994). En studie av Slick, Tan, Strauss og Hultsch (2002) indikerer at simulering av hukommelsesvansker er den av simuleringstrategiene som benyttes hyppigst, og et individ som er motivert til å simulere hukommelsesvansker vil typisk gjøre dette med utgangspunkt i det de oppfatter som reelle hukommelsesvansker (Nies & Sweet, 1994). En ”normal” deltaker eller pasient uten klinisk kunnskap om gjenkjenning og hukommelsesvansker, vil derfor kunne anta at FCT-gjenkjenningsprøven har en høyere vanskelighetsgrad enn det den faktisk har. Dersom personen har til hensikt å simulere kognitiv svikt, vil dette kunne motivere personen til å prestere suboptimalt. En høy grad av tilsynelatende vanskelighet (”apparent difficulty”) vil slik kunne øke sensitiviteten til en test (Bianchini, m.fl., 2001a). I tillegg til tilsynelatende vanskelighet, er det avgjørende for sensitiviteten til testen at den også kan gi inntrykk av å være en reell hukommelsestest. Majoriteten av symptomvaliditetstester benytter seg derfor av FCT *gjenkjenning* for å avdekke simulering.

FCT-prosedyren ble for første gang modifisert og standardisert av Hiscock og Hiscock (1989). Deres prosedyre, Digit Memory Test (DMT), består av 72 ”trials” (testledd; 3 blokker med 24), hvor en presentasjon av en femsifret tallrekke etterfølges av en tovalgsgjenkjenningsprøve. DMT har en viktig fordel ved at det inkluderes økende tidsperioder mellom presentasjon av stimulus og gjenkjenningsprøven, noe som gir en illusjon av at vanskelighetsgraden eskalerer. Flere studier indikerer at bruken av tilsynelatende gradert vanskelighet øker simuleringstendensen på de testleddene som virker å være vanskeligere (Binder, 1993).

Prestasjoner som er signifikant svakere enn sjansenivå er de sterkeste indikatorene for simulering (Iverson & Binder, 2000). Et problem ligger imidlertid i at de ofte har lave sensitivitetsrater (Bickart, Meyer & Connell, 1991; Guilmette, Hart & Giuliano, 1993). Studier viser eksempelvis at simulanter, i likhet med visse pasientgrupper med reelle kognitive svikttegn (for eksempel Huntington-pasienter), vil kunne prestere *likt* eller *bedre* enn sjansenivå, og vil derfor ikke kunne fanges opp av denne metodikken (Iverson & Binder, 2000; Nies & Sweet, 1994). Under sjansenivå fanger kun opp de simulantene med de aller mest åpenbare strategiene (Haines & Norris, 1995).

For å øke sensitivitet, foreslo Binder (1993) å benytte ”bedre-enn-sjansenivå”- cutoff-skårer. Denne tilnærmingen benytter typisk det som kalles ”gulv-effekten” (”empirical floor”) som baserer seg på de svakeste prestasjonene til kjente kliniske pasientgrupper. Med denne type cutoff, fant Guilmette m.fl. (1993) sensitivitets- og spesifisitetsrater på henholdsvis 74 og 100 % på DMT. Sammenlignet med resultater på DMT hvor under-sjansenivå-tilnærmingen

ble benyttet, der sensitivitets- og spesifisitetsrater var på henholdsvis 34 og 100 %, er dette en enorm forbedring i sensitivitet. Lignende økninger i sensitivitetsrater ved bruk av bedre-enn-sjansenivå- cutoff-skårer har blitt påvist i flere andre studier (for eksempel Guilmette m.fl., 1994; Tombaugh, 1997). Nyere FCT-tester har derfor benyttet denne tilnærmingen. Et flertall av valideringsstudiene har imidlertid benyttet simulatordesign (Rogers, 1997b). Denne typen design er blitt kritisert av flere forfattere ettersom at mulighetene for å generalisere konklusjonene til faktiske pasientpopulasjoner reduseres (Bianchini m.fl., 2001a; Haines & Norris, 1995; Nies & Sweet, 1994; Rogers, 1997b). Dette kan medføre negative konsekvenser med tanke på den reelle klassifiseringsevnen til det enkelte instrumentet. En mer omfattende gjennomgang av kritikken mot simulatordesign kommer senere i oppgaven.

FCT-tester antas i dag å være de sterkeste indikatorene på suboptimal prestering. Dagens FCT-tester kan deles inn i to kategorier; a) tallgjenkjenningsprosedyrer og b) alternative prosedyrer (Iverson & Binder, 2000). Et utvalg av FCT-tester vil bli gjennomgått i de følgende avsnitt. Dette utvalget er basert på hvilke tester som antas å være mest i bruk i dagens nevropsykologiske settinger (Slick, Tan, Strauss & Hultsch, 2004), samt søk i fagbasene ISI Web of Knowledge og Google Scholar (oktober 2007) på de symptomvaliditetstester som er nevnt hos Bianchini m.fl. (2001a) og Iverson og Binder (2000). Bianchini m.fl. (2001a) og Iverson & Binder (2001) er blant de to nyeste gjennomgangene av prestasjonsbaserte evalueringprosedyrer som finnes i malingeringlitteraturen. Søket i fagbasene gir et ytterligere inntrykk av hvilke tester i de ulike kategoriene som er mest utbredt og studert (se tabell 1 på neste side). I dette søket ble også motivasjonstester som benytter andre prosedyrer enn FCT funnet. Kun symptomvaliditetstester som benytter FCT-prosedyren er valgt ut til denne oppgaven. De utvalgte testene inkluderer Portland Digit Recognition Test (PDRT), Victoria Symptom Validity Test (VSVT), Test of Memory Malingering (TOMM) og 21-Item Test.

Tabell 1  
Resultater fra søk i fagbaser<sup>a</sup>

Kategori	Test	Google Scholar	ISI Web
FCT	PDRT	252	26
	VSVT	158	15
	CARB	110	17
	CFCT	3	1
	TNM	10	0
	MDMT	0	0
Alt. FCT	TOMM	358	81
	21-Item	117	5
	48-Pictures	3	0
	LMT	33	6
	MMDT	1	0
Andre	Rey FIT	127	22
	DCT	130	12

FCT: Forced Choice Testing; PDRT: Portland Digit Recognition Test; VSVT: Victoria Symptom Validity Test; CARB: Computerized Assessment of Response Bias; CFCT: Computerized Forced-Choice Test; TNM: Test of Neuropsychological Malingering; MDMT: Multi-Digit Memory Test; TOMM: Test of Memory Malingering; 21-Item: 21-Item Test; 48-Pictures: 48-Picture Test; LMT: Letter Memory Test; MMDT: Malingered Memory Deficit Test; Rey FIT: Rey 15-Item Test; DCT: Dot Counting Test.

<sup>a</sup> Søkeord som ble benyttet var det fulle navnet ved hver test, for eksempel "Portland Digit Recognition Test", med unntak for 21-Item Test hvor søkeord inkluderte "force\* AND choice".

### 2.2.1 Tradisjonelle FCT-tester: Tallgjenkjenningsprosedyrer

Hiscocks DMT-prosedyre og tallgjenkjenning er i stor grad rammeverk for flere av dagens symptomvaliditetstester (Bianchini m.fl., 2001a). De har alle et tovalgsformat for hvert testledd. Av de mest anvendte tester i dag som benytter tallgjenkjenningsprosedyren, kan nevnes Portland Digit Recognition Test og Victoria Symptom Validity Test. Sistnevnte er en komputert versjon av tallgjenkjenningsprosedyren.

#### *Portland Digit Recognition Test*

Portland Digit Recognition Test (PDRT; Binder & Willis, 1991, Binder, 1993) er blant de første formelt utviklede forced-choice symptomvaliditetstester og er stadig blant de testene som anvendes oftest (Bianchini m.fl., 2001a; Slick m.fl., 2004). PDRT består av 72 testledd med visuell gjenkjenning av femsifrede tallrekker presentert auditivt. De 72 testleddene er delt inn i to sett med 36 testledd på hver. Lette testledd ("easy items") refererer til de 36 første

testleddene, og vanskelige testledd ("hard items"), til de 36 siste, for å gi et inntrykk av tilsynelatende vanskelighetsgrad. PDRT skiller seg fra Hiscocks DMT ved at den inkluderer distraheringsoppgaver i tidsforsinkelsene (fremfor de tradisjonelle passive forsinkelsesperiodene), hvor pasienten blir bedt om å telle bakover mellom presentasjon av stimulus og gjenkjenningsprøven (Bianchini m.fl., 2001). Testen vil slik ha en tilsynelatende økende vanskelighetsgrad. Distraheringsperiodene øker fra 5 sekunder i de første 18 testleddene til 15 sekunder for den andre fjerdedelen, og 30 sekunder for den siste fjerdedelen.

PDRT er en av de best dokumenterte symptomvaliditetstestene (Bianchini m.fl., 2001a; Doane, Greve & Bianchini, 2005; Iverson & Binder, 2000). I den opprinnelige valideringsstudien sammenlignet Binder og Willis (1991) pasienter med lette hodeskader som hadde pågående erstatnings- eller trygdesaker, pasienter med dokumentert hjerneskade og normale simulatorer, med normale kontroller, pasienter med dokumentert hjerneskade uten pågående erstatnings- eller trygdesaker og pasienter som hadde affektive lidelser uten pågående erstatnings- eller trygdesaker. I alle sammenligninger presterte pasienter som hadde pågående erstatnings- eller trygdesaker svakere enn pasienter som ikke hadde det. Over en fjerdedel av pasientene som hadde pågående erstatnings- eller trygdesaker skåret under den svakeste prestasjonen (det vil si 39/72) til pasientene med dokumenterte skader som ikke hadde slike saker. Cutoff-skåre for lette og vanskelige testledd og "total" ble satt til henholdsvis 19/36, 18/36 og 39/72 (Binder & Willis, 1991). Resultatene fra denne studien og andre lignende studier indikerer en falsk positiv rate på 0 % (det vil si en spesifisitet på 100 %) hos pasienter med moderat til alvorlig traumatisk hjerneskade (TBI) som ikke hadde potensielt motiverende forhold (for eksempel erstatningssaker) (Binder & Willis, 1991; Binder, 1993; Binder & Kelly, 1996; Bianchini, Mathias, Greve, Houston & Crouch, 2001b). Skårene for vanskelige testledd og total ble funnet å være mest effektive. Studier med kjente kliniske grupper indikerer imidlertid at de originale cutoff-skårene til PDRT er for konservative, og at høyere cutoff-skårer vil kunne øke sensitiviteten, uten å forårsake en uakseptabelt lav spesifisitetsrate (95 %) (Greve & Bianchini, 2006b)

Ettersom PDRT har en relativt lang administreringstid (ca 40 minutter), foreslo Binder (1993) bruken av en forkortet PDRT-versjon i situasjoner hvor det foreligger liten grunn til mistanke om suboptimal prestering. Denne forkortede versjonen tillater avslutning av testadministrasjon dersom visse kriterier for prestasjonsnivå er oppfylt. En studie av Doane m.fl. (2005) demonstrerte en overensstemmelse med standardversjonen av PDRT på 99.5 %. Disse resultatene antyder at det foreligger liten fare for falske positive ved anvendelse av denne forkortede versjonen. Studier av andre forkortede versjoner av PDRT har også fått



empirisk støtte med hensyn til deres nytteverdi som symptomvaliditetsmål (Greiffenstein, Baker & Gola, 1994; Youngjohn, Burrows & Erdal, 1995).

PDRT finnes videre i en komputert versjon (PDRT-C; Rose, Hall & Szalda-Petree, 1995) som er direkte analog til standardversjonen. PDRT-C produserer både skårer for totalt antall korrekt responser og responslatens. Total cutoff-skåren alene har indikert sensitivitets- og spesifisitetsnivåer på henholdsvis 64 og 89 % (Rose m.fl., 1995). Disse resultatene er lovende, men det er behov for ytterligere valideringsdata.

### *Victoria Symptom Validity Test*

En annen komputert symptomvaliditetstest er Victoria Symptom Validity Test (VSVT; Slick, Hoop, Strauss & Spellacy, 1996; Slick, Hoop & Thompson, 1997). VSVT benytter en forkortet versjon av Hiscocks DMT-prosedyre, og består av 48 testledd delt inn i to vanskelighetsgrader som varierer med hensyn til antall sifre som målstimulus og gjenkjenningssprøven har til felles. I tillegg til dette, øker intervallet mellom stimulus og respons. I en av de tidligste studiene av VSVT fant Slick m.fl. (1996) at både simulatorer og pasienter med pågående erstatnings- eller trygdesaker hadde signifikant forhøyete feilrater sammenlignet med kontroller og pasienter som ikke hadde pågående erstatnings- eller trygdesaker. Resultatene fra denne studien er konsistente med nyere studier (for eksempel Loring, Lee & Meador, 2005; Loring, Larrabee, Lee & Meador, 2007). I likhet med PDRT, ble de ”vanskeligere” testleddene funnet å være de mest effektive med hensyn til å diskriminere mellom hjerneskadepasienter og simulatorer (Loring m.fl., 2007; Slick m.fl., 1996).

VSVT gir mulighet for bayesiske analyser hvilket fasiliterer tolkning av råskårene. Med bayesiske analyser menes beregninger av sannsynligheten for at en skåre er assosiert med en negativ respons skjevhet (”bias”) med utgangspunkt i en gitt basisrate (Iverson & Binder, 2000). Til tross for at responslatens ikke ble analysert for å avgjøre deres nytteverdi med hensyn til korrekt klassifisering, oppdaget Slick m.fl. (1996) at deltakerne med invalide responser identifisert ut fra feilrater alene, hadde dobbelt så lang responstid som kontrollene. Dette tyder på at responslatens som et kvantitativt mål kan bidra i diskrimineringsprosessen. Denney (1999, i Iverson & Binder, 2000) påpeker at til tross for gunstige empiriske funn, bør cutoff-skårene anses som tentative ettersom de kun er basert på 32 personer med hjerneskade.

I VSVT beskrives skårer som er bedre enn sjansenivå som valide, hvilket impliserer at personens prestasjoner ikke indikerer suboptimal prestering. Men, som tidligere gjennomgått, kan skårer som er bedre enn sjansenivå (men under taknivå), i noen tilfeller være assosiert

med suboptimal prestering. For eksempel viser en bayesisk analyse av VSVT-data at en skåre på 16/24 riktige på de ”lette” testleddene er assosiert med en sannsynlighet på 80 % for viljestyrt utvelgelse av ukorrekte responser. Til tross for dette, har bruken av VSVT god empirisk støtte (Loring m.fl., 2005; Loring m.fl., 2007; Slick m.fl., 1996; Slick m.fl., 1997).

### *2.2.2 Alternative FCT-tester*

Motivert ut fra en intensjon om å forbedre sensitivitet, er det utviklet FCT-tester som varierer fra DMT og tallgjenkjenningsprosedyren ved at de benytter andre typer stimuli enn sifre, eksempelvis bokstavrekker (Letter Memory Test, Malingered Memory Deficit Test), ord (21-Item Test) bilder av ukjente ansikt (Recognition Memory Test) og strektegninger (Test of Memory Malingered, 48-Picture Test). Til tross for få konkrete begrunnelser, antas det at mer abstrakt stimulusmateriale som dette vil kunne øke grad av troverdighet som reell hukommelsestest og tilsynelatende vanskelighet. De skiller seg imidlertid lite fra tallgjenkjenningsprosedyren med tanke på metodologi.

#### *Test of Memory Malingered*

Test of Memory Malingered (TOMM; Tombaugh, 1997; Rees, Tombaugh, Gansler & Moczynski, 1998) er blant de mest studerte symptomvaliditetstestene (Bianchini m.fl., 2001a). TOMM benytter en gjenkjenningsprosedyre i tovalgsformat med 50 strektegninger av hverdagsobjekter som stimuli. Samme stimuli presenteres i ”trial 1” og ”trial 2”. ”Retention trial” administreres 15 minutter etter trial 1 og 2.

Valg av stimuli og prosedyre er basert på funn fra flere studier som tyder på at mennesker har en bemerkelsesverdig høy kapasitet for lagring og gjenhenting av visuell informasjon (Nickerson, 1965, 1968; Shepard, 1967; Standing, Conezio & Haber, 1970, i Tombaugh, 1997). I en studie ble for eksempel to deltakere vist 1100 bilder (5 sekunder på hvert) (Standing m.fl., 1970, i Tombaugh, 1997). 30 minutter senere ble gjenkjenningsevnen testet ved presentasjon av 100 bildepar. Hvert av bildeparene inneholdt ett tidligere vist bilde og ett nytt bilde. De to deltakerne identifiserte henholdsvis 95 og 99 % av bildene. Denne evnen er blitt påvist også hos eldre personer og i nevrologisk svekkete populasjoner (Freed, Corkin, Growden & Nissen, 1989; Hart & O’Shanick, 1993; Hupert & Piercy, 1976, 1978, 1979; Kopelman, 1985; Martone, Butters & Trauner, 1986; Park, Puglisi & Smith, 1986; Winograd, Smith & Simon, 1982, alle i Tombaugh, 1997).

En cutoff-skåre på 90 % riktig på trial 2 og retention trial ble i en studie av Tombaugh (1997) validert gjennom 4 eksperimenter. Resultatene fra denne studien viste at TOMM var

relativt insensitiv for reelle hukommelsesvansker hos pasienter med veldokumenterte hjerneskadene – det vil si at den fremviste en høy grad av spesifisitet. Cutoff-skåren diskriminerte mellom deltakere som simulerte hukommelsesvansker og deltakere som ikke simulerte med en sensitivitet og spesifisitet på 100 %. Nyere studier som har benyttet lignende design har i stor grad replisert funnene til Tombaugh (1997), eksempelvis en studie av Powell, Gfeller, Hendricks og Sharland (2004). Her ble resultater på TOMM fra en kontrollgruppe og to simulatorgrupper sammenlignet. Den ene simulatorgruppen fikk instruksjoner om å simulere hukommelsesvansker etter å ha blitt informert om TBI-symptomatologi, mens den andre fikk lignende simuleringinstruksjoner etter å ha blitt informert om strategier for å unngå å bli oppdaget i en testsituasjon. Totalt ga cutoff-skårene fra valideringsstudien en spesifisitet på 100 % og en sensitivitet på 94 %. Videre, til tross for at gruppen informert om TBI-symptomatologi presterte svakere enn gruppen som fikk informasjon om teststrategier, var testen like sensitiv med hensyn til å avdekke suboptimal prestering i begge grupper. TOMM har vist seg relativt insensitiv med hensyn til alder (Ashendorf, Constantinou & McCaffrey, 2004; Teichner & Wagner, 2004), depresjon og angst (Ashendorf m.fl., 2004; Rees, Tombaugh & Boulay, 2001), laboratorieindusert smerte (Etherton, Bianchini, Greve & Ciota, 2005), og andre nevrologiske tilstander (Teichner & Wagner, 2004; Tombaugh, 1997). Det er imidlertid noen pasienter med objektiv alvorlig nevropatologi som skårer under cutoff (Greve, Bianchini & Doane, 2006c). Med dette grunnlaget forventes det at alle nevrologisk skadete pasienter skårer minst 45, med unntak av pasienter med de aller alvorligste skadene.

En svakhet ved valideringsstudien og studien av Powell m.fl. (2004) beskrevet over, ligger imidlertid i at de har benyttet simulatordesign. Som tidligere nevnt, og som gjennomgås senere i oppgaven, vil et slikt design gi problemer med hensyn til generalisering av resultatene, for eksempel at sensitiviteten til testen kan være noe lavere i reelle populasjoner enn indikert i simulatorstudiene (Greve m.fl., 2006a; Haber & Fichtenberg, 2006; Greve m.fl., 2006c).

Etttersom TOMM er en relativt tidskrevende test (opp mot 40 minutter inkludert tidsforsinkelse mellom trial 2 og retention trial), har noen studier undersøkt klassifiseringsverdien til en forkortet versjon, hvor trial 1 administreres som screening for suboptimal prestering (Gavett, O'Bryant, Fisher, & McCaffrey, 2005; O'Bryant, Engel, Kleiner, Vasterling, & Black, 2007). Gavett m.fl. (2005) fant at 100 % av deltakerne som skåret under standard cutoff på trial 1, gjorde det samme på trial 2 og retention trial, hvilket kan tyde på at validiteten øker lite ved administrering av trial 2 og retention trial. Lignende

resultater kom frem i en studie av O'Bryant m.fl. (2007) som indikerer at trial 1 har en akseptabel klassifiseringsrate og kan derfor benyttes som screeninginstrument. En styrke ved denne studien er at den har benyttet en sammenligning kjente kliniske grupper.

Som supplement til TOMM, er TOMM-2 (Tombaugh, Rees, Munson & Gagnon, 2000) og en komputert versjon av TOMM (Tombaugh, Rees og Girard, 1997) utviklet. Disse TOMM-variantene er foreløpig lite omtalt i faglitteraturen, og de vil derfor ikke gjennomgås her. Med utgangspunkt i det empiriske grunnlaget som er blitt gjennomgått her, antas det at standardversjonen TOMM med sikkerhet kan benyttes som indikator for suboptimal prestering. TOMM er videre lett å administrere, og er av de symptomvaliditetstestene som benyttes mest (Slick m.fl., 2004).

### *21-Item Test*

21-Item Test (Iverson, Franzen, & McCracken, 1991, 1994) består av 21 ord som presenteres muntlig etterfulgt av instruksjoner om å gjenkalle fritt så mange som mulig. Deretter følger en FCT-prosedyre i et tovalgsformat hvor pasienten blir bedt om identifisere målordene. 21-Item Test tar ca 5 minutter å gjennomføre og kan administreres i begynnelsen av en undersøkelse som et screeningsinstrument for suboptimal prestering.

I analoge studier demonstrerte 21-Items Test moderate sensitivetsnivåer hvor mellom 20 og 80 % av deltakerne som ble bedt om å simulere ble identifisert (Iverson & Franzen, 1996; Iverson m.fl., 1991;1994). I en meta-analytisk studie av utvalgte instrumenter, fant Vickery, Berry, Inman, Harris og Orey (2001) at til tross for perfekte spesifisitetsrater, var 21-Items Test den av testene med lavest sensitivitet. En løsning på dette problemet kan imidlertid være modifisering av cutoff-skårer. Videre kan prosedyrerelaterte forskjeller delvis bidra til å forklare funnene, eksempelvis forskjeller i anvendte cutoff-skårer og administrering av 21-Item Test i etterkant av tester med høyere vanskelighetsgrad. Empiriske og meta-analytiske studier gir med andre ord ingen entydig støtte til bruken av 21-Items Test i klinisk sammenheng.

### *2.2.3 Oppsummering*

Som gjennomgått over, er det utviklet en rekke såkalte motivasjonstester som kan benyttes ved vurderinger om suboptimal prestering i nevropsykologiske undersøkelser. For at disse testene skal kunne oppnå tilfredsstillende sensitivetsnivåer, er det avgjørende at de fremtrer som reelle nevropsykologiske tester. Prøver for gjenkjennelse er derfor egnet, ettersom de gir inntrykk av dette. I en meta-analytisk gjennomgang av nevropsykologiske

strategier i vurderinger av malinger, beskriver Nies og Sweet (1994) flere studier som indikerer at prøver for gjenkjennelse er de mest sensitive målene når det gjelder å avdekke suboptimal prestering. FCT-prestasjoner som er svakere enn sjansenivå regnes som det mest validerte og mest robuste målet på suboptimal prestering (for eksempel ved DMT). FCT-prosedyrer har økt sensitivitetsnivåer ved å benytte de aller svakeste prestasjonene til kliniske pasienter (nevrologiske, psykiatriske, psykisk utviklingshemmete pasienter) som sammenligningsgrunnlag (for eksempel PDRT og TOMM).

Bianchini m.fl. (2001a) foreslår flere tiltak som vil kunne øke sensitivitetsratene til de enkelte symptomvaliditetstestene, uten å forsake spesifisitetsratene. Et av tiltakene er å øke graden av tilsynelatende vanskelighet. For eksempel vil dette kunne gjøres ved å benytte en større bredde av stimuli (utover tallrekker og strektegninger), å inkorporere økende vanskelighetsgrader (som ved PDRT), øke cutoff-skårer og inkorporere variabler som er nærmest umulig å simulere. I denne sammenheng nevnes kvantitative mål i komputerte versjoner av symptomvaliditetstester, for eksempel inkonsistens i variabler som responslatens (tid og variabilitet), feilrater, intraindividuell variabilitet og intertestvariabilitet. Det antas at få eller ingen mennesker vil kunne simulere troverdig på slike mål (Ray, Engum, Lambert, Nash & Bracy, 1997). Til slutt bør P300 symptomvaliditetstest nevnes, som er en annen komputert teknikk hvor elektrofysiologiske registreringer kombineres med atferdsmessige responser under gjenkjenningssfasen i en FCT-prosedyre (Bianchini m.fl., 2001a). EEG-registreringer under stimuluspresentasjon resulterer i en identifiserbar "event-related potential" (ERP) som benevnes P300. Denne metoden er særlig appellerende ettersom at det antas at P300 er en relativt automatisk fysiologisk reaksjon som derfor ikke påvirkes av viljestyrt responsmodulering. En ulempe ved denne metoden er imidlertid at den er tidkrevende og kostbar.

### 2.3 Spørreskjema

*K. J.s resultater på symptomsjekklistene indikerte et høyt symptomtrykk. Med unntak av skalaer for psykotisisme og fiendtlighet, var skåringer på samtlige skalaer betydelig eleverte; somatisering, tvang, interpersonlig følsomhet, depresjon, angst, fobisk angst og paranoide vrangforestillinger, hukommelse og trøttbarhet. K. J. krysset med andre ord av for mange ulike symptomer og for høy grad av symptomopplevelse innenfor hver.*

### 2.3.1 Om spørreskjema

Malingeringlitteraturen har i hovedsak fokusert på bruken av symptomvaliditetstester. Mange forfattere argumenterer imidlertid for bruken av flere typer mål i vurderingen av malingering (Haines & Norris, 1995; Iverson & Binder, 2000; Nies & Sweet, 1994; Slick, Sherman & Iverson, 1999). Basert på dette har man sett etter alternative metoder for å avdekke malingering, og i de siste tiårene har man blant annet vendt fokuset mot selvrapporteringsinstrumenter som screeningmetoder for malingering. Flere forfattere anbefaler undersøkelse av selvrapporteringer (Haines & Norris, 1995; Nies & Sweet, 1994; Slick m.fl., 1999; Smith, 1997).

Selvrapporteringsinstrumenter og spørreskjema for identifisering av malingering har tradisjonelt vært utformet med det formål å avdekke simulert psykopatologi (Smith, 1997). Selv om simulering av kognitiv svikt fordrer ulike avdekkingsstrategier enn simulert psykopatologi, har spørreskjema som spesifikk avdekkingsmetode for simulert kognitiv svikt benyttet mye av det samme kunnskapsgrunnlaget. Disse metodene baserer seg for eksempel på empirisk kunnskap om *responderingskarakteristika* til simulanter, noe som igjen har gitt opphav til empirisk etablerte *strategier* for å identifisere malingering (Rogers, 1997a). I de følgende avsnitt beskrives identifiseringsstrategiene som benyttes i spørreskjema for avdekking av simulert kognitiv svikt og psykopatologi.

En av flere identifiseringsstrategier innebærer avkrysning (på skjema) av *lavfrekvente symptomer* (Smith & Burger, 1997; Smith, 1997). Med lavfrekvente symptomer menes *atypiske, sjeldne eller usannsynlige symptomer* (Rogers, 1997a). Symptomene regnes som atypiske, sjeldne eller usannsynlige på bakgrunn av empirisk kunnskap om en gitt skade eller forstyrrelse. Ut fra dette har man kunnet utlede bestemte forventninger om hvilke symptomer en reell pasient vil oppleve. Slick m.fl. (1999) viser videre til resultater som tyder på at simulatorer krysser like mye av for symptomer som er uvanlige som for symptomer som er mye vanligere ved kognitiv sekvele, som for eksempel vansker med å huske telefonnumre.

En annen identifiseringsstrategi gjelder avkrysningen av et *overdrevent antall varierte symptomer*. Det antas at personen kan ha adoptert en "more is better"-strategi (Ridenour, McCoy & Dean, 1998; Rogers, 1997a). Fee og Rutherford (1988, i Ridenour m.fl., 1998) sammenlignet symptomene til pasienter som hadde pågående erstatningssaker med pasienter som ikke hadde det, og fant at førstnevnte pasientgruppe rapporterte opptil tre ganger så mange symptomer som pasientene som ikke hadde pågående erstatningssaker. Rogers (1997a) anbefaler at malingering mistenkes i situasjoner hvor pasienten har krysset av for mer enn to

tredjedeler av symptomene. Relatert til avkrysning av et overdrevent antall av symptomer er tendensen til å krysse av for symptomer som går på kryss og tvers av flere diagnoser og som ikke passer inn i én enkelt diagnostisk enhet (Resnick, 1984, i Smith, 1997).

Hvorvidt det er krysset av for en *overdreven grad av symptomopplevelse*, samt *symptomkronisitet* innenfor det enkelte testledd/det enkelte symptom, er en annen identifiseringsstrategi. Denne strategien er basert på en antagelse om at personer som simulerer eller aggraverer krysser av det mest ekstreme alternativet, og deres totalsum overgår slik gjennomsnittet til den reelle relevante pasientpopulasjonen når det gjelder grad av symptomopplevelse (Smith, 1997; Rogers, 1997a).

Spørreskjemaene som gjennomgås her er Structured Inventory of Malingered Symptomatology (SIMS), som i tillegg til å differensiere mellom aggravert/simulert og reell psykopatologi, også inkluderer skalaer som gir muligheter for å vurdere aggravert/simulert kognitiv svekkelse; Neuropsychological Symptom Inventory (NSI), som er spesifikt utformet for vurderinger om malingering i nevropsykologiske undersøkelser alene; Memory Complaints Inventory (MCI) som er utformet for validitetsvurderinger med hensyn til påstander om hukommelsesvansker; og Malingering Detection Scale (MDS), som kan brukes over en rekke kontekster ettersom den baserer seg på kunnskap om atferd under testing og intervju som identifiseringsstrategi. SIMS og MDS er valgt ut på grunnlag av de instrumentene som er nevnt hos Smith (1997), mens NSI og MCI ble funnet gjennom søk i ISI Web of Knowledge og Google Scholar (oktober, 2007). I dette søket ble også SIMS funnet. Søkestrengen som ble brukt i er: ((inventory AND symptom AND exaggeration) OR (inventory AND symptom AND exaggeration AND neuropsych\*)) OR (inventory AND malingering\*) OR (inventory AND malingering\* AND neuropsych\*). Instrumenter som ikke er relevante for nevropsykologiske undersøkelser ble ikke inkludert. Et ytterligere søk i fagbasene på de enkelte instrumentene gir et inntrykk av hvilke spørreskjemaer som er mest utbredt (se tabell 2 på neste side). I de følgende avsnitt vil samtlige instrumenter bli gjennomgått ytterligere.

Tabell 2  
Resultater fra søk i fagbaser<sup>a</sup>

Spørreskjema	Google Scholar	ISI Web of Knowledge
SIMS	80	12
NSI	17	5
MCI	19	2
MDS	6	1

SIMS: Structured Inventory of Malingered Symptomatology; NSI: Neuropsychological Symptoms Inventory; MCI: Memory Complaints Inventory; MDS: Malingering Detection Scale.

<sup>a</sup> Søkeord som ble benyttet var det fulle navnet ved hver test, for eksempel "Structured Inventory of Malingered Symptomatology"

### 2.3.2 Utvalgte spørreskjema

*K. J.s resultater på Structured Inventory of Malingered Symptomatology (SIMS) indikerte i likhet med resultater på symptomsjekklistene, avkrysning for mange varierte symptomer, i tillegg til at flere av symptomene han krysset av for, anses som svært uvanlige i pasientpopulasjoner for øvrig. Totalskåren på 30 var fordelt på følgende måte; 8 på nevrologisk svekkelse, 8 på amnesi, 8 på lav intelligens og 6 på affektiv forstyrrelse. Med en validert total cutoff-skåre på 14, indikerte resultatene på dette instrumentet høy grad et tilfelle av simulering.*

#### *Structured Inventory of Malingered Symptomatology*

Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997; Smith & Widows, 2005) er et spørreskjema konstruert for å avdekke simulering av flere ulike tilstander (inkludert kognitiv svikt) ved hjelp av 75 testledd. Testleddene er i "sann"/"usann" format. Testledd ble trukket ut fra to ulike kilder; a) spørsmål fra eksisterende instrumenter og b) empirisk kunnskap om malingeringkarakteristika. Når det gjelder første kilde, ble spørsmål fra eksempelvis Minnesota Multiphasic Personality Inventory (validitetsindeksene), Structured Inventory of Reported Symptoms og Wechsler Adult Intelligens Scale Revised (WAIS-R) trukket ut på basis av tidligere empiriske demonstrasjoner av deres differensieringsevne. Disse testleddene ble så redigert for å inkludere bestemte identifiseringsstrategier og for å øke deres sensitivitet til simulering av en bestemt tilstand. Eksempler på identifiseringsstrategier omhandler avkrysning av usannsynlige og atypiske (bizarre) symptomer, og "approximate answers" (det vil si spørsmål som beskriver en falskhet som sann, for eksempel "Ungarn er hovedstaden i Italia" i Lav Intelligens-skalaen). Videre



ble spørsmål som reflekterte empirisk kunnskap om malingeringkarakteristika innlemmet (Smith & Burger, 1997). Spørsmålene ble så organisert inn i fem skalaer, med 15 testledd hver; Psykose (P), Amnestiske forstyrrelser (AM), Lav intelligens (LI), Affektiv forstyrrelse (AF) og Nevrologisk forstyrrelse (N). De fem tilstandene reflekterer det som i eksisterende litteratur vurderes å være de tilstandene som oftest simuleres (Smith & Widows, 2005).

Valideringsstudien til SIMS benyttet et simulatordesign, hvor 476 studenter ble fordelt til en av åtte ulike simuleringsbetingelser (psykose, amnesi, nevrologisk svekkelse, mani, depresjon, lav intelligens, "fake bad" [generell malingering] og kontrollgruppe) (Smith & Burger, 1997). Studentene i de syv simulatorgruppene ble gitt en testvignett hvor de ble bedt om å forestille seg at de var arrestert for grov vold og for å unnsnippe straff, måtte de simulere en spesifikk forstyrrelse. De ble også advart mot å krysse av for symptomer på andre tilstander enn den de forsøkte å simulere, ettersom det kunne føre til at man ble avslørt. Resultatene på SIMS ble sammenlignet med resultater på andre malingeringsmål; F- og K-skalaen i Minnesota Multiphasic Personality Inventory, 16 Personality Factor Questionnaire, samt deler av Malingering Scale som kan administreres i papir-blyant form. Cutoff skåringer ble utviklet for hver av skalaene i tillegg til en totalsum (det vil si summen av alle fem skalaer). Alle simulatorgrupper ble så sammenlignet med kontrollgruppen.

Resultater indikerte lave interkorrelasjoner og god indre konsistens med verdier varierende fra .80 til .88. Mens de enkelte subskalaene viste seg å ha liten diskriminativ verdi, var totalskåren i SIMS den mest effektive indikatoren for malingering blant instrumentene som ble anvendt. Totalskåren demonstrerte en korrekt klassifikasjon av 96 % av simulatorgruppen og 88 % av kontrollgruppen. Forfatterne peker imidlertid på at subskalaene kan gi kvalitativ informasjon om problemområder for pasienten, eventuelt hvilke tilstander personer forsøker å simulere. En svakhet ved studien beskrevet over er den høye basisraten som ble benyttet. Åtte grupper var simulatorgrupper, det vil si 88 % av deltakerne. Å ta utgangspunkt i en lavere basisrate, hvilket er nærere realiteten, vil kunne redusere effektiviteten til instrumentet (dette poenget undersøkes nærmere senere i oppgaven). En annen svakhet er bruken av simulatordesign.

Andre studier av SIMS er blitt gjennomført, både ved hjelp av simulatordesign og sammenligning av kjente kliniske grupper, men da med sikte på å vurdere effektiviteten til SIMS hovedsakelig med tanke på malingering av psykopatologi (Vitacco, Rogers, Gabel & Munizza, 2007; Merkelbach & Smith, 2003; Edens, Otto & Dwyer, 1999). Med unntak av funn fra en studie av Edens m.fl. (1999), er resultatene fra valideringsstudien i stor grad blitt replisert i disse studiene. Edens m.fl. (1999) viser til resultater som tyder på at

spesifisitetsrater var lavest blant deltakere som rapporterte klinisk signifikante nivåer av symptomtrykk, noe som gir en fare for falske positive i kliniske populasjoner.

Det er kun valideringsstudien som har undersøkt SIMS' effektivitet med tanke på å avdekke malingering av kognitive svekkelser (cerebral eller hjerneorganisk svekkelse). På bakgrunn av denne studien, hevder forfatterne at SIMS er en egnet screeningmetode for malingering av kognitive svekkelser, så vel som psykopatologi, som vil kunne bidra til å øke korrekt klassifikasjonsrate (Smith & Widows, 2005). De advarer imidlertid mot å benytte instrumentet alene, og anbefaler en kombinasjon av malingeringsmål, inkludert symptomvaliditetstester, i diagnostiseringsprosessen.

### *Neuropsychological Symptom Inventory*

Spørreskjemaet Neuropsychological Symptom Inventory (NSI; Rattan, Strom & Dean, 1987) ble utviklet for å undersøke pasienters nevropsykologiske funksjon, samt belyse symptomatologi hos den enkelte pasient i et selvrapporteringsformat. For å kunne diskriminere mellom malingering og reell symptomrapportering, baserer NSI seg på identifiseringsstrategier som overdreven avkrysning av antall symptomer og grad av symptomopplevelse. Total utfyllingstid er mellom 15 og 20 minutter, og krever leseferdigheter på barneskolenivå. Instrumentet består av 117 testledd som er delt inn i to kategorier. De første 17 testleddene omhandler den medisinske bakgrunnen til pasienten og pasientens familie, hvor pasienten velger mellom svaralternativer "sant" eller "usant". De resterende 100 testledd omhandler nevropsykologisk symptomatologi, og besvares på en skala, på et enig-uenig kontinuum, hvor (1) er "mesteparten av tiden", (2) er "fra tid til annen", (3) er "svært sjelden" og (4) er "aldri". I to uavhengige studier avdekker faktoranalyser fire domener; generell nevropsykologisk fungering, perseptuell/motorisk fungering, sensorisk/perseptuell fungering og oppmerksomhet/konsentrasjon (McCoy, Hill, Dean & Decker, 1998; Jones, 1993, i Gelder, Titus & Dean, 2002). Cronbach's alfa var .97, og test-retestrelabilitet var .62 (McCoy m.fl., 1998).

Ved en diskriminant funksjonsanalyse er det i flere studier blitt utledet en løgnindeks ("lie index") i NSI (Ridenour, McCoy & Dean, 1996 i Gelder m.fl., 2002). Løgnindeksen har hatt stor suksess med hensyn til å diskriminere mellom simulator- og kontrollgrupper, samt at den har en akseptabel falsk positiv rate. Løgnindeksen består av 10 testledd, hvorav 5 omhandler generell nevropsykologisk fungering, 3 omhandler motorisk fungering, 1 omhandler sensorisk/perseptuell fungering, og 1 omhandler oppmerksomhet/konsentrasjon.

I en studie av Ridenour m.fl. (1998) viser forfatterne til en "hitrate" på 84 % når en cutoff-skåre på 300 ble benyttet, som ga en korrekt klassifisering av 88 % av malingeringstilfellene, og 73 % av reelle pasienter (pasienter med hodepineplager). Gelder m.fl. (2002) viste imidlertid til økte sensitivitetsrater; 95 %. Som forventet, overdrev deltakere i malingeringgruppen symptomene de kryssert av for, samt at de krysset av for et signifikant høyere antall symptomer enn de to pasientgruppene. Dette er konsistent med tidligere funn (Ridenour m.fl., 1998; Ridenour m.fl., 1996, i Gelder m.fl., 2002). Det er viktig å være oppmerksom på at selv om løgnskalaen er sensitiv nok til å differensiere mellom de personene som forsøker å aggrave eller simulere og pasienter som har reelle fysiske plager, er det en fare for å feilaktig gi individer som lider av angst og depresjon som følge av hjerneskade, diagnosen malingering.

På bakgrunn av disse resultatene, konkluderer forfatterne med at NSI i kombinasjon med andre validitetsmål, vil være av stor verdi med hensyn til å forbedre "hit rate" i identifisering av malingering i nevropsykologiske utredninger. Forfatterne påpeker selv at bruken av simulatordesign er en svakhet ved begge NSI-studiene. Begge studiene har i tillegg anvendt grupper som har vært skjeve med hensyn til kjønnsfordeling. Dette kan være en metodologisk ulempe, ettersom det eksisterer funn som tyder på at det forekommer signifikante kjønnsforskjeller med hensyn til malingering (Rogers, Salekin, Sewell, Goldstein & Leonard, 1998).

### *Memory Complaints Inventory*

Memory Complaints Inventory (MCI; Green, Allen & Iverson, 1999) er et selvadministrert komputert spørreskjema som er designet for å kvantifisere selvrapperte hukommelsesvansker. MCI har 58 testledd fordelt på ni skalaer som er utformet for å kunne måle ulike aspekter ved opplevde hukommelsesvansker. Deltakerne i valideringsstudien var 158 pasienter med lette hodeskader henvist til nevropsykologisk undersøkelse på grunn av pågående erstatnings- eller trygdesaker. Deltakerne ble så fordelt til en av to grupper ut fra resultater på to tidligere validerte symptomvaliditetstester; en "reell"-gruppe og en "simulering"-gruppe. Resultatene fra sammenligningen avslørte at "simulering"-gruppen hadde differensiert eleverte MCI-skalaer. En styrke ved denne studien er at den benytter kliniske grupper fremfor simulatordesign, som forbedrer den ytre validiteten til konklusjonene. Det er imidlertid behov for ytterligere undersøkelser av MCI, ettersom det på bakgrunn av søk i fagbaser (Søkeord "Memory Complaints Inventory" i ISI Web of

Knowledge, oktober, 2007), antas å være få eller ingen studier (utover den beskrevet her) av effektiviteten til MCI med hensyn til å bidra i vurderinger om malingering.

### *Malingering Detection Scale*

Malingering Detection Scale (MDS; Barkemeyer & Callon, 1989) er et screeninginstrument som består av 29 testledd som skåres av klinikerens i etterkant av intervju med pasienten. MDS ble utformet med det formål å differensiere mellom nevrologiske pasienter med organiske sykdommer og simulanter, og baserer seg på karakteristika som korrelerer høyt med malingering og som tidligere er blitt beskrevet i litteratur om malingering (Callon, Jones, Barkemeyer & Brantley, 1990, i Smith, 1997).

Instrumentet er delt inn i to hoveddeler. Den første omhandler *pasientens intervjuatferd*, hvor klinikerens observerer hvorvidt spesifikke karakteristika er til stede under intervjuet. Eksempler er hvorvidt pasienten gir spontant uttrykk for en overdrevet tillit til klinikerens ferdigheter, forsøker å manipulere klinikerens, avviser alternative forklaringer med hensyn til symptomenes etiologi, gir bestemte responser til spørsmål, og hvordan pasienten reagerer på uenighet mellom seg selv og klinikerens. Den andre delen omhandler *tilsynelatende formål* med atferden til pasienten, for eksempel ansvarsunngåelse eller oppnåelse av et gode. En cutoff-skåre på 8 eller høyere er assosiert med minst en 95 % sjanse for malingering (Callon m.fl., 1990, i Smith, 1997).

I valideringsstudien ble 122 nevrologiske pasienter delt inn i grupper basert på bestemte kriterier, for eksempel inkonsistens mellom selvrapporterte symptomer og etablert kunnskap om symptomer etter hjerneskade, funn fra nevropsykologisk undersøkelse og laboratoriefunn. Dersom to av de tre kriteriene ble oppfylt, ble deltakeren plassert i malingeringgruppen (n=30). Indre konsistens var på .93 og intertaterreliabilitet med to uavhengige undersøkere var på .94. En diskriminant analyse viste at en cutoff-skåre på 7.6 som klassifiserte 95 % av deltakerne korrekt. Tre deltakere (10 %) i malingeringgruppen og tre reelle pasienter (3 %) ble feilklassifisert. I en kryssvalideringsstudie med 66 pasienter (17 i malingeringgruppen og 49 reelle pasienter), ble 99 % av deltakerne riktig klassifisert, og ingen av de reelle pasientene ble feilklassifisert.

Et problem med valideringsstudien er at grunnlaget for testleddene er uklart, og flere av testleddene ble ikke operasjonelt definert. Eksempelvis var et av atferdsformålene beskrevet som oppnåelse av "et abstrakt gode". Videre lå det et potensial for kontaminering av predikatorer og kriterier ved at fordelingen av pasienter til grupper også tjente til utvikling

av instrumentet (Smith, 1997). I tillegg er det, som forfatterne selv påpeker, et behov for å avgjøre instrumentets generaliserbarhet på tvers av pasientpopulasjoner.

### *2.3.3 Oppsummering*

Til tross for at mistanke om simulert psykopatologi og simulert kognitiv svikt fordrer ulike undersøkelsesmetoder, baserer de ulike spørreskjemaene seg på mange av de samme identifiseringsstrategiene. En svakhet ved spørreskjema utviklet for avdekking av simulert kognitiv svekkelse er at det finnes svært få studier som systematisk har forsøkt å evaluere deres effektivitet, i tillegg til at flere av disse studiene har benyttet simulatordesign. Behovet for ytterligere forskningsdata er derfor stort. Som flere av forfatterne selv påpeker, bør spørreskjemaer benyttes i kombinasjon med andre mål som en bekreftende eller avkreftende informasjonskilde. Alternativt, dersom de benyttes alene som screeningmetode, bør de benyttes som en indikator for hvorvidt ytterligere undersøkelser og vurderinger om malingering bør gjennomføres (Smith, 1997).

Det generelle inntrykket er at bruken av spørreskjema i validitetsvurderinger i Norge fremdeles er forholdsvis sjelden (Schanke m.fl., personlig kommunikasjon, april, 2007). SIMS er imidlertid et av de instrumentene som er inkludert i multisenterprosjektet ”Symptomvaliditet og differensialdiagnostikk i nevropsykologiske undersøkelser” (Schanke m.fl., personlig kommunikasjon, mai 2007). Utvikling og bruk av spørreskjema anses som positivt ettersom det bidrar til større bredde i typer avdekkingsmetoder som er tilgjengelige for klinikerne. Flere ulike typer mål vil kunne styrke påliteligheten i vurderinger om malingering.

### 3. DISKUSJON

*Konklusjonen til nevropsykologen ble trukket på bakgrunn av flere validitetsmål i undersøkelsen. Nevropsykologen hadde funnet flere inkonsistente forhold i resultatene fra standardtestene, samt indikasjoner om simulering fra to spesifikke metoder; en symptomvaliditetstest og et spørreskjema. Funnene ble vurdert å bekrefte hverandre i retning av at testprestasjoner ikke kunne anses som valide gjenspeilinger av K. J. reelle kognitive fungering. Nevropsykologen sluttet seg derfor til at validiteten ved den nevropsykologiske undersøkelsen i sin helhet måtte trekkes i tvil. Billeddiagnostiske metoder hadde ikke funnet noe patologisk, og den eneste relativt sikre konklusjonen fra den nevropsykologiske undersøkelsen var at funnene gjenspeilet simulering av kognitive svekkelser.*

Aldri før har nevropsykologer hatt den bredden i spesifikke metoder som de har til disposisjon i dag, både i antall og varianter. Malingering, som vi har sett, kan være vanskelig å avgjøre ut fra standard nevropsykologisk testing alene. Også inkonsistente testresultater, eller symptomvaliditetsindikatorer, som kan være svært viktige som deler av bekreftende dokumentasjon, har vist seg å være utilstrekkelige alene. I ”Veileder i klinisk nevropsykologi” utarbeidet av NNF i 2002, postuleres det at ”når klienten kan oppnå økonomiske, trygdemessige eller andre fordeler ved å få påvist en sentralnervøs skade eller dysfunksjon, anbefales bruk av spesifikke metoder eller prosedyrer for å belyse troverdigheten av undersøkelsesresultatene”. Terskelmodellen for malingering av nevropsykologiske svekkelser (Threshold Model of Malingered Neuropsychological Impairment), foreslår spesifikke retningslinjer og flere ulike kriterier som vil utløse grundigere undersøkelser ved mistanke om malingering (Pankratz & Binder, 1997). Blant disse kriteriene er for eksempel ”ikke bestått” resultat på spesifikke avdekkingsmetoder. I tråd med dette, hevder Nies og Sweet (1994) at aktuariske prediksjoner om malingering er bedre enn klinisk bedømmelse.

Spesifikke metoder har vist seg å kunne øke nevropsykologens sikkerhet og tiltro til egne beslutninger om validiteten i nevropsykologiske funn. For eksempel, i en studie av Binder og Trueblood (1997), ble 60 nevropsykologer bedt om å bedømme hvorvidt malingering hadde forekommet i fire kliniske testprotokoller hvor malingering allerede var etablert. Noen av psykologene fikk kun resultater fra standard nevropsykologisk testing til å bedømme ut fra, andre fikk FCT-data i tillegg. FCT-data bidro til en signifikant økning i treffsikkerhet (konfidens) med hensyn til diagnostisering. Blant psykologene som ikke mottok

FCT-data, var grad av konfidens like lav i de tilfellene hvor de hadde satt korrekt diagnose som de tilfellene hvor diagnosen var gal.

Det er helt avgjørende å undersøke det metodologiske grunnlaget som nevropsykologiske metoder hviler på ettersom dette vil kunne ha direkte innvirkning på metodens kliniske nytteverdi. For eksempel er det viktig å betrakte fremgangsmåter i valideringen av symptomvaliditetstester og spørreskjema. Etter en vurdering av metodologiske hensyn ved studier av spesifikke avdekkingsmetoder, vil symptomvaliditetstester og spørreskjema sammenlignes og diskuteres, før fokus til slutt settes mot holdninger og praksis ved mistanke om malingering.

### 3.1. Metodologiske hensyn ved studier av spesifikke metoder

Ettersom malingering er en atferd som personen aktivt skjuler, er det vanskelig å identifisere simulanter i reelle pasientpopulasjoner som kan inngå i eksperimentelle grupper (Bianchini m.fl., 2001a). For å omgå dette dilemmaet, har effektiviteten til validitetsmetoder i stor grad blitt undersøkt ved hjelp av friske frivillige (simulatorer), som oftest studenter, som får instruksjon om å simulere en gitt form for svikt/forstyrrelse, og dermed produsere invalide resultater (Rosenfeld, Sands & van Gorp, 2000). På denne måten sikres en høyere grad av indre validitet. Det vil si at man med større visshet kan slå fast at den eksperimentelle manipuleringen er den eneste forklaring på endringer i avhengige variabler (Vickery m.fl., 2004). Simulatordesign, også kalt analoge studier, har imidlertid typisk medført særlig ett klart metodologisk problem; begrensninger i muligheter for å generalisere forskningsresultater til reelle pasientpopulasjoner (ytre validitet). Dette problemet oppstår som resultat av at deltakerne på flere måter er ulike målpopulasjonen (simulatorer versus simulanter), samt at det har vært en tendens til å undervurdere betydningen av basisrater. På bakgrunn av slike metodologiske svakheter, anbefaler forfattere valideringsstudier også benytter kjente kliniske grupper i sitt design, for å sikre ytre validering (Rogers, 1997b). Men også denne typen design er blitt kritisert for å medføre problemer. I de neste avsnittene vil ovennevnte metodologiske utfordringer bli diskutert.

#### 3.1.1 Oversette basisrater

Forekomsten av malingering (basisraten) er vanskelig å avgjøre ettersom simulanten aktivt skjuler sine handlinger. Det finnes likevel flere studier som rapporterer anslag av basisrater. Et av de metodologiske problemene i flere analoge valideringsstudier har vært

tendensen til tilsynelatende å ignorere hvilken innvirkning faktiske basisrater kan ha på presisjonen i klassifiseringen. Dette har typisk skjedd ved at man har tatt utgangspunkt i en basisrate som er for høy (Bianchini m.fl., 2001a; Nies & Sweet, 1994; Rosenfeld m.fl., 2000). For eksempel ble en basisrate på 88 % benyttet i valideringsstudien av SIMS (Smith & Burger, 1997). Det vil si at 88 % av deltakerne var simulatorer. Som konsekvens, vil den reelle klassifiseringsevnen til det enkelte instrumentet være ukjent og/eller feilaktig, og i verste fall står man i fare for feilklassifisere reelle tilfeller.

I en gjennomgang av empiriske studier som rapporterte basisrater, fant Larrabee (2003) en prevalensrate på 40 % hos 1.363 saker henvist til nevropsykologiske undersøkelser med lette hodeskader hvor det eksisterte økonomiske motiver. Her ble det rapportert at basisrater varierte fra så lavt som 15 % til så høyt som 64 %. Greiffenstein, Baker og Gola (1994) fant lignende tall da basisrater ble vurdert ut fra multiple validitetsmål. Mittenberg, Patton, Canyock og Condit (2002) fant i en undersøkelse blant nevropsykologer, at 29 % av personskaadesaker, 30 % av uførhetssaker og 8 % av medisinske saker involverte sannsynlig malingering og symptomaggraving. I to norske studier sammenlignet Aresvik (1999) og Aresvik, Sundet, Schanke, Gjeruldsen og Bergersen (2001) pasienter henvist til nevropsykologisk undersøkelse med spørsmål om løsemiddelskade i forbindelse med erstatnings eller -trygdesaker, med pasienter med hjerneslag som ikke hadde slike saker. Resultatene fra denne studien viste at forekomst av simulering eller overdrivelse var høyere i gruppen med pasienter henvist med spørsmål om løsemiddelskade. Mer spesifikt ble det funnet invalide protokoller i 33 % av tilfellene med løsemiddelskade, og 5,6 % av tilfellene med hjerneslag.

Det antas likevel at forekomsten av malingering i generelle pasientpopulasjoner henvist til nevropsykologisk undersøkelser er lav, og at det er mindretall av pasienter som vil simulere eller aggravere symptomer for å oppnå ytre gevinster (Rogers, 1997b). For eksempel har andre studier funnet noe lavere forekomster, 8 % (Rogers, 1997b) og 11 % (Martens, Donders & Millis, 2001) blant pasienter med hjerneslag.

Klassifiseringsevnen til et (nevropsykologisk) mål vil som nevnt kunne variere enormt, avhengig av den faktiske basisraten til det aktuelle fenomenet og hvilke basisrater som er benyttet i studier av effektiviteten til et instrument (Rosenfeld m.fl., 2000). Dette vil igjen kunne ha direkte negative konsekvenser for klassifisering i praksis. Man kan tenke seg et scenario hvor man har benyttet en basisrate på 50 %, og optimale cutoff-skårer og sensitivitets- og spesifisitetsrater er blitt utledet. Med en hypotetisk forekomst av malingering på 15 % (som er nærmere realiteten) og et utvalg på 1000 personer, vil man kunne forvente at



150 av disse overdrev eller fabrikkerte symptomer, og 850 gjorde det ikke. Videre, med en sensitivitetsrate utledet fra studien på 77 % vil man kunne korrekt klassifisere 115 av de 150 som overdrev eller fabrikkerte, hvorav 35 blir feilklassifisert som ”ærlige”. Med en spesifisitet på 90 %, vil man altså stå i fare for å feilklassifisere 87 av de ”ærlige” 850. Som resultat vil man ha en korrekt klassifikasjonsrate på kun 57 % (det vil si man har 115 korrekte prediksjoner av totalt 202 prediksjoner), og nesten halvparten (43 %) av de 1000 vil feilaktig bli klassifisert som simulanter. Valideringsstudier bør derfor være oppmerksomme på, og ta hensyn til, faktiske basistrater dersom de benytter et simulatordesign.

### 3.1.2 *Simulatorer versus simulanter*

Mens simulatordesign løser noe av problemet i forhold til å finne eksperimentelle deltakere, introduserer simulatordesign et ”simulering-malingering”-paradoks som best beskrives av Rogers (1983, i Rogers & Cruise, 1998); ”asking subjects to comply with instructions to fake in order to study subjects who fake when asked to comply”. Dette paradokset refererer til de ukjente eller begrensede mulighetene for å generalisere fra simulatorer til faktiske simulanter. Dette demonstreres for eksempel i studier av TOMM, hvor valideringsstudien som benyttet simulatordesign fant en sensitivitetsrate på 94 % (Tombaugh, 1997), mens studier som har benyttet sammenligning av kjente grupper, har funnet sensitivitetsrater som har variert fra 45 til 64 % med opprinnelige cutoff-skårer (90 % riktig) (Greve m.fl., 2006a; Haber & Fichtenberg, 2006; Greve, Bianchini & Doane, 2006c). I studiene har imidlertid spesifisitetsraten til TOMM vært stabil på 100 %.

Det argumenteres for eksempel at motivasjonen for å oppnå en relativt beskjeden belønning for å delta i en studie er kvalitativt ulik motivasjonen til en person som står foran muligheten til å oppnå økonomisk gevinst via trygde- eller erstatningssaker (Nies & Sweet, 1994). Incentivene i simulatordesign skiller seg fra reelle incentiver både med hensyn til type, omfang og sannsynlighet for å kunne oppnå dem. I tillegg til dette, vil en simulant stå overfor en fare for å utløse en rekke *negative* konsekvenser dersom han eller hun skulle avsløres, for eksempel strafferettslige, sosiale, økonomiske sanksjoner (Rogers & Cruise, 1998).

Når det gjelder instruksene deltakerne i simuleringstudier har mottatt, har disse blitt kritisert for å være uforståelige, uspesifikke og korte (Rogers, 1997b). Det er derfor uklart om deltakerne har hatt en klar forståelse av hva som forventes av dem, samt at man ikke kan være sikker på at deltakerne har fulgt instruksene. Mange studier har videre presentert scenarioer for deltakerne, i et forsøk på å øke ytre validitet. Rogers (1997b) hevder imidlertid at

inkluderingen av scenarier vil mer sannsynlig bidra til å redusere mulighetene for generalisering.

Relatert til dette, er hensyn om mulige forskjeller i hva slags kunnskap om hjerneskade og nevropsykologiske tester simulatorer og simulanter kan ha tilegnet seg forut for studiene. Man kan for eksempel tenke seg at simulanter prestasjoner i noen tilfeller kan være påvirket av tidligere tilegnet kunnskap om testprestasjoner som er typiske for pasienter med hjerneskade – med andre ord kunnskap som vil kunne gjøre dem bedre rustet til å simulere troverdig (Haines & Norris, 1995; Nies & Sweet, 1994). Det er derfor mulig at reelle simulanter er mer sofistikerte i sin fremgangsmåte enn ”naive” simulatorer. Studier har for eksempel vist at friske deltakere som er bedt om å simulere anvender betydelig mer åpenbare strategier enn mistenkte simulanter, for eksempel ved å prestere signifikant svakere enn reelle pasienter som er mistenkes for å simulere (Haines & Norris, 1995). I tillegg har deltakere i simuleringstudier sjelden fått tid eller muligheter til å forberede simuleringstrategier. Forskningsfunn tyder likevel på at det å ha erfart medisinske og/eller nevropsykologiske undersøkelser ikke nødvendigvis har betydning for hvor ”flink” man er til å simulere på en troverdig måte (Inman & Berry, 2002; Ju & Varney, 2000; Vickery m.fl., 2004). Det er imidlertid tenkelig at reelle pasienter i høyere grad vil kunne være motivert til aktivt å søke informasjon om sekvele ved hjerneskade og nevropsykologiske undersøkelsesmetoder, selv om funn tyder på at dersom de blir gitt nok tid, vil også friske deltakere forberede seg og søke relevant informasjon (Slick m.fl., 2002).

### *3.1.3 Kjente kliniske grupper*

Selv om bruken av kjente kliniske grupper løser mange av problemene som oppstår i analoge design, medfører også denne typen design metodologiske utfordringer. Kjente simulanter er som nevnt vanskelige å identifisere, blant annet på grunn av lav diagnostisk sensitivitet og lav grad av enighet mellom klinikere (Rogers, 1997b). Klinikere har eksempelvis vært uenige med hensyn til hvor stor grad av atferdsendring kan skapes av lette traumatiske hjerneskader, noe som har bidratt til å øke kompleksiteten i klassifiseringen av kognitive svikttegn. Studier med kjente grupper bør kun inkludere de pasientene hvor det er overbevisende grunner til å mistenke malingering. ”Søker trygd eller erstatning” som eneste kriterium for å inkludere en pasient som ”sannsynlig simulant”, anses derfor som utilstrekkelig.

Greiffenstein m.fl. (1994) foreslo i denne sammenheng å operasjonalisere malingering ved at to eller flere av følgende kriterier var oppfylt; 1) to eller flere alvorlig svake resultater

på nevropsykologisk testing, 2) en usannsynlig symptomhistorie, 3) total uførhet i yrkesliv eller sosiale forhold, eller 4) påstander om fjern (remote) hukommelsestap. Disse kriteriene er imidlertid utilstrekkelige, blant annet fordi de ikke inkluderer differensialdiagnostiske kriterier, og fordi visse reelle tilstander etter hodeskade vil kunne tilfredsstillende kriteriene (Slick m.fl., 1999). For eksempel tilfredsstillende ”postcommotio syndrom” (PCS) kriteriene 1) til 3), (Gualtieri, 1995). Kriteriene til Slick m.fl. (1999) for Malingered Neurocognitive Dysfunction (MND) vil derfor være mer hensiktsmessig, og kan anvendes i operasjonaliseringen av malingering (Greve m.fl., 2006a; Greve, Bianchini & Doane, 2006c; O’Byrant, Engel, Kleiner, Vasterling, & Black, 2007). MND-kriteriene ble utviklet i forhold til behovet for spesifikke, objektive diagnostiske kriterier, og forfatterne foreslo definisjoner og spesifikke diagnostiske kriterier for, *definitiv*, *sannsynlig* og *mulig* malingering av kognitive svekkelser (disse vil gjennomgå nærmere senere i diskusjonsdelen). Bianchini m.fl. (2001) anbefaler å sammenligne pasienter som av overbevisende grunner er klassifisert som ”sannsynlig simulert” (som har motivasjonelle forhold) med pasienter klassifisert som ”ikke-simulert” (som også har motivasjonelle forhold). En fare består imidlertid i feilklassifisering på grunn av den sirkulære kvaliteten til slike definisjoner. Ytterligere forskning innen skadevirkninger etter lett traumatisk hjerneskade vil være nyttig med tanke på å forbedre definisjoner av malingering og identifisering av grupper hvor malingering kan forekomme.

Metodologiske svakheter har vært et gjennomgående problem i forskning på, og ved utvikling av, metoder for å avdekke malingering (Bianchini m.fl., 2001; Haines & Norris, 1995; Nies & Sweet, 1994; Rogers, 1997b). Begrensningene ved simulatordesign medfører for eksempel at de faktiske sensitivitets- og spesifisitetsratene til symptomvaliditetstester ikke er kjent. Mens simulatordesign sikrer indre validitet, men mangler ytre validitet, har man det motsatte tilfellet ved bruk av kjente kliniske grupper. Grad av ytre validitet vil her kunne være forbedret, men utilstrekkelige inklusjonskriterier kan sette indre validitet i fare. Det kan derfor synes som om man blir nødt til å foreta en avveining mellom ytre og indre validitet i valg av forskningsdesign. Ytre validering av diagnoser er imidlertid den generelle normen i utviklingen av psykologisk testing, og bør være det også i utviklingen av symptomvaliditetstester. Rogers (1997b) anbefaler derfor en kombinasjon av simulatordesign og sammenligning av kjente kliniske grupper i valideringen av symptomvaliditetsmål. Å kryssvalidere resultater fra forskning på de enkelte instrumentene med ulike kliniske utvalg er også blitt anbefalt (Bianchini m.fl., 2001a; Smith, 1997).

Oppmerksomheten vil nå rettes mot metodenes kliniske nytteverdi, og deres pålitelighet med tanke på klassifisering og diagnostisering.

### 3.2 Sammenligning av symptomvaliditetstester og spørreskjema

Et viktig hensyn som bør overveies ved bruk av spesifikke avdekkingsmetoder, er hva slags slutninger man kan trekke ut fra personens prestasjoner og selvrapporteringer. Iverson (2006) påpeker at begrepene ”å prestere under evne”, ”suboptimal prestering” eller ”ikke optimal prestering” *ikke* er synonyme med ”symptomaggraving”, ”symptomoverdrivelse” og ”simulering”. Til tross for dette finnes det klare tendenser i litteraturen til å bruke disse begrepene om hverandre. Mens ”symptomaggraving”, ”symptomoverdrivelse” og ”simulering” er relativt utvetydige begreper som brukes om selvrapportering av symptomer ved intervju, respondering på spørreskjema, og/eller ved kliniske observasjoner, refererer begrepene ”å prestere under evne”, ”suboptimal prestering” og ”ikke optimal prestering” til atferd under testing. Klinikeren vil kunne trekke en slutning om at suboptimal prestering som er blitt observert hos en pasient er en manifestasjon av symptomaggraving eller -simulering, men denne slutningen er en *sekundær klinisk slutning*, hvor suboptimal prestering er den *primære* kliniske slutning.

Grovt sett kan man se på klassifisering av malingering av kognitiv svikt som en todelt prosess: a) Det må kunne dokumenteres at simulering har forekommet (for eksempel ved hjelp av nevropsykologiske avdekkingsmetoder), og b) Det må avdekkes at personen var motivert til å simulere av ytre insentiver. Først presenteres metodenes rolle i etablering av motivasjon.

#### 3.2.1 Klassifisering; Etablering av motivasjon

En avgjørende del av vurderinger om hvorvidt malingering har forekommet i det enkelte tilfellet, dreier seg om individets intensjon og motivasjon for å simulere (kognitive) symptomer. Iverson (2007) beskriver hovedsakelig tre underliggende årsaker eller motiver for å simulere symptomer. Den første årsaken er malingering. Som tidligere gjennomgått, vil diagnosen ”malingering” (DSM-IV) kunne stilles kun i de tilfeller hvor *ytre* insentiver antas å ha motivert personen til å simulere kognitive symptomer. I både ICD-10 og DSM-IV beskrives det generelt flere typer ytre motiver som kan motivere en person til å aggravere (overdrive eksisterende symptomer) eller simulere symptomer. Blant disse kan nevnes *direkte økonomiske insentiver* (for eksempel utbetaling av trygd eller erstatning), oppnåelse av bedre

leveforhold (som bolig), *unngåelse av yrkesmessige krav* (ofte forsterket av sekundære økonomiske stimuli som for eksempel uførhetstrygd), *unngåelse av verneplikt eller militære oppdrag*, *unngåelse av strafferettslige søksmål*, og *oppnåelse av ulovlige substanser* (WHO, 1992; Roesch & Golding, 1986; Rubinsky & Brandt, 1986; Heubrock & Petermann, 1988; Iverson & Binder, 2000). Både DSM-IV og ICD-10 skiller mellom indre og ytre insentiver, hvor indre motivasjon beskrives som mindre åpenbar eller ukjent, og med et mindre klart formål.

Den andre av de tre underliggende årsakene til simulering dreier seg om atferd som er motivert av et *indre* ønske om å oppnå oppmerksomhet fra ens lege eller familie, differensieres fra malingering (Iverson, 2007). I slike tilfeller vil DSM-IV-diagnosen ”factitious disorder”, eller F68.1 (”Viljestyrt produksjon eller simulering av symptomer eller dysfunksjon, enten fysisk eller psykisk [selvpåført lidelse]”) i ICD-10, være både mer passende og hensiktsmessig (WHO; APA; Rogers, Sewell & Goldstein, 1995 i Heubrock & Peterson, 1998). Vedkommende antas å ha et behov for å inneha rollen som ”den syke” for å kunne oppnå sosial ”belønning” i form av oppmerksomhet, omtanke og sympati fra personer rundt. Det mest alvorlige tilfellene kan klassifiseres som Munchausen syndrom.

Den tredje underliggende årsaken til simulering omhandler situasjoner hvor ”rop om hjelp” rettet mot helsepersonell utarter seg som simulering, uten at dette nødvendigvis klassifiserer til diagnosen ”factitious disorder” (Iverson, 2007). Dette gjelder personer med psykiske vansker (for eksempel somatoforme lidelser, konversjonslidelser, affektive forstyrrelser, andre psykiske vansker [angst, psykose] og somatiske og psykososiale belastninger) og/eller personer som opplever gitte former for symptomer, men som frykter at de ikke tas på alvor av helsepersonell, eller ikke får den oppmerksomhet eller behandling de mener de bør.

Når det gjelder klassifisering av malingering, vil validerte symptomvaliditetstester og spørreskjema kunne gi indikasjoner om *at* pasienten presterer suboptimalt og/eller simulerer kognitive symptomer, men de kan ikke si noe om *hvorfor*. Med andre ord kan de ikke avdekke hva slags motivasjon som ligger til grunn for atferden, eller hvorvidt motivasjonen er av en indre eller ytre art, noe som er avgjørende med hensyn til å kunne differensiere mellom tilstandene beskrevet over og diagnosen ”simulant” (ICD-10) eller ”malingering” (DSM-IV). Simulering ved nevropsykologiske undersøkelser er altså ikke ensbetydende med malingering, noe som understreker behovet for differensialdiagnostiske vurderinger. Differensialdiagnostiske vurderinger om årsaker til simulering behøver derfor ytterligere informasjon fra andre kilder som kan bidra til å etablere hvorvidt motivasjon er av en indre

eller ytre. For eksempel vil sjekklister for psykologiske og emosjonelle symptomer kunne være nyttige her med tanke på å avdekke eventuelle psykiske vansker. Noen av spørreskjemaene diskutert i denne oppgaven vil imidlertid kunne gi noe informasjon om denne typen symptomatologi, som for eksempel SIMS og NSI. Likevel, ettersom disse er utviklet som screeningmetoder for malingering, er det behov for grundigere instrumenter.

### 3.2.2 *Simulering versus symptomaggraving*

En annen differensialdiagnostisk begrensning ved spesifikke avdekkingsmetoder dreier seg om forutsetningene for å skille mellom simulering og aggraving av symptomer. Selv om intensjonell produksjon av falske symptomer og overdrivelse av eksisterende (dokumenterte) symptomer er relatert og begge knyttes til diagnosen malingering, er de ikke synonyme (Iverson, 2007). Ren simulering er uvanlig, mens overdrivelse av symptomer er mer vanlig (Nies & Sweet, 1994). Dette illustreres ved at pasienter kan både ha en sannsynlig organisk betinget kognitiv svikt og en testprotokoll som tyder på symptomaggraving. Svake prestasjoner på symptomvaliditetstester, eller resultater fra spørreskjema som er indikative for simulering, ekskluderer derfor aldri encefalopati. På bakgrunn av slike saker, argumenterer flere forfattere for at malingering ikke forekommer i en "enten–eller" form, men at det rangeres langs et kontinuum, fra mild overdrivelse til grov symptomaggraving (Heubrock & Peterson, 1998). Rogers (1990) foreslår for eksempel en gradering (mild, moderat, alvorlig) som reflekterer alvorlighetsgrad i pasientens forvrengning. Graderinger av malingering har imidlertid ikke vært mulig å operasjonalisere og har derfor mer en heuristisk verdi (Tombaugh, 1996, i Heubrock & Peterson, 1998).

### 3.2.3 *Vurdering om påliteligheten av symptomvaliditetstester og spørreskjema*

Symptomvaliditetstester og spørreskjema er begge utviklet spesifikt for å kunne avdekke simulering i nevropsykologiske undersøkelser. Symptomvaliditetstester er den avdekkingsmetoden som benyttes mest, og som det er mest forskningslitteratur om (se tabell 1 og 2). Bruken av spørreskjema er til sammenligning en ung tilnærming til avdekking av simulert kognitiv svikt som det stadig er relativt lite forskning på, selv om selvrapporeringinstrumenter, inkludert spørreskjema, har en lang tradisjon i avdekking simulert psykopatologi.

Klassifiseringsevnen til symptomvaliditetstester har sin styrke i at de hevdes å kunne avdekke en persons *intensjon* om å simulere kognitive svekkelser. Dette gjøres ved hjelp av forced-choice testing som kan gi sterke indikasjoner om hvorvidt en person intensjonelt

presterer under sin fulle evne på nevropsykologiske tester. Til forskjell fra symptomvaliditetstester, som er prestasjonsbaserte evalueringsprosedyrer, er spørreskjema ment å avdekke malingering ved å sammenligne pasientens svar med selvrapporteringskarakteristika som antas å være typisk ved malingering. For eksempel har man funnet at simulanter typisk krysser av for et høyere antall uvanlige symptomer enn reelle pasienter, som for eksempel ukontrollerbar latter (Slick m.fl., 1999).

Funn fra spesifikke avdekkingsmetoder bør som tidligere nevnt inngå i den diagnostiske vurderingen. Malingering har imidlertid ingen formelle diagnostiske kriterier i DSM-IV eller ICD-10, ettersom malingering ikke klassifiseres som en mental lidelse per se, men som en tilstand som fordrer klinisk undersøkelse (Iverson & Binder, 2000). Slick m.fl. (1999) foreslo på bakgrunn av arbeidet til blant andre Rogers (1990), Greiffenstein, Baker og Gola (1994), Faust og Ackley (1998), spesifikke diagnostiske kriterier for, *definitiv*, *sannsynlig* og *mulig* malingering av kognitive svekkelser (MND). Basert på omfang av funn som støtter diagnosen, gir kriteriene til Slick m.fl. (1999) muligheter for klinikerer å gi uttrykk for grad av sikkerhet i diagnosen. I følge Slick m.fl. (1999) kan diagnostiseringsprosessen deles inn i flere trinn. For diagnosen definitiv MND, må det foreligge et vesentlig ytre insentiv (for eksempel erstatningskrav, søknad om trygd, unngåelse av strafferettslig forfølgning eller militærtjeneste), samt klar og definitiv evidens for overdrivelse eller fabrikkering av kognitiv svekkelse. Atferden må videre kunne beskrives som rasjonell, informert og viljestyrt. I tillegg må alternative forklaringer på atferden utelukkes (for eksempel nevrologiske, psykiatrisk og utviklingsmessige diagnoser).

MND-kriteriene gjenspeiler også på mange måter en vurdering av påliteligheten til ulike avdekkingsmetoder (se tabell 3 på neste side for en oversikt over MND-kriteriene). Blant kriteriene er kriterium B., som omhandler informasjon fra nevropsykologisk testing, herunder spesifikke avdekkingsmetoder, samt inkonsistens mellom testdata og nevropsykologisk kunnskap om mønster i kognitiv fungering, atferdsobservasjoner og komparentopplysninger (Slick m.fl., 1999). Ettersom FCT-tester i dag antas å være de sterkeste indikatorene på suboptimal prestering, er resultater fra denne typen tester inkludert i kriteriene for definitiv og sannsynlig malingering (se tabell 4 på s. 37 for en oversikt over diagnostiske kategorier for MND). *Definitiv negativ respons skjevhet* betegner testprestasjoner som er under sjansenivå ( $p < .05$ ) på en eller flere FCT-prosedyrer. Dette er i tråd med Rogers (1997c) som hevder at prestasjoner som er under sjansenivå FCT-tester er relativt sikre indikatorer på suboptimal prestering. Disse prosedyrene rettfærdiggjøres som

Tabell 3 Oversikt over og forklaring av kriterier for Malingered Neurocognitive Dysfunction (etter Slick m.fl., 1999)

Kriterier for Malingered Neurocognitive Dysfunction (MIND)	Forklaring
A. Tilstedeværelse av vesentlig ytre insentiv	Minst ett identifiserbart og vesentlig ytre insentiv for simulering av symptomer er til stede ved tidspunktet for undersøkelsen
B. Data fra nevropsykologisk testing	Indikasjoner fra nevropsykologisk testing om simulering
B1. Definitiv negativ respons skjevhet	Testprestasjoner som er under sjansenivå ( $p < .05$ ) på $1 \leq$ FCT-prosedyrer.
B2. Sannsynlig respons skjevhet	Prestasjoner som indikerer simulering på $1 \leq$ godt validerte psykometriske tester eller prestasjonsbaserte metoder spesifikt utviklet for å avdekke simulering.
B3. Diskrepans mellom testdata og kunnskap om mønster i kognitiv fungering	Diskrepansen må være konsistent med et forsøk om å overdrive eller fabrikkere nevropsykologisk dysfunksjon (f.eks. sterke prestasjoner på en hukommelsestest, men svake prestasjoner på en oppmerksomhetstest).
B4. Diskrepans mellom testdata og observert atferd	Prestasjoner på $2 \leq$ tester som måler samme domene er inkonsistente med observert kognitivt funksjonsnivå (f.eks. ingen tilsynelatende språkvansker i samtale, men svært svekkete prestasjoner på tester som stiller krav til språk).
B5. Diskrepans mellom testdata og pålitelige informasjonskilder	Prestasjoner på $2 \leq$ tester som måler samme domene er inkonsistente med personens daglige kognitive fungering beskrevet av minst én pålitelig komparentinformant.
B6. Diskrepans mellom testdata og dokumentert bakgrunnshistorie	Usannsynlig svake prestasjoner på $2 \leq$ standardiserte kognitive tester som er inkonsistent med dokumentert nevrolgisk eller psykiatrisk historie.
C. Data fra selvrapportering	Atferd som indikerer malingering (bekreftende dokumentasjon), men som alene er utilstrekkelige for å stille diagnose.
C1. Diskrepans: Medisinsk-/psykososial sykdomshistorie	Rapporterte symptomer er inkonsistente med dokumentert sykdomshistorie
C2. Diskrepans: Kunnskap om mønster i kognitiv fungering	Rapporterte symptomer er usannsynlige i forhold til antall, mønster, alvorlighetsgrad, forventninger basert på nevropsyk. kunnskap om skadetype.
C3. Diskrepans: Observasjoner av atferd	Rapporterte symptomer er inkonsistente med observert atferd.
C4. Diskrepans: Komponentopplysninger	Rapporteringer er inkonsistente med informasjon fra andre pålitelige informanter.
C5. Indikasjoner om simulering på validerte selvrapporteringsinstrumenter	C3, C4 tilfredsstiltes og godt validerte validitetsmål i selvrapporteringsinstrumenter indikerer sterkt overdrevne eller fabrikkererte symptomer.
D. Atferd som tilfredsstiller B- og C-kriterier som ikke fullt kan forklares av psykiatriske eller nevrolgiske forstyrrelser eller psykisk utviklingshemming.	Atferd som tilfredsstiller kriterier fra B- og C-grupper er produkter av informerte, rasjonelle og viljestyrte handlinger som er utført med den hensikt å oppnå eksterne insentiver (kriterium A), men som ikke kan forklares av psykiatriske eller nevrolgiske forstyrrelser eller psykisk utviklingshemming, eller andre årsaker for simulering (for eksempel behov for å inneha rollen som "den syke").



Tabell 4 Oversikt over diagnostiske kategorier for Malingered Neurocognitive Dysfunction (etter Slick m.fl., 1999)

Kriterienr.	Definitiv	Sannsynlig	Mulig
1. kriterie	A	A	A
2. kriterie	B1	$2 \leq \text{av B2-B6}$ eller $1 \text{ av B2-B6}$ og $2 \leq \text{av C1-C5}$	$1 \leq \text{av C1-C5}$
3. kriterie	D	D	D eller kriterier for definitiv eller sannsynlig MND er tilfredsstilt, men ikke D

Note: Se tabell 3 for forklaringer på de ulike kriteriene.

validitetsindikatorer, til tross for metodologiske begrensninger i valideringsstudier, på grunnlag av uangripelige matematiske sannsynlighetsestimater. Under-sjansenivå-tilnærmingen gir perfekte spesifisitetsrater (100 %), ettersom det kun er personer som simulerer som presterer svakere enn sjansenivå.

*Sannsynlig negativ respons skjevhet* beskriver prestasjoner på en eller flere godt validerte spesifikke avdekkingsmetoder som indikerer simulering. Blant disse metodene kan nevnes bedre-enn-sjansenivå symptomvaliditetstester. Til forskjell fra under-sjansenivå-tester, bør større forsiktighet utøves i forhold til å konkludere på bakgrunn av FCT-tester som benytter bedre-enn-sjansenivå-strategier (Rogers, 1997c). Dette er fordi valideringsstudier til instrumenter som benytter den strategien har brukt simulatordesign. Disse testene bør derfor benyttes som én av flere validitetsindikatorer. En kombinasjon av flere validerte symptomvaliditetstester vil kunne øke validiteten og forbedre sensitivitets- og spesifisitetsrater i vurderinger om eventuell simulering (Bianchini m.fl., 2001a; Trueblood & Schmidt, 1993). Denne anbefalingen støttes av studier som har undersøkt effekten av simulering på enkelttester alene (Nies & Sweet, 1994). Resultatene demonstrerer at simulering på én enkelt test er enklere enn å simulere svekkelse gjennom den nevropsykologiske undersøkelsen i sin helhet. For eksempel, har friske deltakere vist seg inkonsistente med hensyn til å opprettholde simulatorstatusen gjennom den totale undersøkelsestiden (Strauss m.fl., 2002; Nies & Sweet, 1994). Videre er det ikke avgjort at pasienten vil betrakte den enkelte testen som relevant for typen svekkelse han eller hun forsøker å simulere, og han eller hun vil dermed prestere normalt. Kriteriene for sannsynlig og mulig MND er basert på samme kriterier som for definitiv MND, men for disse diagnosene er det færre sikre holdepunkter.

Som beskrevet over, anses FCT-prosedyrene for å være de mest pålitelige metodene i avdekkingen av malingering, særlig de som benytter under-sjansenivå-tilnærmingen, ettersom

disse er de best validerte av de spesifikke avdekkingsmetodene. Disse inkluderes derfor i høyeste grad i diagnostisering av malingering hos Slick m.fl. (1999). Til sammenligning, hevder forfatterne at det til nå ikke finnes noen objektive indikasjoner basert på selvrapporteringsmønstre som er tilstrekkelig validerte. Data fra selvrapportering er likevel ett av de diagnostiske kriteriene (kriterium C.) for både sannsynlig og mulig MND, og funn fra *godt validerte* selvrapporteringsinstrumenter (C5.) er inkludert her (for eksempel Minnesota Multiphasic Personality Inventory). Etersom spørreskjema for simulert kognitiv svikt ikke ennå kan anses å være godt validerte, vil de falle utenfor dette kriteriet. Dette får implikasjoner i praksis – man bør være forsiktig med å legge for mye vekt på funn fra spørreskjema i vurderinger om simulert kognitiv svikt.

Selv om MND-kriteriene ble utviklet for åtte år siden, tyder gjennomgangen av eksisterende forskning på spørreskjemaer i denne oppgaven på at vurderingen til Slick m.fl. (1999) om påliteligheten ved spørreskjema fortsatt er solid. Til tross for at flere av studiene som er blitt gjennomgått i denne oppgaven har vist til høye sensitivitets- og spesifisitetsrater, gjelder de samme metodologiske begrensningene som ellers i forskning på avdekkingsmetoder for malingering; flertallet av valideringsstudiene har benyttet simulatordesign. En ytterligere svakhet er at det ikke på langt nær finnes like mange valideringsstudier av spørreskjema, som det som finnes av studier av symptomvaliditetstester. Som resultat er spørreskjemaer langt fra like godt validerte som mange av symptomvaliditetstestene.

Til sitt forsvar påpeker forfatterne av studier av spørreskjema at instrumentene ikke er ment å benyttes i et vakuum, eller som eneste grunnlag for å avgi en klinisk slutning om hvorvidt det har forekommet malingering. De hevder at de kan og bør benyttes i kombinasjon med flere validitetsmål, eller som en screeningmetode hvor funn som indikerer malingering utløser grundigere undersøkelser. Det er videre en gjennomgående anbefaling om å benytte informasjon fra flere *typer* indikatorer (Nies & Sweet, 1994; Slick m.fl., 1999). Videre, til tross for kritikken av spørreskjema, kan informasjonen som oppnås fra spørreskjema være nyttig. For eksempel kan bruken av selvrapporteringer ses som nok en kilde til kvantifiserbar informasjon som er direkte sammenlignbart med data fra friske kontroller, pasientgrupper og simulantgrupper. Informasjonen fra spørreskjemaer kan videre sammenlignes med informasjon fra andre kilder, for eksempel komparentopplysninger eller observasjoner i sykehusavdelingen, og kan sammenlignes over tid. Slike sammenligninger kan gi holdepunkter for eventuelle inkonsistente forhold. En fordel ved flere av spørreskjemaene, er også at de kan gi et innblikk i områder og symptomer som pasienten sliter med. For eksempel, kan resultater på NSI indikere tilstedeværelse av psykiatriske symptomer som ofte kan

påvirke prestasjoner på nevropsykologiske tester (Gelder m.fl., 2002). Likevel, metoden er avhengig av ytterligere forskning og validering for å bli et pålitelig instrument i vurderinger om malingering. Man bør også være klar over at noen av spørreskjemaene, for eksempel SIMS og NSI, har vist en tendens til lave spesifisitetsrater blant deltakere som lider av angst og depresjon som følge av hjerneskade. Dette vil i verste fall kunne føre til falske positive klassifiseringer av malingering.

### 3.3 Holdninger og praksis ved mistanke om malingering

Ettersom at malingering er et sensitivt tema og klassifisering kan medføre alvorlige konsekvenser for den enkelte pasient, er det viktig at den enkelte kliniker betrakter sine egne holdninger og praksis i situasjoner hvor han eller hun har funnet grunn til å mistenke malingering. Dette gjelder både med tanke på eventuelle fordommer, samt valg og bruk av spesifikke avdekkingsmetoder.

#### 3.3.1 *Mistanke og moralisering*

På bakgrunn av at man har funnet høyere forekomst av malingering i visse populasjoner, har flere forfattere postulert at malingering i høyere grad bør mistenkes hos pasienter fra disse populasjonene, og at det i disse tilfellene bør administreres spesifikke avdekkingsmetoder. For eksempel er basisraten funnet å være høyere blant pasienter med lette hodeskader som har pågående erstatnings eller -trygdesaker, enn blant pasienter med dokumenterte moderate eller alvorlige hjerneskader (Binder & Rohling, 1997). Nies og Sweet (1994) viser til studier hvor resultatene tyder på at forekomsten er høyere hos personer med sosiopatiske trekk og personer med rus- og alkoholproblematikk. Bianchini, Curtis og Greve (2006) fant videre at jo større økonomiske insentiver, jo høyere forekomst av diagnostiserbar malingering. Forfatterne fra disse studiene konkluderer med at klinikerer i høyere grad bør mistenke malingering i tilfeller hvor slike personlige og økonomiske omstendigheter foreligger.

ICD-10- og DSM-IV-definisjonen av malingering angir få spesifikke kriterier for situasjoner hvor malingering bør mistenkes utover motivasjonelle forhold. DSM-IV angir for eksempel fire kriterier for når man skal mistenke malingering. De fire kriteriene består av (1) rettsmedisinske kontekster, (2) tilfeller hvor objektive funn ikke støtter opp under påstander om symptomer eller uførhet, (3) usamarbeidsvillige pasienter, og (4) tilstedeværelse av antisosial personlighet (APA, 1994). På bakgrunn av dette og de ovennevnte basisratene,

anbefales det at man i større grad mistenker malingering hos individer med visse personlighetstrekk; avhengighet i forhold til alkohol og andre rusmidler, og sosiopatiske trekk (Nies & Sweet, 1994).

Rogers (1990) kritiserer denne tankegangen og DSM-modellen for å være puritansk, og for å ha et moralistisk perspektiv på malingering. Han påpeker det han mener er implisitte bedømmelser av "sletthet" ("badness"). Objektivitet går på denne måten tapt i det klinikere foretar bedømmelser angående moralen i det de selv oppfatter som "bedragerisk" og "slett" atferd, og sosiale fordømmelser vil som konsekvens kunne følge. Ashers (1972, i Cunnien, 1997) observasjoner av klinikere og deres behandling av malingering kan være illustrerende i denne sammenheng: "The pride of a doctor who has caught a malingerer is akin to that of a fisherman who has landed an enormous fish".

Rogers (1990) påpeker videre at malingering på ingen måte er begrenset til visse populasjoner, at disse populasjonene ikke har "monopol" på malingering, eller er "flinkere" til å simulere. Malingering er videre, i følge Rogers, vanskelig å vurdere korrekt ettersom det overlapper med en rekke hverdagslige handlinger (for eksempel skulke skole eller arbeid, finne på unnskyldninger for å unnsnippe sosiale sammenkomster). På bakgrunn av dette og sin kritikk mot DSM-modellen, foreslår han en adaptasjonsmodell ("adaptational model") i stedet for DSM-modellen (Rogers, 1990). Adaptasjonsmodellen forklarer malingering som et *valg* i en beslutningsteoretisk kontekst, og begrenser ikke malingering til spesifikke populasjoner. Adaptasjonsmodellen antar at beslutningen foretas i tre omganger: (1) en person oppfatter undersøkelsen/behandlingen som ufrivillig eller ugunstig, (2) personen oppfatter at han eller hun vil kunne tape noe gjennom å bli undersøkt eller at han eller hun har noe å vinne ved å simulere eller aggrave symptomer, og (3) personen ser ikke andre mer effektive måter å oppnå det ønskete målet på. Malingering i militæret kan være et godt eksempel på adaptasjonsmodellen, hvor konteksten er ugunstig (innkalling til militærtjeneste) og belønningen for å "skulke" er store (unngåelse av kampsituasjoner og mulig død). Malingering bør derfor ikke betraktes som en mental lidelse, men heller som en responsstil som kan konseptualiseres i sosialpsykologiske termer.

Relatert til dette er bruken av spesifikke metoder og prosedyrer, som for eksempel symptomvaliditetstester og spørreskjema, hvor det bør vises aktsomhet mot å la metodene bli en slags "moralisk test" hvor "strykarakter" rettfærdiggjør sosial fordømmelse. Malingering-diagnosen bør i alle tilfeller overveies grundig på grunn av muligheten for alvorlige sanksjoner, og hver enkelt kliniker bør derfor betrakte egne fordommer og praksis i validitetsvurderinger. I denne sammenheng vil det også være nyttig å være klar over at

forekomsten av malingering i generelle pasientpopulasjoner henvist til nevropsykologiske undersøkelser i realiteten er lav (Rogers, 1997b). I ICD–10 hevdes det for eksempel at ”simulering er forholdsvis vanlig i rettslige og militære sammenhenger, og heller uvanlig i det sivile liv” (WHO, 1992).

### 3.3.2 Valg av metoder

I tillegg til å se på egne holdninger og fordommer, er det avgjørende at den enkelte kliniker foretar fornuftige og informerte valg av metoder som skal inngå i utredningen. Det er pr i dag ingen ”gullstandard” med hensyn til hvilke spesifikke metoder som bør inngå i en validitetsvurdering, og metodevalg gjøres ofte på grunnlag av individuell erfaring og kompetanse. I noen tilfeller kan dette være uheldig. En undersøkelse blant nevropsykologer viste for eksempel at de fleste anvendte spesifikke avdekkingsmetoder i situasjoner hvor malingering kunne mistenkes, men et flertall av disse anvendte en motivasjonstest (Rey 15 Item Test) som over flere omganger har vist seg å være blant de minst effektive instrumentene, med uakseptabelt lave sensitivitets- og spesifisitetsnivåer (Slick m.fl., 2004). Individuelle variasjoner i bruken av metodene vil også kunne påvirke sensitiviteten til en test. For eksempel vil det å gi en advarsel i en nevropsykologiske undersøkelse om at forsøk på å simulere vil kunne oppdages, kunne resultere i mer sofistikert malingering, fremfor å redusere forekomsten av det (Youngjohn, Lees-Haley & Binder, 1999).

Sensitivitetsnivåer vil videre kunne svekkes gjennom langvarig bruk av bestemte metoder, for eksempel i rettssaker og med tanke på de store mengdene publisert litteratur som finnes om metodene og deres fremgangsmåter. På denne måten vil metodene kunne bli kjent for andre enn nevropsykologer og bli sårbare for manipulering. Det er derfor avgjørende at man har en stadig utvikling av validerte metoder og et mangfold å velge fra.

Av hensyn til å opprettholde sensitivitetsnivåer, vil det være direkte klinisk relevant for den enkelte psykolog å oppdatere seg rutinemessig på relevante empiriske funn, samt å gjøre seg kjent med de begrensningene som eksisterer ved spesifikke avdekkingsmetoder. Det er på dette grunnlaget valg av metoder bør foretas. På denne måten vil man bedre kunne bruke metodene på en kompetent, ansvarlig og informert måte. En slik praksis vil videre være i tråd med ”Prinsipperklæring om evidensbasert psykologisk praksis” vedtatt av Norsk Psykologforening (NPF, 2007).

#### 4. KONKLUSJON

I denne oppgaven er det funnet at det er et behov for spesifikke avdekkingsmetoder i vurderinger av malingering på grunn av viktigheten av tilstrekkelig dokumentasjon i det enkelte tilfellet. Symptomvaliditetstester og spørreskjema er to slike metoder som begge er spesifikt utviklet for å avdekke simulering av kognitiv svikt i nevropsykologiske undersøkelser. De anvender imidlertid ulike avdekkingsstrategier som igjen baserer seg på ulike informasjonskilder; testprestasjoner og selvrapporterte symptomer.

Som det er blitt beskrevet, er det pr i dag ingen spesifikke avdekkingsmetoder som *alene* kan fastslå malingering med sikkerhet, verken symptomvaliditetstester eller spørreskjema. De kan likevel gi betydningsfulle indikasjoner, og i kombinasjon med flere spesifikke avdekkingsmetoder, vil de være nyttige i nevropsykologiske undersøkelser for å kunne foreta en så pålitelig vurdering som mulig. På bakgrunn av gjennomgangen av eksisterende forskningslitteratur, er det videre grunn til å anse symptomvaliditetstester som de mest pålitelige av de to metodene. Fremtidig forskning vil kunne gi ytterligere svar på om spørreskjema vil kunne oppnå en lignende status, men pr dags dato er det ingen spørreskjema som har oppnådd dette.

Litteratur om symptomvaliditet vektlegger nødvendigheten av multimodale tilnærminger i vurderinger om validitet, malingering og differensielle diagnoser (Haines & Norris, 1995; Iverson & Binder, 2000; Nies & Sweet, 1994; Rogers, 1997c; Slick m.fl., 1999). Det anbefales derfor at undersøkelser, i tillegg til nevropsykologiske spesifikke metoder og validitetsindikatorer, inneholder anamnestic informasjon (for eksempel sykehistorie, medisinske data), informasjon om skole- og yrkeshistorie (eventuelle lærevansker), tidligere og aktuelle livspåkjenninger, egenrapportering av symptomer og symptomutvikling og komparentopplysninger (informasjon fra familie, venner).

Det anbefales videre at hver enkelte kliniker objektivt vurderer egne fordommer knyttet til malingering, samt egen praksis med hensyn til valg av hvilke metoder som skal inngå i en validitetsvurdering, og optimal bruk av disse.

I kasustilfellet beskrevet i denne oppgaven, som er en reell sak fra en nevropsykologisk klinikk i Øst-Norge, konkluderte nevropsykologen med at pasienten mest sannsynlig simulerte de påståtte kognitive svekkelsene. I dette tilfellet ga funn fra både symptomvaliditetstester og spørreskjema indikasjoner om simulering. Det understrekes at konklusjonen i kasustilfellet ikke ble trukket på bakgrunn av én symptomvaliditetstest eller ett spørreskjema alene, men ut fra en *kombinasjon* av disse og andre forhold i den

nevropsykologiske undersøkelsen. Dette illustrerer hvordan bruk av flere og varierte spesifikke metoder som symptomvaliditetstester og spørreskjema kan styrke evnen til korrekt klassifisering i kliniske tilfeller.

Nytteverdien til spesifikke metoder, som ved andre nevropsykologiske metoder, avgjøres av deres evne til korrekt klassifisering. Klassifiseringsevnen måles, som vi har sett, ved vurderinger om sensitivitet og spesifisitet. Mangel på sensitivitet og spesifisitet vil i verste fall kunne medføre feilklassifisering, henholdsvis falske negative og falske positive. I praksis vil forskjellen mellom falske positive og falske negative bestå i hvem som vil lide under feildiagnostiseringen (Bianchini m.fl., 2001a). Når det gjelder falske positive, er det pasienten og hans eller hennes nærmeste som berøres. Hva angår falske negative er det samfunnet som berøres. Falske positive funn som indikerer simulering eller overdrivelse av symptomer kan ha økonomiske, yrkesmessige, sosiale og emosjonelle konsekvenser for personen, og vil kunne forfølge personen som et urettmessig ”stempel” i andre situasjoner hvor helsevesenet er involvert. Falske negative funn vil svekke muligheter for å avdekke for eksempel malingering, og kan redusere tilgjengeligheten av ressurser og derved skade troverdigheten til den nevropsykologiske metoden.

Mye av forskningen som finnes pr i dag, og som er gjennomgått i denne oppgaven, er gjennomført i USA. Den amerikanske litteraturen er i stor utstrekning basert på strafferettssaker, hvor målet er å avsløre hvorvidt personen tildekker sannheten ved å påberope seg en psykiatrisk lidelse eller hukommelsestap, og saker som innebærer økonomiske forhold, hvor man ønsker å unngå økonomisk ytelse til mennesker som ikke oppfyller de gitte kriteriene (Slick m.fl., 2002). Malingeringlitteraturen har hatt et særlig fokus på personer som simulerer kognitive tilstander for å unndra seg kriminelle handlinger. Dette er ikke vanlig i Norge, selv om det kan forekomme. I Norge er validitetsvurderinger hovedsakelig aktuelt i saker som dreier seg om erstatningssaker eller trygdeytelser, hvor motivasjonelle forhold kan påvirke prestasjonene, og hvor bruken av flere ulike avdekkingsmetoder blir nødvendig. Det primære er at pasienter får en så riktig vurdering som mulig, og at metodene som benyttes er pålitelige.

## REFERANSER

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM)* (4<sup>th</sup> ed.).
- Aresvik, J. L. (1999). *The comparison external incentives on data validity in the neuropsychological examination. A comparison of three clinical groups*. Levert som hovedoppgave ved Psykologisk Institutt, Universitetet i Oslo, veileder Sundet, K. S.
- Aresvik, J. L., Sundet, K. S., Schanke, A. K., Gjeruldsen, M F., & Bergersen, H. (2000). The influence of financial compensation on the validity of neuropsychological test profiles. *Poster presented at the 23<sup>rd</sup> Annual Mid-Year Conference, INS, Brussels, Belgium, July 12-15, 2000*.
- Ashendorf, L., Constantinou, M., & McCaffrey R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives Of Clinical Neuropsychology, 19*, 125-130.
- Barkemeyer, C. A., & Callon, E. B. (1989). *Malingering Detection Scale*. Baton Rouge, LA: North Street.
- Bianchini, K. J., Mathias C. W., & Greve, K. W. (2001a). Symptom Validity Testing: A critical review. *The Clinical Neuropsychologist, 15*, 19-45.
- Bianchini, K. J., Mathias C. W., Greve, K. W., Houston, R. J., & Crouch, J. A. (2001b). Classification accuracy of the Portland digit recognition test in traumatic brain injury. *The Clinical Neuropsychologist, 15*, 461-470.
- Bianchini, K. J., Curtis, & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: A dose-response relationship? *The Clinical Neuropsychologist, 20*, 831-847.
- Bickart, W. T., Meyer, R. G., & Connell, D. K. (1991). The symptom validity technique as a measure of feigned short-term memory deficit. *American Journal Of Forensic Psychology, 9*, 3-11.
- Binder, L. M. (1993). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of clinical and experimental neuropsychology, 15*, 170-182.
- Binder, L. M., & Rohling, M. L. (1996). Money matters: A meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal Of Psychiatry 153* 7-10.
- Binder, L. M., & Trueblood, W. (1997). Psychologists' accuracy in identifying neuropsychological test protocols of clinical malingerers. *Archives of Clinical Neuropsychology, 12*, 13-27.
- Cunnen, A. J. (1997). Psychiatric and medical syndromes associated with deception. I Rogers, R. (ed.). *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 23-46). New York: Guilford Publications, Inc.



- Doane, B. M., Greve, K. W., & Bianchini, K. J. (2005). Agreement between the abbreviated and standard Portland Digit Recognition Test. *The Clinical Neuropsychologist*, *19*, 99-104.
- Edens, J. F., Otto, R. K., & Dwyer, T. (1999). Utility of the Structured Inventory of Malingered Symptomatology in identifying persons motivated to malingering psychopathology. *Journal of the American Academy of Psychiatry and the Law*, *27*, 387-396.
- Etcoff, L. M., & Kampfer, K. M. (1996). Practical guidelines in the use of symptom validity and other psychological tests to measure malingering and symptom exaggeration in traumatic brain injury cases. *Neuropsychology Review* *6*. 171-201.
- Etherton, J. L., Bianchini, K. J., Greve, K. W., & Ciota, M. A. (2005). Test of Memory Malingering performance is unaffected by laboratory-induced pain: implications for clinical use. *Archives Of Clinical Neuropsychology*, *20*, 375-384.
- Faust, D., & Ackley, M. A. (1998). Did you think it was easy? Some methodological suggestions for the investigation and development of malingering detection techniques. I Reynolds, C. R. (ed.). *Detection of Malingering during Head Injury Litigation* (1-54). New York: Plenum Press.
- Faust, D., Hart, K., Guilmette, T. J., & Arkes, H. R. (1998). Neuropsychologists capacity to detect adolescent malingerers. *Professional Psychology-Research And Practice*, *19*, 508-515.
- Franzen, M. D., Iverson, G. L., & MacCracken, L. M. (1990). The detection of malingering in neuropsychological assessment. *Neuropsychology Review*, *1*, 247-279.
- Gavett, B. E., O'Bryant, S.E., Fisher, J. M., & McCaffrey R. J. (2005). Hit rates of adequate performance based on the test of memory malingering (TOMM) trial 1. *Applied Neuropsychology*, *12*, 1-4.
- Gelder, B. C., Titus, J. B., & Dean, R. S. (2002). The efficacy of neuropsychological symptom inventory in the differential diagnosis of medical, psychiatric and malingering patients. *International Journal of Neuroscience*, *112*, 1377-1394.
- Green, P., Allen, L. M., & Iverson, G. L. (1999). Utility of the memory complaints inventory for identifying symptom exaggeration in mild to moderate traumatic brain injury. *Archives of clinical neuropsychology*, *14*, 743-744.
- Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation, *Legal and Criminological Psychology*, *11*, 131-145.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218-224.
- Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Adult Scale Intelligence Scale.

- Validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology*, 18, 245-260.
- Greve, K. W., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., Swift, D. A., & Ciota, M. (2006a). Classification accuracy of the Test of Memory Malingering in persons reporting exposure to environmental and industrial toxins: Results of a known-groups analysis. *Archives of Clinical Neuropsychology*, 21, 439-448.
- Greve, K. W., & Bianchini, K. J. (2006b). Classification accuracy of the Portland Digit Recognition Test in Traumatic Brain Injury: Results of a known-groups analysis. *The Clinical Neuropsychologist*, 20, 816-830.
- Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006c). Classification accuracy of the Test of Memory Malingering in traumatic brain injury: Results of a known-groups analysis. *Journal of clinical and experimental neuropsychology*, 28, 1176-1190.
- Gualtieri, C. T. (1995). The problem of mild head injury. *Neuropsychiatry, Neuropsychology And Behavioral Neurology*, 8, 127-136.
- Guilmette, T. J., Hart, K. J. & Giuliano, A. J. (1993). Malingering detection: The use of a forced-choice method in identifying organic versus simulated memory impairment. *The Clinical Neuropsychologist*, 7, 56-69.
- Guilmette, T. J., Hart, K. J., Giuliano, A. J., & Leninger B. E. (1994). Detecting simulated memory impairment: Comparison of the Rey 15-Item Test and the Hiscock Forced-Choice Procedure. *The Clinical Neuropsychologist*, 8, 283-294.
- Haber, A. H., & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head trauma sample. *The Clinical Neuropsychologist*, 20, 524-532.
- Haines, M. E., & Norris, M. P. (1995). Detecting the malingering of cognitive deficits: An update. *Neuropsychology Review*, 5, 125-148.
- Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application. *European Journal Of Psychological Assessment* 14, 211-225.
- Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, 11, 967-974.
- Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, 17, 1-23.
- Iverson, G. L., Franzen, M. D., McCracken, L. M. (1991). Evaluation of an objective assessment technique for the detection of malingered memory deficits. *Law And Human Behavior*, 15, 667-676.

- Iverson, G. L., Franzen, M. D., & McCracken, L. M. (1994). Application of a forced-choice memory procedure designed to detect experimental malingering. *Archives Of Clinical Neuropsychology*, *9*, 437-450.
- Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal Of Clinical And Experimental Neuropsychology*, *18*, 38-51.
- Iverson, G. L., & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal Of Head Trauma Rehabilitation*, *15*, 829-858.
- Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Psychology*, *13*, 77-90.
- Iverson, G. L. (2007). Identifying exaggeration and malingering. *Pain Practice*, *7*, 94-102.
- Ju, D., & Varney, N. R. (2000). Can head injury patients simulate malingering? *Applied Psychology*, *7*, 201-207.
- Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*, 410-425.
- Liebert, R. M & Liebert, L. L. (1995). *Science and behavior. An introduction to methods of psychological research* (4<sup>th</sup> ed.). New York: Prentice Hall International, Inc.
- Loring D. W., Lee G. P., & Meador, K.J. (2005). Victoria symptom validity test performance in non-litigating epilepsy surgery candidates. *Journal of Clinical and Experimental Neuropsychology*, *27*, 610-617.
- Loring D. W., Larrabee G. J., Lee G. P., & Meador, K.J. (2007). Victoria symptom validity test performance in a heterogenous clinical sample. *Clinical Neuropsychologist*, *21*, 522-531.
- Martens, M., Donders, J., & Millis, S. R. (2001). Evaluation of invalid response sets after traumatic head injury. *Journal of Forensic Neuropsychology*, *2*, 1-18.
- McCoy, K. D., Hill, S. K., Dean, R. S., & Decker, S. (1998). Reliability and construct validation of a neuropsychological symptom inventory. *International Journal of Neuroscience*, *96*, 127-139
- Merkelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, *18*, 145-152.
- Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology*, *18*, 261-276.
- Mittenberg, W., Patton, C., Canyok, E. M., & Condit, D. C. (2002). Baserates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*, 1094-1102.

- Nies, K. J., Sweet J. J. (1994). Neuropsychological assessment and malingering - a critical-review of past and present strategies. *Archives Of Clinical Neuropsychology*, 9, 501-552.
- Norsk Nevropsykologisk Forening (2002). *Veileder i klinisk nevropsykologi*. Norsk Psykologforening.
- Norsk Psykologforening (2007). *Prinsipperklæring om evidensbasert psykologisk praksis*. Hentet 17.oktober 2007 fra <http://www.psykol.no/Default.aspx?did=9122669>.
- O'Bryant, S. E., Engel, L. R., Kleiner, J. S., Vasterling, J. J., & Black, F. W. (2007). Test of Memory Malingering (TOMM) Trial 1 as a screening measure for insufficient effort. *The Clinical Neuropsychologist*, 21, 511-521.
- Pankratz, L. (1983). A new technique for the assessment and modification of feigned memory deficit. *Perceptual And Motor Skills*, 57, 367-372.
- Pankratz, L., & Binder, L. M. (1997). Malingering on Intellectual and Neuropsychological Measures. I R. Rogers (ed.), *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 223-239). New York: The Guilford Press.
- Powell, M. R., Gfeller, J. D., Hendricks, B. L., & M. Sharland (2004). Detecting symptom- and test-coached simulators with the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, 19, 693-702.
- Rattan, G., Strom, D. A., & Dean, R. S. (1987). The efficacy of a Neuropsychological Symptom Inventory in the differential diagnosis of neurological, depressed and normal patients. *Archives of Clinical Neuropsychology*, 2, 257-264.
- Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment*, 10, 10-20.
- Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives Of Clinical Neuropsychology*, 16, 501-506.
- Reinvang, I. (1997). Klinisk nevropsykologi: status og utfordringer. *Tidsskrift for Norsk Psykologforening*, 34, 290-296.
- Ridenour, T. A., McCoy, K. D., & Dean, R. S. (1998). Discriminant function analysis of malingers' and neurological headache patients' self-reports of neuropsychological symptoms. *Archives of Clinical Neuropsychology*, 13, 561-567.
- Rogers, R. (1990). Models of feigned mental illness. *Professional Psychology: Research and Practice*, 21, 182-188.
- Rogers, R. (1997a). Structured interviews and dissimulation. In R. Rogers (ed.), *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 301-327). New York: The Guilford Press.

- Rogers, R. (1997b). Researching dissimulation. In R. Rogers (ed.), *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 398- 426). New York: The Guilford Press.
- Rogers, R. (1997c). Current status of clinical methods. In R. Rogers (ed.), *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 373-397). New York: The Guilford Press.
- Rogers, R., & Cruise, K. R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior*, 22, 273-285.
- Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law And Human Behavior*, 22, 353-367.
- Rose, F. E., Hall, S., Szalda-Petree, A. D., & Bach, P. J. (1995). A comparison of four tests of malingering and the effects of coaching. *Archives of Clinical Neuropsychology*, 13, 349-363.
- Rosenfeld, B., Sands, S. A., & van Gorp, W. G. (2000). Have we forgotten the base rate problem? Methodological issues in the detection of distortion. *Archives of Clinical Neuropsychology*, 15, 349-359.
- Slick, D. J., Hoop, G., Strauss, E., & Spellacy, F. (1996). Victoria Symptom Validity Test: Efficiency for detection of feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Psychology*, 18, 911-922.
- Slick, D. J., Hoop, G., & Thompson, G. (1997). *The Victoria Symptom Validity Test*. Odessa, FL; Psychological Assessment Resources.
- Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria of malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, 13, 545-561.
- Slick, D. J., Tan, J. E., Strauss, E. & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, 16, 495-505.
- Slick, D. J., Tan, J. E., Strauss, E., & Hultsch, D. F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465-473.
- Smith, G. P. (1997). Assessment of Malingering with Self-Report Instruments. In R. Rogers (ed.), *Clinical Assessment of Malingering and Deception* (2<sup>nd</sup> ed., pp 351-370). New York: The Guilford Press.
- Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, 25, 180-189.

- Smith, G. P., & Widows, M. R. (2005). *Structured Inventory of Malingered Symptomatology (SIMS). Professional Manual*. Florida: Psychological Assessment Resources, Inc.
- Strauss, E., Slick, D. J., Levy-Bencheton, J., Hunter, M., MacDonald, S. W. S., & Hultsch, D. A. (2002). Intraindividual variability as an indicator of malingering in head injury. *Archives of Clinical Neuropsychology, 17*, 423-444.
- Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingered (TOMM): normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives Of Clinical Neuropsychology, 19*, 455-464.
- Tombaugh, T. N. (1997). The Test of Memory Malingered (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*, 260-268.
- Tombaugh, T. N., Rees, L. M., Munson J., & Gagnon, M. (2000). The Test of Memory Malingered 2 (TOMM-2). *Archives of Clinical Neuropsychology, 15*, 818.
- Tombaugh T. N., Rees L., & Girard M. (1997). Comparison of the computerized and paper-and-pencil versions of the test for memory malingered (TOMM) using a simulation design. Abstract from the 17th Annual Meeting, Las Vegas. *National Academy of Neuropsychology*.
- van Gorp W. G., Humphrey, L. A., Kalechstein, A., Brumm, V. L., McMullen, W. J., Stoddard, M., & Pachana, N. A. (1999). How well do standard clinical neuropsychological tests identify malingering? A preliminary analysis. *Journal Of Clinical And Experimental Neuropsychology, 21*, 245-250.
- Verdens Helseorganisasjon (2004). *Den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer*. Oslo: Fagbokforlaget
- Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E. & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology, 19*, 37-48.
- Vickery, C. D., Berry, D. T. R., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology, 16*, 45-73.
- Vitacco, M. J., Rogers, R., Gabel, J., & Munizza, J. (2007). An evaluation of malingering screens with competency to stand trial patients: A known-groups comparison. *Law and Human Behavior, 31*, 249-260.
- Youngjohn J., Burrows, L., & Erdal, K. (1995). Brain-damage or compensation neurosis - the controversial post-concussion syndrome. *Clinical Neuropsychologist, 9*, 112-123.
- Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14*, 511-515.