

## Op een symptoomvaliditeitstest presteert de patiënt normaal: Wat nu?

- ▶ Harald Merckelbach
- ▶ Ingrid Rooijackers
- ▶ Brechje Dandachi-FitzGerald
- ▶ Isabella Niesten

- **Samenvatting** — Stel een patiënt beantwoordt aan het DSM-5-profiel van simuleren, maar presteert vervolgens normaal op symptoomvaliditeitstests (SVT's). Zijn clinici dan bereid om hun verdenking af te zwakken? We onderzochten dat door experts ( $N = 42$ ) een fictieve gevalbeschrijving van een patiënt voor te leggen. Daarin werden toespelingen gemaakt op het DSM-5-profiel van simuleren. Vervolgens ontvingen de experts in stappen SVT-resultaten die het idee van simuleren tegenspreken. Telkens kruisten ze aan hoe groot ze de kans achtten dat de patiënt simuleerde. Experts bleken deze kans niet naar beneden bij te stellen. De eenmaal ontstane indruk dat de patiënt klachten overdrijft, vat snel post, maar is moeilijk te corrigeren.

### Inleiding

Wanneer moet de clinicus er ernstig rekening mee houden dat een patiënt klachten simuleert? De DSM-5 (American Psychiatric Association, 2013) zegt dat zoiets verstandig is als zich een combinatie van de volgende kenmerken voordoet: de patiënt werkt niet goed mee, heeft antisociale trekken, is verwickeld in een juridische procedure en zijn klachten contrasteren met wat andere bronnen daarover zeggen. Maar vanwege de geringe afbakening is dit profiel op grote groepen patiënten van toepassing (Berry & Nelson, 2010; Niesten e.a., 2015). Zo bezit iets meer dan de helft van alle forensische patiënten ten minste twee van deze kenmerken, terwijl de prevalentie van klachtenoverdrijving in die groep zelden zo'n omvang bereikt (Van Impelen e.a., 2016). Daarom: als clinici niet ook andere informatie meenemen in hun diagnostische afwegingen, zullen ze met het DSM-profiel in de hand op ruime schaal eerlijke patiënten

voor simulanten gaan houden, dat wil zeggen, foutpositieve beslissingen nemen (Rogers, 1990).

Om vast te stellen of patiënten simuleren, zijn klinici gebaat bij symptoomvaliditeitstests (SVT's). Het gaat daarbij om twee typen tests. Het ene type bestaat uit vragenlijsten waarmee valt na te gaan of de patiënt overmatig veel symptomen rapporteert. Een voorbeeld is de Structured Inventory of Malingered Symptomatology (SIMS; zie voor een recent overzicht: Van Impelen e.a., 2014). De SIMS somt een groot aantal atypische klachten op. Wie er meer dan zestien aankruist, laadt de verdenking op zich symptomen te overdrijven. In een onderzoek met experimentele simulanten en eerlijke proefpersonen bleek het afkappunt van zestien een behoorlijke sensitiviteit (meer dan 90% van de simulanten werd geïdentificeerd) en een relatief laag percentage foutpositieven (minder dan 10%) op te leveren (Van Impelen e.a., 2014).

Het andere type SVT bestaat uit redeneer- of geheugentaken, waarop zelfs jonge kinderen en patiënten met een hersenbeschadiging redelijk goed presteren. Een voorbeeld is de Amsterdamse Korte-Termijn Geheugentaak (AKTG; Schagen e.a. 1997; Schmand e.a., 1999), die uit dertig onderdelen bestaat. Elk onderdeel omvat drie stappen. Eerst ziet de patiënt vijf woorden, zoals boom, auto, fiets, step en ijskast. Vervolgens moet er een simpele rekensom worden gemaakt, zoals  $3 \times 3 = ?$  En daarna krijgt de patiënt weer vijf woorden, bijvoorbeeld paard, auto, fiets, step en roos, waarvan er drie ook al tijdens de eerste stap opdoken. De opdracht aan de patiënt is om de drie terugkerende woorden aan te wijzen. De maximale score op deze geheugentest is zodoende  $30$  (onderdelen)  $\times 3$  (woorden) =  $90$ . Wie onder de  $85$  scoort, zou er wel eens opuit kunnen zijn om moedwillig slecht te presteren. Bij het afkappunt van  $85$  horen hoge percentages van simulanten die door mand vallen (circa 90%) en een relatief bescheiden percentage fout-positieven ( $< 12\%$ ). Bij het conservatievere afkappunt van  $82$  daalt het percentage fout-positieven zelfs naar de  $2\%$ .

Instrumenten zoals de SIMS en de AKTG sonderen naar klachtenoverdrijving en onderprestatie, maar de oorzaak daarvan kan velerlei zijn. Moedwillige symptoomfabricatie is slechts één mogelijkheid; het kan ook zijn dat iemand beroerd presteert, omdat hij bijvoorbeeld de tests lukraak invult of de vragen niet begrijpt (Merten & Merckelbach, 2013). Niettemin hebben SVT's eerst en vooral de reputatie dat je er simulanten mee kunt opsporen. Op deze manier worden ze ook door experts gewaardeerd. Zo legden Trueblood en Binder (1997) neuropsychologen gevalsbeschrijvingen voor van simulanten en patiënten met hersenletsel. De beschrijvingen omvatten ook de afwijkende uitslagen van simulanten op

SVT's. Experts betoonden zich gevoelig voor deze informatie en gebruikten haar om simulanten te identificeren.

Zo'n nadruk op identificatie oftewel de sensitiviteit waarmee simulanten te detecteren zijn, kan iets aan het zicht onttrekken wat minstens zo belangrijk is: de ontkrachtende bewijswaarde van normale prestaties op SVT's. Wanneer een patiënt normaal scoort op een SVT – bijvoorbeeld onder de 16 op de SIMS en/of boven de 82 op de AKTG – pleit dat voor de aanname dat de patiënt zijn klachten op een eerlijke manier presenteert. In technische zin spreken we dan over de negatief voorspellende waarde. Neem de SIMS: deze test heeft een negatief voorspellende waarde die de 0,85 vaak overstijgt, wat betekent dat de kans dat iemand *niet* simuleert 85% of meer is indien zijn SIMS-score normaal is (< 16; Van Impelen e.a. 2014).<sup>1</sup> Vergelijkbare overwegingen gelden voor de AKTG: uitgaande van de eerder genoemde sensitiviteit en specificiteit (en een basisfrequentie van 30%), bereikt de negatief voorspellende waarde makkelijk de 90%. Ons punt is dat niet alleen afwijkende, maar ook normale scores op SVT's het verdienen om te worden verdisconteerd in het uiteindelijke diagnostische oordeel van experts. De vraag is of dat wel gebeurt.

Wat daaraan in de weg zou kunnen staan, is de ongunstige bijklank van begrippen als simuleren, veinzen, malingering et cetera. De DSM-5 en zijn voorgangers dragen aan die negatieve connotaties bij. Want afgezien van de ruime typering die zij van simuleren geven, roept hun typering associaties op met dubieuze karaktertrekken (Rogers, 2008). Dat kan aanzetten tot een vorm van diagnostische tunnelvisie (Oskamp, 1965; Wedding & Faust, 1989), waarbij bedenkelijke aanwijzingen – bijvoorbeeld: de patiënt is ooit in aanraking geweest met justitie – zodanig als ankers gaan fungeren dat ontlastende evidentie – normale SVT-scores – niet meer worden meegewogen. We onderzochten of experts inderdaad naar zo'n tunnelvisie neigen. Daarbij hanteerden we een benadering die we ontleenden aan Oskamp (1965). Experts kregen trapsgewijs de beschikking over diagnostische informatie en we keken hoe hun diagnostisch oordeel zich gaandeweg ontwikkelde. Oskamp vond overigens in een steekproef van psychodiagnostische experts dat trapsgewijze informatie experts zelfverzekerder, maar niet accurater maakt. Deze wankele

---

<sup>1</sup> De negatief voorspellende waarde van een test is niet alleen afhankelijk van haar sensitiviteit en specificiteit, maar ook van de basisfrequentie waarmee het fenomeen dat de test wil detecteren zich in de populatie voordoet. Zo is bij een sensitiviteit van 77%, een specificiteit van 90% en een basisfrequentie van 30% de negatief voorspellende waarde 0,90. De negatief voorspellende waarde neemt toe als de basisfrequentie omlaaggaat. Zie Rosenfeld e.a. (2000).

relatie tussen accuraatheid en zelfverzekerdheid is overigens wel vaker beschreven in de literatuur over neuropsychologische diagnostiek (Garb & Schramke, 1996).

## Methodie

### Experts

Via een sneeuwbalmethode nodigden we 42 psychologen en psychiaters uit om zich te buigen over een casus (zie onder). Sommige experts waren als forensisch deskundige werkzaam (forensische groep:  $n = 19$ ; 9 vrouwen). Anderen werkten vooral als behandelaar binnen de geestelijke gezondheidszorg (klinische groep:  $n = 23$ ; 16 vrouwen). De gemiddelde leeftijd van onze experts was 39 jaar ( $SD = 11$ ) en door de bank genomen werkten ze al 12 jaar ( $SD = 10$ ) op de plek waar ze nu zaten.

### Casus

De casus werd via e-mail of op papier aangeboden. Het was een hypothetisch geval dat startte met achtergrondinformatie die toespelingen maakte op de DSM-typologie van simuleren.<sup>2</sup> Zo werd verteld dat: het ging om een 55-jarige patiënt, die sinds kort klaagde over migraineachtige hoofdpijnen, maar ook zei al langere tijd last te hebben van intrusies verband houdend met een traumatisch voorval; dat betrokkene sinds 1995 illegaal in ons land verbleef, zodoende het Nederlands goed machtig was, maar dat nu uitzetting dreigde; dat de patiënt aantekeningen had in het justitieel documentatieregister wegens onder meer openbare dronkenschap; en dat zijn huisarts er bij hem op had aangedrongen om een neuroloog te consulteren, maar dat de patiënt dat advies niet had opgevolgd. De verwijsvraag luidde: 'Hoe valide zijn de symptomen van deze patiënt? Om daarachter te komen praat u met de patiënt en neemt u een aantal tests bij hem af.'

Nadat de experts dit allemaal hadden gelezen, beoordeelden ze (1) hoe realistisch ze de casus vonden, (2) of ze in dit stadium dachten met een simulant van doen te hebben en (3) hoeveel vertrouwen ze in hun oordeel stelden. De experts gebruikten daarvoor schalen die liepen van 0 ('helemaal niet realistisch; de kans is nihil dat dit een simulant is; ik ben totaal niet zeker van mijn oordeel') tot 10 ('zeer realistische casus; de kans is zeer groot dat dit een simulant is; ik ben zeer zeker van mijn oordeel').

---

2 De volledige casus is verkrijgbaar bij de eerste auteur.

### Trapsgewijze informatie

Daarna kregen de experts in vijf rondes nieuwe informatie. In ronde 1 (SIMS = 14) was dat een korte toelichting op de SIMS en de mededeling dat de patiënt een score van 14 haalde, dus aan de veilige ('normale') kant van het afkappunt zat. In ronde 2 (Hobby) volgde betrekkelijk neutrale informatie over de hobby's van de patiënt (met de honden wandelen). In ronde 3 (AKTG = 87) werd eerst een korte schets van de AKTG gegeven, waarna de mededeling volgde dat de patiënt een geruststellende score op deze test haalde (namelijk 87). In ronde 4 (Interview) werden summier enkele interviewgegevens besproken: de patiënt rapporteerde dat de klachten waren verergerd sinds er de dreiging van uitzetting was, hij zei bang te zijn voor een tumor in zijn hoofd en beschreef dat de klachten zo'n één à twee keer per week optraden en dan een hele dag aanhielden. In ronde 5 (Psychometrie) werden kort enkele psychometrische details ter sprake gebracht. Het ging om de scores van de patiënt op veelgebruikte instrumenten om klachten te inventariseren. Het kwam erop neer dat de patiënt veel klachten rapporteerde op de SCL-90 en een lijst met PTSS-symptomen.

Na elke ronde vulden de experts steeds de 11-puntschalen in over hoe groot zij de kans (K) achtten dat de patiënt een simulant was en hoe zeker (Z) ze van dit oordeel waren (zie boven). Bovendien gaven ze telkens antwoord op de vraag of ze de verdenking van simulatie zouden noemen in hun eindrapport aan de verwijzer.

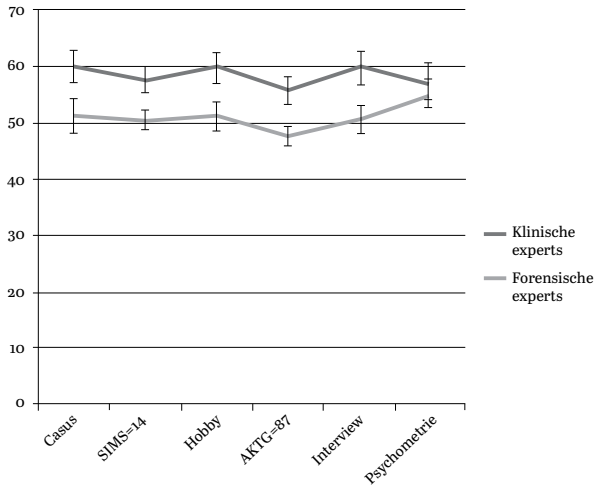
### Analyse

Om het aantal toetsen beperkt te houden, middelden we kans- en zekerheidsoordelen [ $((K+Z)/2) \times 10$ ] en het was deze score die we aan een 2 (groepen)  $\times$  6 (achtergrondinformatie; rondes 1-5) -variantieanalyse met herhaalde metingen op de laatste factor onderwierpen. Met  $\chi^2$ -toetsen keken we naar het percentage experts dat simuleren als verdenking zou noemen in hun eindrapport en hoe dat percentage fluctueerde over de rondes.

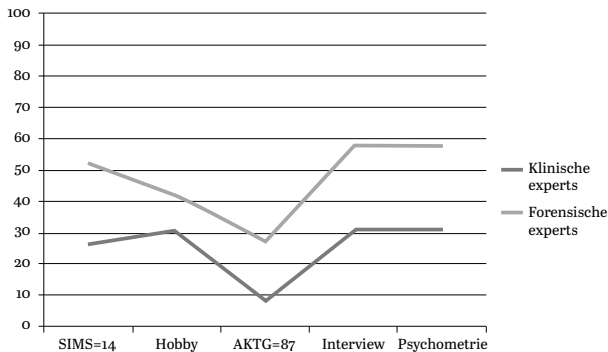
### Resultaten

Beide groepen vonden de casus realistisch – forensische experts: 8,1 (SD = 0,9); klinische experts: 7,6 (SD = 1,1) – en er waren wat dat betreft geen verschillen tussen de groepen [ $t(40) = 1,6, p = 0,10$ ]. Figuur 1 laat zien hoe de verdenking op simuleren [ $= ((K + Z)/2) \times 10$ ] over de verschillende rondes verliep. Er zijn drie opvallende aspecten aan dit patroon. Om te beginnen was de achterdocht over de gehele linie sterker bij de forensische

dan bij de klinische experts: 58 (SD = 2,1) versus 51 (SD = 1,9),  $F(1, 38) = 6,4$ ,  $p < 0,02$ ,  $\text{partial } \eta^2 = 0,14$ . Ten tweede: de verdenking op simuleren fluctueerde nauwelijks onder invloed van de trapsgewijs aangeboden informatie –  $F(5, 190) = 1,5$ ,  $p = 0,14$  – en bleef meestal iets boven het middelpunt van de schaal (50) hangen. De experts waren dus amper gevoelig voor de informatiewaarde van de SVT-resultaten. En ten derde: forensi-



**FIGUUR 1** Verloop van verdenking op simuleren na trapsgewijze onthulling van informatie bij forensische ( $n = 19$ ) en klinische ( $n = 23$ ) experts. Y-as: het gemiddelde van kans op simuleren en zekerheid in oordeel  $\times 10$



**FIGUUR 2** Percentage experts dat overweegt om simuleren als diagnostische optie te noemen in een eindrapport

sche en klinische experts verschilden daarin niet en betoonden zich even onverschillig voor deze informatie:  $F(5, 190) = 1,1, p = 0,36$ .

Figuur 2 toont de percentages van experts dat in een eindrapport de verdenking van simuleren zou noemen. Dit keer waren er geen overtuigende verschillen tussen forensische en klinische experts (alle  $\chi^2$ 's (1) < 3,2, alle  $p$ 's > 0,07). Het meest in het oog springende aspect van Figuur 2 is dat aan het begin en aan het einde van de rit het percentage experts dat melding zou maken van simulatie grofweg hetzelfde was.

## Discussie

Toegegeven: ons panel van experts was klein, de opzet exploratief en de gevalsbeschrijving summier. De experts kregen louter schriftelijke informatie en ze waren niet in de gelegenheid om hun eigen vragen te stellen en hun eigen tests in te zetten. Bovendien lieten we experts zich uitspreken over het geval van slechts één patiënt en de trapsgewijze onthulling van test- en interviewgegevens was niet *counterbalanced*. Onze resultaten zijn onder een te kunstmatige conditie verkregen om er ferme uitspraken mee te doen over de zorgvuldigheid van diagnostische beslissingen.

Dat gezegd hebbend, menen we wel dat ons onderzoek licht werpt op een kwestie waarover nooit eerder is gepubliceerd: hoe wegen experts normale prestaties op SVT's? Zijn ze bereid om de ontkrachtende waarde daarvan te verdisconteren in hun oordeel? De hier gepresenteerde gegevens doen vermoeden van niet. Ze laten zien dat normale prestaties op SVT's nauwelijks de verdenking van simuleren kunnen wegnemen.

Men zou kunnen tegenwerpen dat onze experts een relatief lage verdenkingsgraad hanteerden, die zich in de buurt van het midden ophield (50). Daaruit zou diagnostische behoedzaamheid kunnen worden afgeleid. Men zou daarom kunnen zeggen dat voor experts die al voorzichtig zijn, normale uitslagen van SVT's geen toegevoegde waarde hebben. Dat zo'n sympathieke duiding minder waarschijnlijk is, blijkt wel uit Figuur 2. Daarin is immers te zien dat het percentage experts dat overwoog om de verdenking van simuleren in een eindrapport te noemen bijna voortdurend substantieel was (> 25%).

Aan deze constatering willen we een viertal overwegingen toevoegen. Ten eerste: de DSM-5 hanteert een slecht afgebakende typologie van simuleren en deze typologie heeft bovendien een pejoratieve ondertoon (Berry & Nelson, 2010; Niesten e.a., 2015). Het is bij uitstek zo'n combinatie die een vorm van tunnelvisie in gang kan zetten, waarbij een eerste

diagnostische impressie zich niet meer laat corrigeren door ontkrachtende informatie (Oskamp, 1965; Berner & Graber, 2008).

Ten tweede: de verdenking op simuleren was over de gehele linie geprononceerder bij forensische experts dan bij klinische experts. Dat is in het licht van het vorige punt niet verwonderlijk. De basisfrequentie van simuleren ligt in de forensische omgeving immers hoger dan in de reguliere gezondheidszorg (Niessen e.a., 2015) en forensische experts zullen meer bekend zijn met de literatuur – de DSM-5 voorop – over het fenomeen.

Ten derde: wat we hierboven beschreven is een variant op meer fundamenteel onderzoek dat het bestaan van een asymmetrische buigzaamheid in onze oordeelsvorming laat zien: mensen zijn soepeler in het accepteren dan in het afwijzen van een hypothese (Gilbert, 1991). Dat geldt ook voor de diagnostische hypothesen van medische en psychologische experts (Berner & Graber, 2008). In het geval van mogelijke klachtensimulatie wordt deze asymmetrie op twee manieren aangewakkerd. Allereerst wekt de DSM-5 de indruk dat simuleren een categorale aangelegenheid is: óf de (pseudo)patiënt simuleert, óf de patiënt is eerlijk. De diagnosticus staat zodoende voor de opgave om een serie van probabilistische aanwijzingen – de resultaten van tests en interviews – te wegen en te vertalen naar een categorale beslissing. Mensen – ook experts – zijn daar niet goed in en hebben vaak de neiging om bijvoorbeeld het *'beyond reasonable doubt'*-criterium naar beneden bij te stellen (Magnussen e.a., 2014). Dat kan makkelijk zo uitpakken dat experts al met geringe aanwijzingen genoeg nemen om tot de indruk te geraken dat de patiënt zijn klachten overdrijft.

Een ander punt dat deze asymmetrie aanwakkert, heeft te maken met de manier waarop degenen die SVT's ontwikkelen, praten en schrijven over hun tests. Hun meta-analyses en handleidingen besteden veel aandacht aan de sensitiviteit en de geringe foutpositieve percentages van SVT's (zie voor een illustratie: Sollman & Berry, 2011). De nadruk ligt zodoende op het opsporen van simulanten. Maar minstens zo belangrijk is de negatief voorspellende waarde van SVT's. Dit is niet de plek om bij de fitnesses van dat begrip stil te staan (zie daarvoor Rosenfeld e.a., 2000). Het komt erop neer dat de mate waarin een normale SVT-score wijst op een eerlijke symptoompresentatie bij instrumenten als de SIMS en de AKTG doorgaans rond de 0,85 cirkelt.<sup>1</sup> Indien de patiënt op twee van zulke SVT's normaal presteert, is de kans dat hij behoort tot de categorie van foutnegatieven – de patiënt is in werkelijkheid een simulant, maar wordt geïdentificeerd als eerlijke patiënt – in de orde van grootte van 2 à 3% (uitgaande van de in dit geval redelijke aanname dat de tests onafhankelijk



zijn). Zo'n patroon van testresultaten zou billijkerwijze de verdenkingsgraad moeten temperen.

### Conclusie

Onze resultaten doen vermoeden dat het die uitwerking niet heeft. Dat valt in de allereerste plaats testontwikkelaars aan te rekenen. Zij zouden de betekenis van de negatief voorspellende waarde meer moeten benadrukken in hun handleidingen en meta-analyses. De kwestie is belangrijk. Zo vonden Dandachi-FitzGerald e.a. (2016) dat bij de groep van ziekenhuispatiënten die wordt verwezen naar de neuropsycholoog zich de constellatie van aanvankelijke verdenking op symptoomoverdrijving en vervolgens normale scores op SVT's heel regelmatig voordoet. Schrijven deze neuropsychologen uiteindelijk in hun rapport dat met het oog op de testuitslagen symptoomoverdrijving onwaarschijnlijk is? Dat zou de enige juiste conclusie zijn, maar op grond van de hierboven geschetste resultaten sluiten we niet uit dat ze zullen vasthouden – al is het omfloerst – aan hun eerste impressie. Ook hier geldt wat the Schotse empiricus Alexander Bain (1818-1903) al weer enige tijd geleden schreef: 'The great master fallacy of the human mind is believing too much.'

**Harald Merckelbach** Sectie Forensische Psychologie, Faculteit der Psychologie en Neurowetenschappen, Universiteit Maastricht,

**Correspondentieadres:** Prof. dr. H. Merckelbach, Universiteit Maastricht, Faculteit der Psychologie en Neurowetenschappen, Postbus 616, 6200 MD, Maastricht, h.merckelbach@maastrichtuniversity.nl.

**Ingrid Rooijackers** PsyQ Maastricht.

**Brechje Dandachi-FitzGerald** Afdeling Psychiatrie, MUMC+/Universiteit Maastricht.

**Isabella Niesten** sectie Forensische Psychologie, Faculteit der Psychologie en Neurowetenschappen, Universiteit Maastricht.

### Literatuur

- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders. (5th. ed.)*. Arlington, VA: American Psychiatric Association Publishing.
- Berner, E.S. & Graber, M.L. (2008). Overconfidence as a cause of diagnostic error in medicine. *American Journal of Medicine*, 121, 2-23.
- Berry, D.T. & Nelson, N.W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury and Law*, 3, 295-303.
- Dandachi-FitzGerald, B., Merckelbach, H. & Ponds, R.W.H.M. (2016). Neuropsychologists' ability to predict symptom validity. *Journal of Clinical and Experimental Neuropsychology*, 7-9-2016, DOI: 10.1080/13803395.2016.1223278.
- Garb, H.N. & Schramke, C.J. (1996). Judgment research and neuropsychological

- assessment: A narrative review and meta-analysis. *Psychological Bulletin*, 120, 140-153.
- Gilbert, D.T. (1991). How mental systems believe. *American Psychologist*, 46, 107-119.
- Impelen, A. van, Merckelbach, H., Jelicic, M. & Merten T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *Clinical Neuropsychologist*, 28, 1336-1365.
- Impelen, A. van, Merckelbach, H., Niesten, I., Huhnt, B. & Jelicic, M. (2016). Biased symptom reporting and antisocial behaviour. In voorbereiding.
- Magnussen, S., Eilertsen, D.E., Teigen, K.H. & Wessel, E. (2014). The probability of guilt in criminal cases: Are people aware of being beyond reasonable doubt? *Applied Cognitive Psychology*, 28, 196-203.
- Merten, T. & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law*, 6, 122-137.
- Niesten, I.J.M., Nentjes, L., Merckelbach, H. & Bernstein D.P. (2015). Antisocial features and faking bad: A critical review. *International Journal of Law and Psychiatry*, 41, 34-42.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- Rogers, R. (1990). Development of a new classificatory model of malingering. *Bulletin of the American Academy of Psychiatry and the Law*, 18, 32-333.
- Rogers R. (2008). An introduction to response styles. In: R. Rogers (red.), *Clinical assessment of malingering and deception* (pp. 3-13). New York: Guilford.
- 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.
- Schagen, S., Schmand, B., De Sterke, S. & Lindeboom, J. (1997). Amsterdam Short-Term Memory Test: A new procedure for the detection of feigned memory deficits. *Journal of Clinical and Experimental Neuropsychology*, 19, 43-51.
- Schmand, B., De Sterke, S. & Lindeboom, J. (1999). *Amsterdamse Korte Termijn Geheugen Test*. Amsterdam: Pearson.
- Sollman, M.J. & Berry, D.T.R. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology*, 26, 774-789.
- Trueblood, W. & Binder, L.M. (1997). Psychologists' accuracy in identifying neuropsychological test protocols of clinical malingerers. *Archives of Clinical Neuropsychology*, 12, 13-27.
- Wedding, D. & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology*, 4, 233-265.