

## Summary

The research described in this thesis focuses on poor symptom validity of patients clinically referred for psychological assessment. The main objectives of this thesis were to gain more insight into the frequency of poor symptom validity in clinical assessments, how it relates to patients' symptom reports on standard clinical instruments (e.g., memory tests and psychological symptom questionnaires), whether symptom validity is an uni- or multidimensional phenomenon, and the current practices and beliefs of neuropsychologists regarding symptom validity assessment. These aims are addressed in the five studies of this thesis.

**Chapter 1** provides a general introduction on the topic of symptom validity. This chapter addresses the definition of symptom validity, methods to measure symptom validity, and the overlap and demarcation with the concept of malingering. The chapter ends with the aims and outline of this thesis.

**Chapter 2** describes a cross-sectional study on symptom validity in a sample of 183 patients with various psychiatric diagnoses seen in a mental health care institution. All patients were referred for neuropsychological assessment for treatment purposes. A validity test that measures cognitive underperformance and a validity test that measures symptom over-reporting were part of the test battery. In total, 13% of the sample evidenced either underperformance or symptom over-reporting. In 8% of the assessments both validity tests were failed. Underperformance and symptom over-reporting were only modestly correlated. Symptom over-reporting was strongly, positively related to the score on a psychological symptom list. The subsample that only failed the performance validity test (PVT), performed worse on memory tests. These results show that poor symptom validity occurs frequently, and that underperformance and over-reporting are only loosely coupled dimensions. The relation between over-reporting and the score on a psychological symptom list was the most robust.

**Chapter 3** presents a study with the same design as in Chapter 2 but executed in a hospital setting. The Psychology departments of five hospitals participated in this study. The medical specialist referred patients for psychological assessment. With the exception of patients with evident cognitive deficits, clinicians administered the Structured Inventory of Malingered Symptomatology (SIMS), a validity test that targets over-reporting of symptoms. In neuropsychological assessments the Amsterdam Short-Term Memory Test (ASTM), a validity test for cognitive underperformance, was also administered. In total, the data of 469 assessments were analyzed. Background variables, such as age, education, diagnosis upon referral, and involvement in litigation, were taken into account. The sample comprised five main diagnostic groups: neurological conditions, morbid obesity, medically unexplained symptoms (MUS), psychiatric disorders, cognitive complaints not otherwise specified, and medical conditions with cognitive complaints. Symptom over-reporting occurred in 12%-19% of the main diagnostic groups. An exception being the group of morbidly obese patients. These patients were all referred for psychological assessment as part of a screening to determine the eligibility for a gastric by-pass surgery. In this group only one patient

scored deviant on the SIMS. The SIMS score was positively associated with the score on a self-report psychological symptom list. In 29.3% of the neuropsychological assessments the score on the ASTM indicated underperformance. The score on the ASTM was negatively associated with memory performance. Our study did not find an association between poor symptom validity and the presence of financial incentive, such as receiving disability benefits or involvement in litigation. These study findings underscore the importance of symptom validity testing in clinically referred assessments. More research is needed to advance our understanding of which factors contribute to poor symptom validity.

**Chapter 4** describes three studies with the central question of whether symptom validity can best be understood as an uni- or multidimensional concept. Many clinicians conceptualize symptom validity monolithically, and by implication assume that when people over-report symptoms, they also underperform on cognitive tests. Another wide-spread assumption is that malingerers have a preference for feigning memory disorders. In fact, most validity tests are memory-oriented PVTs. These PVTs are also used to determine symptom validity in psychological assessments of patients that do not present with memory disorders as their primary complaint. We wanted to test both assumptions. In the first study participants read a case-vignette and were interviewed about the symptoms that they would and would certainly not feign. Non-specific somatic symptoms and depression were mentioned most frequently as preferred symptom to feign. Almost 10% of the sample indicated that they would certainly not feign memory deficits. In study 2, participants were asked to read a case-vignette and subsequently perform a short assessment in which the Medical Symptom Validity Test (MSVT; a memory-oriented PVT) and the SIMS, a broad index of symptom over-reporting, were administered. The participants were free in their choice of feigned symptomatology. The SIMS correctly identified 91% of the experimental malingerers, whereas the MSVT only identified 45% correct. In study 3, both tests were administered in a small sample of 21 psychiatric patients. Only one of the three parameters of the MSVT was correlated to the score on the SIMS. Our findings indicate that that memory-oriented PVTs like the MSVT may be useful to detect feigned memory-deficits, but perform less well as a screening tool for other types of feigned psychopathology. The clinical implication is that validity tests aimed at the specific symptoms that a patient reports need to be administered.

**Chapter 5** investigates clinicians' ability to predict distorted symptom presentation. This is an important topic since validity tests are not per default incorporated in test batteries. This study is part of the larger research project described in Chapter 3. In total, 31 neuropsychologists predicted - after their interview but before testing started - the outcome of subsequent validity testing. In this study we focused on the neuropsychological assessments, because in these assessments two validity tests - the ASTM and the SIMS - were administered. We matched the prediction of the clinician (i.e., unproblematic versus somewhat problematic) against the cases that either pas-

sed or failed both tests. Only in 76% of the cases the clinical prediction agreed with the validity test outcome. Of the 152 patients for whom neuropsychologists had predicted non-distorted symptom presentations, 14 patients (9.2%) failed both validity tests. Of the 51 patients for whom neuropsychologists had predicted at least somewhat problematic validity, 35 patients (68.6%) passed both tests. These findings support the notion that it is better to routinely test for distorted symptom presentation. Validity tests have incremental value in that they may correct initial clinical judgment.

**Chapter 6** focuses on the practices and beliefs of neuropsychologists with respect to symptom and performance validity. We conducted a survey among neuropsychologists ( $N = 515$ ) from six Western European countries: Germany, the Netherlands, Italy, Norway, Finland, and Denmark. We queried the respondents about the tools they used to evaluate symptom credibility in clinical and forensic assessments and other issues related to response validity testing. The majority of the respondents demonstrated technical knowledge about validity testing. However, only a minority of the respondents systematically based their validity judgment on validity tests (i.e., in the total sample, only 45% always included a validity test in forensic assessments, and only 12% did so in clinical assessments). There were notable differences between the countries. For example, in the Netherlands 70% of the respondents always included a validity tests in forensic assessments, whereas in Finland only 14% of the respondents did so. Also, a sizeable minority of the respondents relied on outdated notions (e.g., the idea that clinicians can determine the validity of obtained data based on intuitive judgment). There is little consensus among neuropsychologists on how to instruct patients when they are administered validity tests and how to handle test failure. Our findings indicate that the issues regarding how to administer and communicate the validity test results to patients warrants systematic research.

**Chapter 7** provides a general discussion of the results of the studies described in this thesis. The clinical implications are discussed, and recommendations for future research are made. A key implication is that validity of diagnostic data in clinical assessments should be standardly tested. The first order inference when validity measures are failed, is that the obtained diagnostic data cannot be interpreted as these data do not validly reflect a person's cognitive abilities and experienced suffering. For the future, it is important to develop response validity measures for physical symptoms, such as pain and fatigue. Ideally, in the future all self-report questionnaires and cognitive tests will have built-in validity measures.

However, besides development of new detection measures and strategies, more fundamental research on the concept of symptom validity is needed. The dynamics operating behind symptom validity are poorly understood. It is notoriously difficult to distinguish between fabricated and genuinely experienced symptoms that cannot be fully medically explained. Several lines of research on deception have shown that persons who start with simulating may gradually become to believe in their feigned symptoms.

The fluidity between feigned and genuine symptoms implies that poor symptom validity is not mere a condition that needs to be ruled-out before psychological data can be interpreted. It is a phenomenon that warrants intervention of clinicians, because of its potential to escalate into genuinely experienced symptoms. More studies addressing the practical issues of response validity in clinical assessments are needed, such as studies on interventions aimed at preventing that patients engage in symptom magnification and studies on how to best provide feedback.

## **Samenvatting**

Het onderzoek beschreven in dit proefschrift richt zich op de validiteit van (neuro)psychologische onderzoeken van patiënten met uiteenlopende aandoeningen, gezien binnen een ziekenhuissetting of GGZ instelling.

Psychologisch onderzoek wordt verricht om een beter beeld te krijgen van de symptomen en functioneren om tot een onderbouwde diagnose en een behandeladvies te komen. Voor het psychologisch onderzoek is het noodzakelijk dat patiënten goed meewerken. Patiënten moeten zich inspannen om zo goed mogelijk te presteren bij de cognitieve tests en de vragenlijsten eerlijk en accuraat invullen. Echter niet iedere patiënt zal dit doen. Er zijn tal van vertekeningen mogelijk. Patiënten kunnen onwillig zijn en zich daarom niet inspannen om een lijst woorden te onthouden. Patiënten kunnen ziekte willen aantonen om zich te verzekeren van medische aandacht of om een bepaalde voordelen te behalen. Ook omgekeerd kan een patiënt er belang bij hebben om zich zo gezond mogelijk te presenteren en psychologische symptomen en beperkingen in het functioneren te minimaliseren. Hierbij kan bijvoorbeeld gedacht worden aan een psychologische screening die mede bepaalt of iemand in aanmerking komt voor een bepaalde medische behandeling. Kortom, er zijn tal van redenen waarom de patiënt zich anders kan voordoen dan hij is. Wanneer een dergelijke vertekening aan de orde is, dan volgt hier logischerwijs uit dat de verkregen gegevens – de zelf-gerapporteerde klachten, de levensgeschiedenis en de scores op de cognitieve testen - inboeten aan geldigheid. De vraag naar de geldigheid van gerapporteerde klachten of voldoende inzet bij testonderzoek, staat ook wel bekend als symptoomvaliditeit. Om symptoomvaliditeit te meten, bestaan allerlei methodieken. Er zijn op zichzelf staande testen en ingebouwde indicatoren binnen gangbare cognitieve tests en psychologische vragenlijsten. Deze methodieken richten zich op het vaststellen of er sprake is van onderpresteren bij cognitieve testen en van over- of onderrapporteren van psychologische symptomen. Met onderpresteren wordt bedoeld dat iemand onder zijn feitelijk niveau van kunnen presteert.

De vraag naar symptoomvaliditeit bij psychologisch onderzoek en hoe de diagnosticus dit objectief kan bepalen, werd aanvankelijk vooral gesteld binnen het forensische domein. Bij gerechtelijke procedures kan het hebben van klachten of ziektes betekenen dat de onderzochte een veroordeling ontloopt, arbeidsongeschikt wordt verklaard of dat een aanzienlijke letselschadebedrag wordt toegekend. De vraag naar symptoomvaliditeit bij de psychologische onderzoeken in de reguliere klinische zorg is - waar dergelijke directe materiele of strafrechtelijke belangen een minder directe rol spelen – is echter nog weinig onderzocht.

De centrale doelstelling van het onderzoek in dit proefschrift is het kaart brengen van hoe vaak verminderde symptoomvaliditeit voorkomt binnen de patiëntzorg, met welke factoren verminderde symptoomvaliditeit samenhangt, of symptoomvaliditeit een multidimensioneel verschijnsel is en hoe neuropsychologen in hun dagelijks handelen met symptoomvaliditeit omgaan. Deze doelen zijn geadresseerd in de vijf studies van dit proefschrift.

**Hoofdstuk 1** geeft een algemene inleiding op het onderwerp en de doelstellingen van dit proefschrift. In dit hoofdstuk komen de definitie van het begrip symptoomvaliditeit, methoden om symptoomvaliditeit te meten, en de overlap en afbakening met het begrip malingering aan bod. Ook worden de onderzoeksvragen besproken.

**Hoofdstuk 2** beschrijft een cross-sectionele studie naar symptoomvaliditeit in een steekproef van 183 patiënten met verschillende psychiatrische aandoeningen gezien binnen een GGZ instelling. Alle patiënten waren voor behandeling verwezen voor neuropsychologisch onderzoek. Bij deze onderzoeken is een validiteitstest gericht op het meten van onderpresteren bij cognitieve testen en validiteitstest gericht op het overrapporteren van klachten afgenomen. In totaal behaalde 13% van de patiënten een score die wijst op onderpresteren of op overrapporteren. Door 8% van de patiënten werd op beide validiteitstesten afwijkend gescoord. Onderpresteren en overrapporteren van klachten bleken slechts matig samen te hangen. Er was een sterk, positief verband tussen overrapporteren van klachten en de score op een lijst voor psychische klachten. De subgroep die enkel de test voor onderpresteren slecht maakte, presteerde slechter op een indexscore voor geheugen. De resultaten van deze studie laten zien dat verminderde symptoomvaliditeit regelmatig voorkomt. Het onderpresteren en overrapporteren van klachten hangt slechts matig met elkaar samen en het overrapporteren is het sterkst gekoppeld aan verhoogde scores op de psychologische klachtenlijst.

In **hoofdstuk 3** wordt een onderzoek gepresenteerd met dezelfde opzet als in hoofdstuk 2 maar uitgevoerd in een ziekenhuissetting. De afdelingen Psychologie van vijf ziekenhuizen participeerden in de studie. De patiënten waren door de medisch specialisten verwezen voor diagnostiek. Patiënten met evidente cognitieve stoornissen uitgezonderd, namen psychologen bij ieder psychologisch onderzoek de SIMS af, een validiteitstest voor het overrapporteren van symptomen. Bij de neuropsychologische onderzoeken namen psychologen ook de AKTG af, een onderpresteertaak. In totaal zijn de data van 469 psychologische onderzoeken geanalyseerd. Hierbij zijn ook achtergrondgegevens van de patiënten, zoals leeftijd, opleiding, diagnose bij aanmelding, en lopende juridische procedures meegenomen. Er waren 5 diagnostische hoofdgroepen: neurologische aandoeningen, morbide obesitas, somatisch onvoldoende verklaarde klachten (SOLK), psychiatrische aandoeningen, cognitieve klachten niet nader gespecificeerd en medische aandoeningen met cognitieve klachten. Het overrapporteren van klachten kwam voor bij 12%-19% van de diagnostische hoofdgroepen. Een uitzondering hierop vormden de groep patiënten met morbide obesitas. Bij deze groep diende het psychologisch onderzoek om de geschiktheid te beoordelen voor een maagverkleiningsoperatie. In deze groep scoorde slechts 1 patiënt afwijkend hoog op de SIMS. De score op de SIMS hield positief verband met de score op een zelfrapportage lijst voor psychische klachten. Bij 29.3% van de neuropsychologische onderzoeken werd afwijkend laag gescoord op de AKTG. De score op de AKTG hing negatief samen met de geheugenprestatie. In onze studie werd geen verband gevonden tussen ver-

minderde symptoomvaliditeit en de aanwezigheid van financiële belangen, zoals het ontvangen van een uitkering of het verwickeld zijn in een juridische procedure. Deze bevindingen onderschrijven het belang van het meten van symptoomvaliditeit bij regulier klinische psychologische onderzoeken. Meer onderzoek is nodig om te begrijpen welke factoren een rol spelen bij verminderde symptoomvaliditeit.

**Hoofdstuk 4** beschrijft drie verschillende studies met als centrale vraag of personen die geen valide beeld tonen van hun problematiek zowel onderpresteren bij testonderzoek als overrapporteren van klachten. De vraag is of beide gedragingen – onderpresteren en overrapporteren – behoren tot één dimensie of dat het twee aparte dimensies zijn van het overstijgende concept symptoomvaliditeit. Deze vraag is belangrijk omdat de meeste validiteitsmaten zich richten op onderpresteren bij geheugentaken en ook worden toegepast om symptoomvaliditeit te bepalen bij aandoeningen waar de geheugenklachten niet primair zijn, zoals bijvoorbeeld bij patiënten met een somatisch onvoldoende verklaarbare klachten.

In de eerste studie werden studenten aan de hand van een gevalsbeschrijving gevraagd welke symptomen ze wel en welke ze zeker niet zouden veinzen. Aspecifieke lichamelijke symptomen en depressie werden het meest frequent genoemd als symptomen om te veinzen. Bijna 10% van de deelnemers rapporteerde zeker geen geheugenproblemen te veinzen. Bij de tweede experimentele studie, vroegen we studenten om zich in te leven in een gevalsbeschrijving en vervolgens te veinzen bij psychologisch onderzoek, waarbij de MSVT, een maat voor onderpresteren bij geheugen, en de SIMS werden afgenomen. De deelnemers werden vrijgelaten in de keuze van welke aandoening ze wilde veinzen. Onder deze omstandigheden werd met de SIMS 91% van de experimentele simulanten correct geïdentificeerd. De MSVT classificeerde slechts 45% correct. In de derde studie werden beide testen afgenomen bij een kleine steekproef van 21 psychiatrische patiënten. Slechts één van de drie subtesten van de MSVT bleek samen te hangen met de score op de SIMS. Tezamen genomen laten deze bevindingen zien dat geheugen georiënteerde validiteitstesten zinvol kunnen zijn om geveinsde geheugenstoornissen op te sporen, maar dat ze minder geschikt zijn om andere vormen van geveinsde psychopathologie vast te stellen. De bevindingen van de drie studies wijzen op een meerdimensionaal concept van symptoomvaliditeit. De klinische implicatie is dat klinici symptoomvaliditeitstests moeten gebruiken die betrekking hebben op de specifieke klachten die de patiënt rapporteert.

**Hoofdstuk 5** onderzoekt het vermogen van klinici om van tevoren in te schatten welke patiënt al dan niet een valide beeld zal geven van zijn functioneren bij neuropsychologisch onderzoek. Deze vraag is belangrijk omdat klinici niet standaard validiteitstesten afnemen bij psychologische onderzoeken in de reguliere zorg. Doorgaans maakt de clinicus een inschatting of validiteitstesten aan de testbatterij toegevoegd moeten worden of niet. De vraag is of dit een verstandige werkwijze is. Bij de in hoofdstuk 3 beschreven studie in 5 ziekenhuizen, hebben klinici na de intake en voor aanvang van het testonderzoek, een voorspelling gedaan over hoe de patiënt zou sco-

ren op de symptoomvaliditeitstests (voldoende – matig – onvoldoende). Voor deze studie hebben we ons beperkt tot de neuropsychologische onderzoeken, omdat bij deze onderzoeken twee symptoomvaliditeitstesten – de AKTG en de SIMS – zijn afgenomen. De inschatting van de klinici (voldoende versus matig / onvoldoende) is afgezet tegen de uitkomst van de twee validiteitstesten (normale scores op beide taken versus afwijkende scores op beide taken). In totaal beoordeelden 31 klinici de uitkomst van validiteitstesten van 203 neuropsychologische onderzoeken. De uitkomsten laten zien dat de aanname van veel psychologen dat de 'klinische blik' voldoende is om het overrapporteren van klachten en symptomen te signaleren niet houdbaar is. Slechts in 76% van de gevallen kwam de uitkomst van de klinici overeen met de uitkomst van de validiteitstesten. Van de 152 patiënten bij wie de klinici vooraf een niet afwijkende symptoomvaliditeit hadden ingeschat, scoorden 14 patiënten (9%) afwijkend op beide testen. Omgekeerd scoorden 35 van de 51 patiënten (69%) bij wie de klinici verminderde symptoomvaliditeit hadden voorspeld, normaal bij beide testen. Het standaard met objectieve maten beoordelen van de validiteit verdient daarom de voorkeur boven het per casus klinisch beslissen om deze objectieve maten al dan niet te gebruiken.

**Hoofdstuk 6** richt zich op hoe neuropsychologen denken over symptoomvaliditeit en hoe ze met dit onderwerp omgaan in hun dagelijkse praktijk. Om dit te onderzoeken hebben een enquête uitgezet onder neuropsychologen in zes West-Europese landen: Duitsland, Nederland, Italië, Noorwegen, Finland, en Denemarken. In totaal vulden 515 neuropsychologen de vragenlijst in. De meerderheid van de neuropsychologen had technische kennis over symptoomvaliditeit. Echter klinici gebruiken SVTs nog relatief weinig: 45% van de respondenten gebruikt een SVT in vrijwel ieder forensisch onderzoek en 12% van de respondenten gebruikt SVTs in vrijwel ieder klinisch psychologisch onderzoek. Hierin waren er wel verschillen tussen de verschillende landen. Zo stelde in Nederland 70% van de respondenten bij vrijwel ieder forensisch onderzoek een SVT te gebruiken, in Finland slechts 14% van de respondenten. Ook bleek een minderheid nog verouderde opvattingen over symptoomvaliditeit erop na te houden, zoals het idee dat klinici de validiteit kunnen vaststellen met hun klinische blik. Verder bleek consensus te ontbreken over hoe klinici patiënten vooraf moeten instrueren over SVTs en over hoe klinici dienen te handelen wanneer een patiënt een afwijkende score op een SVT behaalt. Ook lieten de resultaten zien dat er nog geen consensus bestaat over hoe de uitkomsten het beste gecommuniceerd kunnen worden. Deze kwesties aangaande het toepassen van SVTs bij het testonderzoek en de communicatie van afwijkende uitslagen verder onderzoek.

In **hoofdstuk 7** worden de resultaten bediscussieerd. Het belang van standaard meten van symptoomvaliditeit bij klinische onderzoeken wordt benadrukt. Een belangrijke klinische implicatie van de bevindingen in dit proefschrift is dat validiteit formeel beoordeeld moet worden in het diagnostisch onderzoek. Zoals het meerwaarde heeft om tal van psychische symptomen en cognitieve functies te meten in plaats van te

volstaan met een indruk, zo heeft het meerwaarde om de symptoomvaliditeit te meten. Voor toekomstig onderzoek is het belangrijk dat er validiteitsmaten worden ontwikkeld voor lichamelijke klachten, zoals pijn en vermoeidheid. Idealiter zal in de toekomst iedere test en vragenlijst ingebouwde validiteitsindicatoren hebben. Naast de ontwikkeling van nieuwe instrumenten, is meer fundamenteel onderzoek naar het concept van symptoomvaliditeit nodig. Het is notoir moeilijk om onderscheid te maken tussen gefabriceerde symptomen en oprecht ervaren symptomen die niet medisch verklaard kunnen worden. Verschillende lijnen van onderzoek hebben laten zien dat er geen scherp onderscheid bestaat tussen beide condities. Personen die beginnen met het voorwenden van klachten kunnen langzamerhand ook echt zelf gaan geloven in de klachten. Het is alsof de persoon gaandeweg het doe-alsof karakter van de klachten uit het oog verliest. Dit laat zien dat verminderde symptoomvaliditeit meer behelst dan enkel vertekening van het psychologisch onderzoek. Het is niet alleen een verschijnsel dat klinici moeten uitsluiten voordat ze overige testuitslagen kunnen interpreteren. Het is een verschijnsel dat een interventie van klinici behoeft omdat het potentieel kan leiden tot een toename van ervaren klachten. Of anders gezegd dat iemand zichzelf steeds meer gaat ervaren als een patiënt met beperkingen. Meer onderzoek is nodig om te begrijpen hoe het opschuiven van voorwenden naar zelf ervaren van klachten werkt en welke interventies dit kunnen couperen. Terugkoppeling van verminderde symptoomvaliditeit aan de patiënt is een interventie die wetenschappelijke aandacht verdient.