

Screenen met SCAN? Liever niet.

GLYNIS BOGAARD^a
EWOUT MEIJER^b
HARALD MERCKELBACH^c



Panopticon, 37 (3), 197-210
© 2016 Maklu | ISSN 0771-1409 | Mei 2016

- a Promovenda (doctoraatsstudent), sectie Forensische Psychologie, Faculteit der Psychologie en Neurowetenschappen, Universiteit Maastricht (Corresp.: glynis.bogaard@maastrichtuniversity.nl).
- b Universitair docent, sectie Forensische Psychologie, Faculteit der Psychologie en Neurowetenschappen, Universiteit Maastricht.
- c Professor, sectie Forensische Psychologie, Faculteit der Psychologie en Neurowetenschappen, Universiteit Maastricht.

ABSTRACT

Screening with SCAN? I don't think so.

In general, people perform very mediocre when faced with the challenge of detecting deceit. It is therefore understandable that people have searched for tools that might improve our truth finding. One of these tools, which was developed by a former polygraph examiner, is the Scientific Content Analysis (SCAN). Thirty years after its introduction, this method is widely used by both public and private organisations. However, this broad application calls for the rather urgent question as to whether SCAN is actually a reliably and accurate tool. On the basis of scientific literature we evaluated SCAN on its diagnostic merits. Our conclusions are straightforward; SCAN does not meet the requirements to be used as diagnostic tool for detecting deception. The defects are too substantial and the use of SCAN does not encourage a more objective analysis of statements.

Keywords: Verbal lie detection, deception detection, Scientific Content Analysis, SCAN

Kernwoorden: Verbale leugendetectie, Scientific Content Analysis, SCAN

LIEGEN EN LEUGENDETECTIE

Al sinds de jaren '50 van de vorige eeuw zoeken wetenschappers naar hulpmiddelen waarmee leugenaars te ontmaskeren zijn. Het belang van zulke hulpmiddelen is evident: onderzoek laat immers telkens weer zien dat we veel moeite hebben met het detecteren van leugens en doorgaans niet beter presteren dan kop-of-munt (BOND & DEPAULO, 2006, 2008; ELAAD, 2003; VRIJ, EDWARD, & BULL, 2001). Opvallend genoeg geldt dit ook voor experts die in hun dagelijkse werk te maken hebben met het verhoren van verdachten, getuigen, slachtoffers, etc.. Om oordelen over hun leugenachtigheid dan wel eerlijkheid accurater te maken, zijn verscheidene methoden ontwikkeld die de pretentie hebben om experts te helpen bij hun oordeelsvorming.

Een van deze methoden is de Scientific Content Analysis (SCAN), in Nederlandse vertalingen ook de Wetenschappelijke Inhoudsanalyse genoemd (SAPIR, 2005). SCAN werd in 1984 ontwikkeld door Avinoam SAPIR, een voormalig luitenant-polygrafist uit Israël. Tegenwoordig geeft hij SCAN-cursussen over de hele wereld. Landen waarin de cursussen regelmatig worden gegeven, zijn onder andere het Verenigd Koninkrijk, Australië, Israël, Mexico, Qa-

tar en België (zie <http://www.lsiscan.com/>). SCAN maakt zelfs deel uit van allerlei leerprogramma's die onder auspiciën van de *FBI*, *U.S. Army Military Intelligence*, de *Royal Canadian Mounted Police*, de *Royal Australian Police College* en de *American Society for Industrial Security* worden aangeboden (BOCKSTAELE, 2008a). Sinds 1999 wordt de SCAN-cursus ook jaarlijks in Oost-Vlaanderen door het Centrum voor Politiestudies aangeboden (BOCKSTAELE, 2008b). We kunnen dus stellen dat SCAN wereldwijd wordt toegepast en haar weg heeft gevonden naar zowel publieke als private instanties.

Dat roept de nogal dringende vraag op of deze brede toepassing van SCAN kan bogen op robuuste evidentie. Is deze methode betrouwbaar en nauwkeurig genoeg? In dit artikel gaan we die vraag beantwoorden. Om daarop alvast een voorschot te nemen: onze conclusie zal zijn dat de betrouwbaarheid en nauwkeurigheid van de SCAN ondermaats zijn. Om duidelijk te maken hoe we tot die conclusie komen, bespreken we aan de hand van de wetenschappelijke literatuur ter zake eerst hoe de SCAN-methode werkt en daarna in welke mate SCAN van waarde kan zijn voor het politionele opsporingswerk.

SCAN: VERANKERING

Wie de *SCAN-website* (www.lsiscan.com) raadpleegt, komt te weten dat SCAN een analysemethode is die kijkt naar de inhoud van op schrift gestelde verklaringen van getuigen, slachtoffers en verdachten. Het idee is dat je aldus uitspraken kunt doen over de waarachtigheid van deze verklaringen en 'elke zaak snel en makkelijk' kunt oplossen. Een SCAN-analyse toont volgens de *website* namelijk aan 'of de ondervraagde de waarheid spreekt of liegt, welke informatie de ondervraagde achterhoudt en of de ondervraagde betrokken was bij de misdaad' (Nederlandse vertaling van de tekst op de *website* door eerste auteur). Ook het SCAN-handboek maakt duidelijk dat de methode twee belangrijke doelen nastreeft: (1) informatie verzamelen en (2) misleiding opsporen (SAPIR, 2005, p.11).

SAPIR stelt dat ware en onware verklaringen van elkaar verschillen, hetgeen hem voor het eerst zou zijn opgevallen tijdens zijn werk als polygrafist bij de Israëlische politie (BOCKSTAELE, 2008b). Ware verklaringen, zo luidt de redenering, komen namelijk uit het geheugen, terwijl onware verklaringen vanuit de verbeelding ontstaan. Dit leidt tot inhoudelijke en structurele verschillen tussen beide soorten verklaringen (SAPIR, 2005). SCAN staat overigens niet alleen in haar nadruk op de inhoudelijke elementen van verklaringen. Zo'n nadruk is ook typerend voor andere analysemethoden, zoals Reality Monitoring (RM; JOHNSON & RAYE, 1981) en Criteria Based Content Analysis (CBCA; STELLER & KÖHNKEN, 1989; Undeutsch, 1967)

ZES FASEN

Een typische SCAN-procedure bestaat uit zes fases. De eerste fase is de kennismaking met de te verhoren persoon (initiële fase). In de tweede fase wordt de persoon gevraagd een verklaring op schrift te stellen, waarbij hij/zij antwoordt op de vraag 'wat is er gebeurd?'. Op deze manier wordt de persoon uitgenodigd om een verklaring in eigen woorden te gieten, zonder enige tussenkomst van de verhoorder. Deze versie wordt de 'pure versie' genoemd en dient als uitgangspunt voor de SCAN-analyse.

In de derde fase wordt de verklaring onderzocht op de aan- of afwezigheid van verscheidene criteria. Sommige criteria worden door middel van een kleurcode herkenbaar gemaakt; voor andere worden woorden in de verklaring omcirkeld, onderlijnd of omkaderd (BOCKSTAELE, 2008a). Een voorbeeld van een criterium is het gebruik van voornaamwoorden. SAPIR stelt dat leugenaars zich emotioneel proberen te distantiëren van een verklaring. Hierdoor zijn ze minder geneigd om persoonlijke voornaamwoorden te gebruiken en des te meer de derde

persoon bezigen (BOCKSTAELE, 2008b; SAPIR, 2005). Iemand die de waarheid vertelt, zal zinnen formuleren in eerste persoon enkelvoud (vb. 'ik zag de dief nog net wegrennen richting het stadscentrum'). Uit elke afwijking (vb. 'de dief rende weg richting het stadscentrum') zou dan kunnen worden afgeleid dat de persoon niet de (gehele) waarheid vertelt. Appendix A geeft een overzicht van de meest gebruikte SCAN-criteria, gebaseerd op (BOGAARD, MEIJER, VRIJ, BROERS, & MERCKELBACH, 2014b; VRIJ, 2008).

In de vierde fase wordt de SCAN-analyse nagelopen en worden er open vragen geeneerd, om zo belangrijke details in de verklaring op te helderen. Aan de hand van deze vragen wordt vervolgens een gedetailleerd verhoor afgenomen (fase vijf). De laatste fase bestaat uit een nagesprek om te achterhalen of de betrokkene een eerlijk of een leugenachtig persoon is. Tijdens dit gesprek zal gevraagd worden om gedetailleerd uit te weiden over enkele zaken die aan bod kwamen tijdens het verhoor. Het achterliggende idee is dat leugnars meer moeite hebben met het beantwoorden van deze vragen dan eerlijke mensen en dat dit tijdens het gesprek duidelijk zal worden (SAPIR, 2005).

RANDVOORWAARDEN

Instrumenten als de SCAN hebben een diagnostische pretentie: de gedachte is dat je met zo'n instrument het onderscheid tussen leugenachtige en eerlijke uitspraken beter kunt maken dan zonder zo'n instrument. Voor diagnostische tests bestaat er in de wetenschap een helder beoordelingskader (zie bijvoorbeeld voor medische diagnostiek MAXIM, NIEBO, & UTELL, 2014). Het belangrijkste punt in dit verband is dat het instrument voldoende betrouwbaarheid en validiteit bezit (GIESBRECHT & PETERS, 2008). Betrouwbaarheid verwijst naar de mate waarin een test onder gelijke omstandigheden en op dezelfde manier toegepast tot vergelijkbare conclusies komt. Dus als meerdere beoordelaars de SCAN toepassen op dezelfde verklaring, zouden zij tot eenzelfde conclusie dienen te komen.

Bij validiteit gaat het om zowel criterium-validiteit als construct-validiteit. De eerste betreft de mate waarin een test meet wat hij beweert te meten. Op de SCAN toegepast wil dat zeggen dat het instrument in staat moet zijn om (1) bruikbare informatie te genereren die je anders niet zou hebben verkregen en (2) ware van onware verklaringen te onderscheiden. Criterium-validiteit wordt vaak uitgedrukt als nauwkeurigheid. Om een schatting te kunnen geven van de nauwkeurigheid is een 'gouden standaard' vereist (Maxim et al., 2014). Je moet weten welke verklaringen waar en welke onwaar zijn en pas dan valt er een beargumenteerde uitspraak over de betrouwbaarheid van een instrument als de SCAN te doen. Afgaan op enkel het consumentenoordeel van de gebruikers is hachelijk. De geschiedenis van de geneeskunde is bijvoorbeeld geplaveid met diagnostische tests die in de ogen van hun bedenkers uitmuntend waren, maar afgezet tegen een gouden standaard een bedenkelijke kwaliteit bleken te bezitten. We benadrukken dit punt om duidelijk te maken waarom we zoveel waarde hechten aan wetenschappelijke studies – doorgaans laboratorium experimenten – waarin de SCAN werd getoetst tegen een gouden standaard. Construct-validiteit, tot slot, verwijst naar de onderbouwing van de test; het gaat dan om de vraag of de test – in dit geval SCAN – in een deugdelijke theorie is verankerd.

Voldoet SCAN aan de randvoorwaarden van betrouwbaarheid, criterium-validiteit en construct-validiteit? Om die vraag te beantwoorden, beroepen wij ons bij voorkeur op artikelen gepubliceerd in wetenschappelijke tijdschriften. Zulke artikelen zijn door collega-wetenschappers beoordeeld, en voldoen daarmee aan minimale eisen op het gebied van bijvoorbeeld methodologie. In discussies over de merites van SCAN wordt soms echter ook verwezen naar een aantal publicaties dat niet in wetenschappelijke vaktijdschriften ver-

scheen en dat bijvoorbeeld te vinden is in overheidsrapporten en afstudeerscripties. Voor de volledigheid zullen we hieronder eveneens stil staan bij deze categorie van publicaties.

Om een volledig overzicht van de SCAN literatuur te verkrijgen, zochten we via *Google Scholar* naar literatuur met de volgende zoektermen 'SCAN, Scientific Content Analysis, lie detection, deception en leugendetecie'. Op de zoekperiode werd geen beperking gezet. In totaal leverde deze zoektocht acht artikelen op waarin empirische gegevens over SCAN werden gerapporteerd (BOGAARD, MEIJER, & VRIJ, 2014; BOGAARD, MEIJER, VRIJ, BROERS, & MERCKELBACH, 2014a; BOGAARD, MEIJER, et al., 2014b; DRISCOLL, 1994; NAHARI, VRIJ, & FISHER, 2012; PORTER & YUILLE, 1996; SMITH, 2001; VANDERHALLEN, JASPAERT, & VERVAEKE, 2015). In ons persoonlijk archief beschik- ten we bovendien over de stukken van BOCKSTAELE (2008a, 2008b en 2015a) en hun literatuur- lijst heeft ons uiteindelijk geleid naar de afstudeerscripties van FLEMINGS (2009) en VAN GEEST (2008).

Betrouwbaarheid

Zoals gezegd verwijst betrouwbaarheid naar de mate waarin meerdere beoordelaars die SCAN toepassen op dezelfde verklaring tot een vergelijkbare conclusie komen. Een drietal studies (een overheidsrapport en twee wetenschappelijke publicaties) keek naar deze betrouwbaarheid en wel bij rechercheurs die in het gebruik van de methode waren getraind. De eerste studie – een overheidsrapport uitgevoerd in opdracht van het Britse *Home Office* – vergeleek vijf groepen van experts die ware en gelogen verklaringen moesten beoorde- len (SMITH, 2001). Drie van deze groepen bestonden uit beoordelaars die in wisselende mate ervaring hadden met het toepassen van SCAN (beginners, regelmatige en ervaren SCAN- analisten). De twee andere groepen bestonden uit beginnende en ervaren rechercheurs die geen gebruik maakten van SCAN. Iedereen werd gevraagd om dezelfde 27 verklaringen te beoordelen op hun waarachtigheid. Deze verklaringen werden eerder door Amerikaanse politiebambten op basis van verkregen bekentenissen en de mate van ondersteunend bewijs ingedeeld als *waar (4)*, *gelogen (20)* of *onbeslist (3)*. De resultaten lieten zien dat er weinig overeenkomst was tussen de verschillende beoordelaars als hen werd gevraagd om op basis van SCAN uitspraken te doen over de waarachtigheid van dezelfde verklaringen. Beoordelaars gebruikten doorgaans compleet verschillende criteria om tot hun conclusies te komen. De mate van overeenstemming tussen beoordelaars over de aan- of afwezigheid van SCAN-criteria in verklaringen varieerde tussen 0 en 40%, wat wil zeggen dat de beoor- delars het over de aanwezigheid van het merendeel van de criteria oneens waren.

Onlangs onderzochten ook wij (BOGAARD *et al.*, 2014) de betrouwbaarheid van SCAN in 82 zedenzaken van de Nederlandse politie. Alle verklaringen kwamen tot stand volgens de richtlijnen van SCAN en voor elke zaak waren twee SCAN-analyses beschikbaar. Voor elk van de 82 verklaringen – en voor elk SCAN criterium apart – vergeleken we steeds de SCAN analyses van de twee beoordelaars. Meer bepaald vergeleken we of de twee beoordelaars het met elkaar eens waren over de aan- of afwezigheid van ieder SCAN-criterium. In totaal onderzochten we de betrouwbaarheid van 15 criteria, waarbij 4 SCAN-analisten waren be- trokken. Onze resultaten toonden dat de mate van overeenstemming per SCAN criterium varieerde tussen 0% en 79%. Als we de overeenstemming voor alle SCAN criteria middelden, kwamen we uit op een overeenstemming van 31%. Voor het overgrote deel van de criteria waren beoordelaars het dus met elkaar oneens of de verklaringen er nu wel of niet aan voldeden.

VANDERHALLEN *et al.* (2015), tot slot, vroegen aan studenten, ervaren rechercheurs (beiden zonder SCAN) en SCAN-getrainde rechercheurs om uitspraken te doen over de waarchtig- heid van vier verklaringen. Enkel de laatste groep werd gevraagd de verklaringen te analy-

seren met behulp van SCAN, de overige groepen kwamen tot hun conclusie zonder SCAN. De onderzochte verklaringen handelden over een verkeersongeval waarbij de schrijver betrokken was. Slechts voor twee van de 12 onderzochte criteria vonden de auteurs een matige tot goede betrouwbaarheid, terwijl alle andere criteria een lage betrouwbaarheid hadden, variërend tussen 0 en 56%.

Criterion-validiteit

Het is psychometrisch gezien onmogelijk dat een test die in de handen van verschillende beoordelaars steeds andere conclusies genereert – lage betrouwbaarheid dus – wél valide is (MAGNUSSON, 1967). De mate van betrouwbaarheid markeert de bovengrens voor de mate van validiteit. Vanwege de teleurstellende overeenstemming tussen SCAN-gebruikers onderling mag men dus geen hoge verwachtingen koesteren over de validiteit van het instrument.

Belangrijk is dat we bij het bespreken van de validiteit een onderscheid maken tussen twee soorten studies: studies met rechercheurs die de originele SCAN-training hebben gevolgd en studies met studenten of andere leken die door de onderzoeker werden getraind in het gebruik van SCAN. Het belangrijkste verschil is dat bij deze laatste groep de betrouwbaarheid doorgaans hoog is, omdat met een vaste lijst van SCAN-criteria wordt gewerkt, waarbij elk criterium helder en duidelijk is geformuleerd en wordt gescoord met behulp van een vast scoringssysteem.

Ondertussen is er een zestal studies naar de validiteit van SCAN gepubliceerd in *peer-reviewed* tijdschriften. Twee studies die niet gepubliceerd zijn in zulke tijdschriften worden wel vaak aangehaald als bewijs voor de bruikbaarheid van SCAN (namelijk DRISCOLL, 1994; SMITH, 2001). We zullen daarom ook deze studies – en hun beperkingen – uitvoerig bespreken. Tabel 1 geeft een kort overzicht van de bevindingen van de studies die aan de orde komen in dit artikel. We staan eerst stil bij studies waarin SCAN-getrainde rechercheurs gevraagd werd verklaringen te analyseren en daarna bij studies waarin studenten werden getraind in het toepassen van SCAN.

TABEL 1. OVERZICHT VAN DE BEVINDINGEN VAN ALLE SCAN STUDIES¹

Auteur	Jaartal	Beoordelaars	Aantal verklaringen	Aantal criteria	Wetenschappelijke publicatie	Gouden standaard	Valide
DRISCOLL	1994	onderzoekers	30	10	nee	nee	ja
PORTER & YUILLE	1996	onderzoekers	60	3	ja	ja	nee
SMITH	2001	politie + SCAN ²	27	13	nee	nee	ja
NAHARI, VRIJ & FISHER	2012	onderzoekers	61	13	ja	ja	nee
BOGAARD, MEIJER & VRIJ	2014	studenten	128	11	ja	ja	nee
BOGAARD, MEIJER, VRIJ, BROERS & MERCKELBACH	2014a	politie + SCAN en studenten	64	12	ja	ja	nee
BOGAARD, MEIJER, VRIJ, BROERS & MERCKELBACH	2014b	politie + SCAN	82	28	Ja	Nvt	nee
VANDERHALLEN, JASPAERT & VERVAEKE	2015	politie + SCAN	4	13	ja	ja	nee

1 De studie van VAN GEEST (2008) is niet opgenomen in de lijst aangezien het een ongepubliceerde scriptie is.
 2 Politie + SCAN houdt in dat deze politiebeamten ten minste de SCAN basis training gevolgd hebben.

Eén van de eerste studies naar de validiteit van SCAN stipten we hierboven ook al aan, namelijk het overheidsrapport van SMITH (2001). Haar resultaten toonden dat de drie SCANGroepen (beginners, regelmatige en ervaren SCAN-analisten) niet van elkaar verschilden en in staat waren om minimaal 80% van de ware en 75% van de onware verklaringen correct te beoordelen. De SCAN-analisten scoorden daarmee niet beter dan de ervaren rechercheurs zonder SCAN, maar wel beduidend beter dan de beginnende rechercheurs. SCAN bezat dus alleen een toegevoegde waarde ten opzichte van deze laatste groep. Opvallend in deze studie was dat elke groep in staat was om leugens boven kans niveau te detecteren.

Er zijn diverse kanttekeningen te plaatsen bij de uitkomsten van dit rapport. De belangrijkste is wel dat de verklaringen gecategoriseerd werden op basis van ondersteunend bewijs. Dat is geen waterdichte methode om de gouden standaard vast te stellen. Gebrek aan bewijs betekent allerminst dat een verklaring gelogen is. Bovendien blijft het onduidelijk of de twijfelachtige verklaringen wel of niet meegenomen werden in de analyses en hoe deze mogelijk de resultaten beïnvloedden (ARMISTEAD, 2011).

De meest recente wetenschappelijke publicatie over SCAN is van VANDERHALLEN *et al.* (2015) en keek niet alleen naar de betrouwbaarheid (zie boven), maar ook naar de validiteit van SCAN. Studenten en ervaren rechercheurs zonder SCAN bleken beter in staat om ware verklaringen te herkennen dan de rechercheurs die SCAN toepasten (82% vs. 57%). Bij het herkennen van onware verklaringen verschilden rechercheurs met SCAN (78%) niet significant van de rechercheurs zonder SCAN (63%), maar beide groepen presteerden wel beter dan studenten (47%). Hoewel rechercheurs met en zonder SCAN dus niet van elkaar verschilden in het identificeren van onware verklaringen, was de eerste groep wel veel slechter in het identificeren van ware verklaringen. SCAN lijkt dus vooral aan te zetten tot achterdocht en te leiden tot vals-positieve oordelen (waren verklaringen voor leugenachtig houden). Merk op dat bij deze studie de gouden standaard onbetwistbaar was: de gebruikte verklaringen konden met zekerheid als waar of gelogen worden ingedeeld.

De oudste studie naar de validiteit van SCAN-criteria betreft die van DRISCOLL (1994). Hij verzamelde 30 verklaringen en aan de hand van het beschikbare bewijs deelde hij ze in als *schijnbaar accuraat* (waarheid) en *twijfelachtig* (gelogen). De verklaringen werden gescoord aan de hand van tien SCAN-criteria, en dat resulteerde in een correcte classificatie van 84% van de verklaringen. Net zoals bij de studie van SMITH, is het probleem van dit onderzoek het ontbreken van een gouden standaard: het is voor de gebruikte verklaringen in deze studie onduidelijk of ze echt waar of gelogen zijn. Dit bemoeilijkt de interpretatie van DRISCOLLS bevindingen.

PORTER AND YUILLE (1996) publiceerden het eerste artikel over SCAN in een wetenschappelijk tijdschrift. Zij losten het probleem van de gouden standaard op door proefpersonen een geënceneerd misdrijf te laten plegen. Hierover moesten ze later een ware of een gelogen verklaring afleggen. Deze verklaringen werden geanalyseerd aan de hand van drie SCAN-criteria: structuur, missende informatie en het gebruik van de eerste persoon enkelvoud-verleden tijd. Ware en onware verklaringen verschilden niet in de mate waarin ze voldeden aan deze criteria. Hoewel de gouden standaard in deze studie vaststond, maakten de onderzoekers geen gebruik van een zuivere SCAN-procedure. De verklaringen van de proefpersonen werden namelijk eerst mondeling afgenomen en pas daarna getranscribeerd. Mogelijk heeft dit een nadelige invloed gehad op de kwaliteit van de tekst, en dus ook op de SCAN-analyse.

Een decennium later onderzochten NAHARI *et al.* (2012) of SCAN van waarde kon zijn als leugendetectie-methode. Hun studie verscheen eveneens in een *peer-reviewed* tijdschrift. En net zoals PORTER en YUILLE (1996) lieten NAHARI *et al.* (2012) hun proefpersonen naar het lab komen om een examen te stelen of een onschuldige handeling uit te voeren. Proefpersonen

schreven daarna een ware of gelogen verklaring over wat ze hadden gedaan. Een totaal van 13 SCAN-criteria werden getest, maar ware en onware verklaringen verschilden voor geen van deze criteria van elkaar. Ook wanneer alle criteria werden samengenomen, was er geen verschil.

Ook BOGAARD, MEIJER, and VRIJ (2014) en BOGAARD, MEIJER, *et al.* (2014a) onderzochten SCAN in het lab en publiceerden hierover in wetenschappelijke tijdschriften. In beide studies moesten proefpersonen een ware en een onware autobiografische verklaring opschrijven. Twaalf SCAN-criteria werden getest, maar andermaal kon SCAN geen onderscheid maken tussen ware en onware verklaringen. De nauwkeurigheid van SCAN werd in beide studies vergeleken met de twee eerder genoemde analysemethoden, namelijk RM en CBCA. Deze methoden waren wél in staat om ware van onware verklaringen te onderscheiden. Blijkbaar waren er dus inhoudelijke verschillen tussen beide type verklaringen die door de SCAN niet, maar door RM en CBCA wel werden opgepikt.

Construct-validiteit

De tweede vorm van validiteit betreft de theoretische verankering van de SCAN. SAPIR stelt dat ware verklaringen uit het geheugen komen, terwijl onware verklaringen vanuit de verbeelding ontstaan waardoor er inhoudelijke en structurele verschillen ontstaan (SAPIR, 2005). Dit idee heeft veel weg van de UNDEUTSCH-hypothese, waarop de CBCA is gebaseerd (STELLER, 1989; STELLER & KÖHNKEN, 1989; UNDEUTSCH, 1967). Ook die hypothese gaat ervan uit dat mensen anders vertellen over echte gebeurtenissen dan over bij elkaar gefantaseerde gebeurtenissen, omdat het grondmateriaal waaruit correcte herinneringen en fantasieën zijn opgebouwd zou verschillen. Maar de UNDEUTSCH-hypothese houdt geen rekening met *waarom* en *wanneer* deze verschillen zouden optreden (SPORER, 1997), en dat verzuim geldt ook voor de rationale achter SCAN. Stel bijvoorbeeld dat een verdachte of getuige een onware verklaring aflegt waarin hij of zij zelf heilig gelooft. Dat komt voor, zo weten uit de lijvige literatuur over pseudoherinneringen (zie bijvoorbeeld PORTER & BAKER, 2015). Kunnen SCAN en CBCA dan de hand leggen op het onware karakter van de verklaring? Meer in het algemeen rijst de vraag of waarheidsgetrouwe beschrijvingen van wat men gezien heeft en beschrijvingen van wat men gefabuleerd heeft wel verwijzen naar mentale compartimenten die strikt zijn gescheiden. Onderzoek doet namelijk vermoeden dat perceptie en imaginatie nogal met elkaar overlappen (DENIS & KOSSLYN, 1999; KOSSLYN *et al.*, 1999). In weerwil van deze kritiek is de CBCA, in tegenstelling tot SCAN, wel redelijk in staat om een onderscheid tussen ware en onware verklaringen te maken.

CONTROVERSE

Hierboven bespraken wij het – weinig hoopgevende – onderzoek naar de betrouwbaarheid en validiteit van SCAN. Niet iedereen interpreteert dit onderzoek even negatief (zie bijvoorbeeld BOCKSTAEL 2008a;2008b;2015). Hieronder staan we kort stil bij waarom pleitbezorgers van de SCAN positief oordelen over het instrument.

Laboratoria zijn niet echt

Een vaak gehoord kritiekpunt is dat veel van het onderzoek dat we hierboven bespraken in een kunstmatige laboratoriumomgeving tot stand kwam, en dat dit type onderzoek derhalve weinig zegt over de complexe werkelijkheid (BUCKLEY, 2012). Ons komt dit voor als een kanttekening die redelijker klinkt dan ze feitelijk is. Weinigen zouden haar plaatsen bij

labonderzoek naar pakweg een virus dat ernstige aandoeningen veroorzaakt. Wat telt, is de eerder beschreven construct-validiteit. Die bepaalt namelijk hoe een methode getest kan worden. Als – zoals bij SCAN – de nadruk ligt op of een verhaal wel of niet uit het geheugen afkomstig is, dan zou die kwestie in beginsel ook in het laboratorium – waar deelnemers bijvoorbeeld gevraagd wordt om een ware en een verzonden negatieve gebeurtenis te rapporteren – te bestuderen moeten zijn. En natuurlijk, perfecte laboratoriumstudies bestaan niet. Maar dat maakt de teleurstellende resultaten van de imperfecte studies die wél zijn gedaan nog niet irrelevant. Zoiets volhouden komt neer op wat DAWES (2009; p. 25) ‘arguing from the vacuum’ noemt: ‘Supportive evidence is simply hypothesized while negative evidence that has actually been collected is ignored.’

De SCAN analisten waren niet ervaren genoeg

Een tweede verklaring die volgens pleitbezorgers de matige onderzoek prestaties van de SCAN zou kunnen verklaren, is de variabiliteit in ervaring van de SCAN-analisten. Lage betrouwbaarheid zou dan komen doordat de SCAN-analisten verschillen in de mate waarin zij ervaring hebben met SCAN; en of zij bijvoorbeeld de 3-daagse of de 5-daagse opleiding volgde. Was er gebruik gemaakt van SCAN-analisten met voldoende ervaring en opleiding – is dan de redenering – dan waren de resultaten zeker anders geweest (BOCKSTAELE, 2015).

Het is een kanttekening die ons niet overtuigd. Cursisten ontvangen immers na de 3-daagse cursus een officieel SCAN-certificaat, hetgeen ze in staat zou moeten stellen om SCAN betrouwbaar toe te passen. Beweren dat de cursus ontoereikend is, roept dat de vraag op wanneer een SCAN-analist dan wel voldoende getraind is om de methode toe te passen. Wordt dat niet van tevoren vastgelegd, dan dreigt een circulaire redenering: SCAN werkt altijd, want als het niet werkt betekent dat slechts dat de analisten niet ervaren genoeg waren. Belangrijk, tot slot, is om op te merken dat de data van SMITH (2001) geen steun biedt voor de ervaringshypothese. De drie groepen SCAN analisten (beginners, regelmatige en ervaren SCAN-analisten) presteerden allemaal gelijkaardig.

SCAN is geen leugendetectie

Het laatste punt van kritiek op het empirisch onderzoek naar SCAN komt er op neer dat SCAN – volgens sommigen van haar pleitbezorgers – in het geheel niet de pretentie heeft om te kunnen uitmaken of een persoon al dan niet liegt (zie bijvoorbeeld BOCKSTAELE 2008a; 2008b; 2015). Het komt ons als een curieus argument voor, al was het maar omdat op de ‘officiële’ *website* (www.lscan.com) van het instrument dit valt te lezen: ‘SCAN is a new, innovative and effective technique for obtaining information and detection of deception by analyzing the words people use’. Die mededeling staat niet op zichzelf. De beschrijving van de SCAN-cursus aangeboden via www.politiestudies.be luidt als volgt: ‘Door analyse (met SCAN) van de eigenhandig geschreven woorden is het mogelijk dit bedrog te achterhalen.’ Dit laat geen twijfel over de pretentie van SCAN.

De eerste auteur woonde trouwens de driedaagse SCAN-cursus in Gent bij en constateerde dat SCAN ook aan de cursisten werd voorgehouden als een leugendetectie-methode.

Volgens de *maar-het-is-geen-leugendetectie*-redenering zou SCAN slechts helpen met het detecteren van zogenaamde *hotspots*; dit zijn opmerkelijke delen binnen een verklaring die meer uitleg vereisen. Nadat deze *hotspots* zijn geïdentificeerd, kunnen vervolgens vragen gegenereerd worden. Met deze specifieke vragen zou het mogelijk zijn om betere informatie te verkrijgen (BOCKSTAELE, 2015). Ook hier geldt dat er geen onderzoek bestaat dat deze bewering kan ondersteunen. De afstudeerscriptie van VAN GEEST (2008) liet zien dat het

al dan niet gebruiken van SCAN geen verschil maakte voor het aantal *hotspots* dat werd geïdentificeerd. Noch had SCAN enig effect op de kwantiteit of kwaliteit van de vervolgvragen. Die bevinding is overigens niet zo verbazingwekkend. Hoe zou immers een methode die geen onderscheid kan maken tussen leugens en waarheid, wel goede vervolgvragen kunnen genereren?

ACHTERDOCHT

Bovendien rijst de vraag of SCAN in haar zoektocht naar kwestieuze passages niet aanzet tot overmatige achterdocht en dus vals positieve fouten. De gegevens van VANDERHALLEN en collega's (2015) laten zien dat wanneer rechercheurs die SCAN gebruiken verkeerde conclusies trekken, ze vooral ware verklaringen voor leugenachtig houden; rechercheurs zonder SCAN maakten dit type fout veel minder. Het gaat hier om een type fout dat in de opsporing zo veel mogelijk vermeden dient te worden (VOLOKH, 1997).

In dit opzicht sluit het empirisch kennisbestand over SCAN naadloos aan bij onderzoek naar de invloed van de serie 'Lie to me' op het herkennen van leugens (LEVINE, SEROTA, & SHULMAN, 2010). Kijkers van deze serie versleten waarheidsprekers vaker voor leugenaars dan mensen die deze serie niet hadden bekeken. Belangrijk hierbij is dat de nonverbale signalen die in de serie aan bod kwamen als zijnde indicatief voor liegen, in feite daarvoor niet diagnostisch zijn, net zomin als de SCAN-criteria. Wie desondanks van zulke signalen gebruik maakt, wordt overmatig argwanend, een fenomeen dat als *investigator bias* te boek staat (KASSIN, GOLDSTEIN, & SAVITSKY, 2003). De plicht om aan te tonen dat dit toch wel akelige risico voldoende kan worden uitgesloten, rust op de schouders van pleitbezorgers van SCAN.

Onze samenvatting van de literatuur laat in ieder geval zien dat SCAN haar pretenties niet kan waarmaken. Het onderzoek ter zake laat maar één conclusie toe: SCAN voldoet niet aan de randvoorwaarden om als wetenschappelijke verankerde methode – *bewezen effectief* zo men wil – door het leven te gaan. Dat heeft te maken met de lage betrouwbaarheid, en de gebrekkige construct- en criterium-validiteit. Of in ronder Nederlands: SCAN wordt door experts op nogal uiteenlopende wijze toegepast, en het instrument meet niet wat het beweert te meten.

WAAROM WERKT HET NIET?

Waarom is SCAN niet in staat om ware van onware verklaringen te onderscheiden? Daarvoor zijn vijf mogelijke verklaringen. Ten eerste ontbreken belangrijke criteria die te maken hebben met de mate van details en zintuigelijke waarnemingen. Van deze criteria staat vast dat ze een redelijke aanwijzing in de richting van waarachtigheid opleveren (AMADO *et al.*, 2015; MASIP, SPORER, GARIDO, & HERRERO, 2005; VRIJ, 2005). In tegenstelling tot wat sommige auteurs beweren (zie bijvoorbeeld BOCKSTAELE 2008a, 2008b), zijn deze criteria niet terug te vinden in het SCAN-handboek (SAPIR, 2005), en maken ze geen deel uit van de SCAN training.

Een tweede verklaring betreft de afwijkende interpretatie die SCAN aan verscheidene criteria geeft in vergelijking met CBCA en RM. Zo worden enkele criteria door SCAN beschreven als indicatief voor bedrog, terwijl dezelfde criteria volgens CBCA en RM op waarachtigheid duiden. BOGAARD en collega's (2014b) geven een uitgebreid overzicht van deze contradicties. Logischerwijs kunnen beide interpretaties niet correct zijn. Of SCAN heeft het bij het rechte eind, of RM en CBCA. RM en CBCA zijn veel beter ingebed in de wetenschappelijke literatuur; een zoektocht in de literatuur toont aan dat er inmiddels ongeveer 50 publicaties over CBCA en 25 publicaties over RM publicaties zijn verschenen, niet-Engelstalige studies daargelaten. Hier tegenover staan slechts acht SCAN publicaties. Nog belangrijker: zowel

RM (MASIP *et al.*, 2005) als CBCA (AMADO *et al.*, 2015; VRIJ, 2005) zijn accurater in het maken van een onderscheid tussen ware en onware verklaringen dan SCAN. Op basis van het voorgaande is het redelijk om aan de wijze waarop RM en CBCA criteria uitleggen tenminste het voordeel van de twijfel te geven.

Ten derde speelt onduidelijkheid over het aantal SCAN criteria een rol. Zelfs 30 jaar na de ontwikkeling van SCAN bestaat er onder wetenschappers noch gebruikers consensus over het precieze aantal criteria dat SCAN zou moeten omvatten. Verschillende bronnen rapporteren een uiteenlopend aantal criteria, variërend van tien tot 28 criteria (voor een overzicht zie BOGAARD, MEIJER, *et al.*, 2014b). Het SCAN-handboek geeft hierover in ieder geval geen uitsluitel. Nergens in het boek is een overzicht van de criteria terug te vinden en het ontbreekt evenzeer aan richtlijnen over hoe deze criteria precies geïnterpreteerd dienen te worden. Neem als voorbeeld het criterium *structuur van een verklaring*. Volgens SAPIR dient elke verklaring te bestaan uit een introductie (20%), een middenstuk (50%) en een afsluiting (30%); afwijkingen van dit stramien zouden duiden op misleiding. Maar hoe zeer moeten verklaringen hiervan afwijken opdat een SCAN-beoordelaar kan zeggen dat ze niet meer voldoen aan het criterium? Het is een onopgehelderde kwestie.

BOGAARD *et al.* (2014b) toonden dat een aantal SCAN-criteria, die verspreid in het handboek te vinden zijn, zelden of nooit voorkomen in verklaringen. Het gaat om criteria als *disproportionele of buitencontextuele vermeldingen* en *specifieke activiteiten*. Deze criteria lijken dus van weinig belang te zijn, maar ze kunnen wel ruis creëren. Wetenschappers hebben om dezelfde redenen de CBCA bekritiseerd. Criteria dienen betrouwbaar gescoord te kunnen worden en voldoende aanwezig te zijn in verklaringen. Als dit niet het geval is, moeten ze opnieuw worden gedefinieerd of moeten ze worden verwijderd uit de methode (HOROWITZ *et al.*, 1997). Ditzelfde geldt ook voor SCAN.

Een vierde punt heeft er mee te maken dat SCAN-experts doorgaans enkel de afzonderlijke SCAN-criteria interpreteren en SCAN niet als geheel toepassen. Dit laatste is uiteraard ook moeilijk, omdat er onduidelijkheid bestaat over het aantal criteria en hun interpretatie. Desalniettemin waarschuwen wetenschappers voor het trekken van conclusies op basis van afzonderlijke criteria: de psychometrie leert dat betrouwbaarheid een functie is van het aantal items. Hoe meer valide items of criteria, hoe betrouwbaarder de uitkomst van een instrument (zie bijvoorbeeld AMADO *et al.*, 2015; MASIP *et al.*, 2005; VRIJ, 2005 voor onderzoek van CBCA en RM criteria).

Een vijfde punt heeft niet alleen betrekking op SCAN, maar geldt in meer of mindere mate ook voor RM en CBCA. Deze leugendetectie-methoden maken gebruik van een passieve manier van informatie vergaren. Het stellen van de open vraag 'Wat is er gebeurd?' zoals SCAN dat voorschrijft, beïnvloedt de verkregen informatie niet of nauwelijks. Hier laten deze methoden een kans liggen. Zo blijkt liegen ons doorgaans meer moeite te kosten dan het vertellen van de waarheid (VRIJ, FISHER, MANN, & LEAL, 2006). Als verhoorder kun je hiervan gebruik maken door vragen te stellen of taken te geven die het nog moeilijker maken voor leugenaars om een heldere verklaring te geven (e.g., verhaal achterstevoren vertellen, oogcontact blijven houden, gebruik van schetsen). Voorts blijken leugenaars hun verklaringen zo goed mogelijk voor te bereiden alvorens een verhoor in te gaan (HARTWIG, GRANHAG, & STRÖMWALL, 2007). Ook hier kan de verhoorder op inspelen door onverwachte vragen te stellen. Het idee is dat leugenaars harder moeten nadenken dan waarheidssprekers om deze vragen te beantwoorden. Dit zijn slechts enkele voorbeelden van strategieën waarbij de verhoorder een actieve rol speelt en hij/zij invloed kan uitoefenen op de verkregen informatie. Het gebruik van deze strategieën resulteert doorgaans in meer aantoonbare signalen voor liegen (voor een overzicht zie VRIJ & GRANHAG, 2012).

TEN SLOTTE

De SCAN voldoet niet aan de vereisten om als praktisch hulpmiddel in de opsporing te worden toegepast. De gebreken zijn substantieel en het gebruik van SCAN zet niet aan tot een meer objectieve analyse van de inhoudelijke kenmerken, maar jaagt wel een *investigator bias* aan met alle vals-positieven van dien. We begrijpen hoe belangrijk waarheidsvinding is in het forensische veld en hoe aantrekkelijk methodes zoals SCAN zijn, die pretenderen uitkomst te bieden. Die pretentie kan door SCAN niet worden waar gemaakt waardoor wij het gebruik ervan ontraden. Dat maakt ook de weg vrij voor methodes die beter zijn veranderd in de wetenschappelijke literatuur, zoals CBCA en RM. Ook daarvoor geldt trouwens dat ze geen wondermiddelen zijn, dat ze altijd een foutenmarge bezitten en dat hun bruikbaarheid afhangt van de basisfrequentie waarmee zich ware en onware verklaringen in een gegeven context aandienen.

Screenen met SCAN? Liever niet, luidt ons advies. Best mogelijk dat de pleitbezorgers van SCAN de argumentatie daarvoor niet overtuigend vinden. Maar dan is een redelijk verzoek aan hen om een stopregel te specificeren. Hoe ziet het onderzoek dat hen ernstig zou doen twifelen aan de merites van SCAN dan wel uit?

REFERENTIES

- AMADO, B. G., ARCE, R., & FARIÑA, F. (2015). UNDEUTSCH hypothesis and Criteria Based Content Analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context*, 7, 1-10.
- ARMISTEAD, T. W. (2011). Detecting deception in written statements: The British Home Office study of scientific content analysis (SCAN). *Policing: An International Journal of Police Strategies & Management*, 34, 588-605.
- BOCKSTAELE, M. (2008a). De SCAN als middel tot waarheidsvinding *Het Tijdschrift voor de Politie [Journal of Police]*, 70, 8-13.
- BOCKSTAELE, M. (2008b). Scientific Content Analysis (SCAN). Een nuttig instrument bij verhoren? In L. SMETS & A. VRIJ (Eds.), *Het analyseren van de geloofwaardigheid van verhoren: Het gebruik van leugendeteciethoden [The analysis of the credibility of interrogations: The use of lie detection methods]* (pp. 105-156). Brussels, Belgium: Politeia.
- BOCKSTAELE, M. (2015). Vragen bij de kwaliteit van rechtspsychologisch onderzoek over verhoortechnieken. *Panopticon*, 36, 375-384.
- BOGAARD, G., MEIJER, E., & VRIJ, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN. *Journal of Investigative Psychology and Offender Profiling*, 11, 151-163. doi: 10.1002/jip.1409
- BOGAARD, G., MEIJER, E., VRIJ, A., BROERS, N. J., & MERCKELBACH, H. (2014a). Contextual Bias in Verbal Credibility Assessment: Criteria Based Content Analysis (CBCA), Reality Monitoring (RM), and Scientific Content Analysis (SCAN). *Applied Cognitive Psychology*, 28, 79-90. doi: 10.1002/acp.2959
- BOGAARD, G., MEIJER, E., VRIJ, A., BROERS, N. J., & MERCKELBACH, H. (2014b). SCAN is largely driven by 12 criteria: Results from field data. *Psychology, Crime and Law*, 20, 430-449. doi: 10.1080/1068316X.2013.793338
- BOND, C. F., & DEPAULO, B. M. (2006). Accuracy of deception judgments. *Personality and Individual Differences*, 10, 214-234.
- BOND, C. F., & DEPAULO, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, 134, 477-492.

- BUCKLEY, J. P. (2012). Detection of deception researchers needs to collaborate with experienced practitioners. *Journal of Applied Research in Memory and Cognition*, 1, 126-127.
- DAWES, R. (2009). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- DENIS, M., & KOSSLYN, S. M. (1999). Scanning visual mental images: A window on the mind. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 18, 409-465.
- DRISCOLL, L. (1994). A validity assessment of written statements from suspects in criminal investigations using the SCAN technique. *Police Studies*, 4, 77-88.
- ELAAD, E. (2003). Effect of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology*, 17, 349-363.
- GIESBRECHT, T., & PETERS, M. V. J. (2008). Tests en taken: De achtergrond. In T. GIESBRECHT, C. DE RUITER & M. JELICIC (Eds.), *Forensisch psychodiagnostisch gereedschap* (pp. 13-20). Amsterdam: Hartcourt.
- HARTWIG, M., GRANHAG, P. A., & STRÖMWALL, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, 13, 213-227.
- HOROWITZ, S. W., LAMB, M. E., ESPLIN, P. W., BOYCHUK, T. D., KRISPIN, O., & REITER-LAVERY, L. (1997). Reliability of criteria-based content analysis of child witness statements. *Legal and Criminological Psychology*, 2, 11-21.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67-85.
- KASSIN, S. M., GOLDSTEIN, C. C., & SAVITSKY, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, 27, 187-203.
- KOSSLYN, S. M., PASCUAL-LEONE, A., FELICIAN, O., CAMPOSANO, S., KEENAN, J. P., THOMPSON, W. L., . . . ALPERT, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science*, 284, 167-170.
- LEVINE, T. R., SEROTA, K. B., & SHULMAN, H. C. (2010). The impact of Lie to Me on viewers' actual ability to detect deception. *Communication Research*, 37, 847-856.
- MAGNUSSON, D. (1967). *Test theory*. Reading, MA: Addison-Wesley.
- MASIP, J., SPORER, A. L., GARIDO, E., & HERRERO, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime and Law*, 11, 99-122.
- MAXIM, L. D., NIEBO, R., & UTELL, M. J. (2014). Screening tests: a review with examples. *Inhalation toxicology*, 26, 811-828.
- NAHARI, G., VRIJ, A., & FISHER, R. P. (2012). Does the truth come out in the writing? SCAN as a lie detection tool. *Law and Human Behavior*, 36, 68-76. doi: 10.1037/h0093965
- PORTER, S., & BAKER, A. T. (2015). CSI (Crime Scene Induction): Creating False Memories of Committing Crime. *Trends in Cognitive Sciences*, 19, 716-718.
- PORTER, S., & YUILLE, Y. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context *Law and Human Behavior*, 20, 443-458.
- SAPIR, A. (2005). *The LSI course on scientific content analysis (SCAN)*. Phoenix, AZ: Laboratory for Scientific Interrogation.
- SMITH, N. (2001). Reading between the lines: an evaluation of the Scientific content Analysis technique (SCAN). *Police Research Series Paper 135*, 1-42.
- SPORER, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373-397.
- STELLER, M. (1989). Recent developments in statement analysis. In J. C. YUILLE (Ed.), *Credibility Assessment* (pp. 135-154). Dordrecht: Kluwer Academic Publishers.
- STELLER, M., & KÖHNKEN, G. (1989). Criteria Based Statement Analysis. In D. C. RASKIN (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York: Springer

- UNDEUTSCH, U. (1967). Beurteilung der glaubhaftigkeit von Aussagen. In U. UNDEUTSCH (Ed.), *Handbuch der psychologie Vol 11: Forensische Psychologie*. Göttingen, Germany: Hogrefe.
- VAN GEEST, E. (2008). *Misleidingen in verklaringen: De SCAN-techniek*. Katholieke Universiteit Leuven.
- VANDERHALLEN, M., JASPAERT, E., & VERVAEKE, G. . (2015). SCAN as an investigative tool. *Police Practice and Research*, 1-15.
- VOLOKH, A. . (1997). N guilty men. *University of Pennsylvania Law Review*, 146, 173-216.
- VRIJ, A. (2005). Criteria Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11, 3-41.
- VRIJ, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: Wiley.
- VRIJ, A., EDWARD, K., & BULL, R. (2001). Police officer's ability to detect deceit: The benefit of indirect deception detection measures. *Legal and Criminological Psychology*, 6(185-196).
- VRIJ, A., FISHER, R., MANN, S., & LEAL, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, 10, 141-142.

APPENDIX A. SCAN CRITERIA (ONTLEEND AAN VRIJ (2008A)).

- (1) **Ontkennen van beschuldiging:** Wanneer iemand direct zijn schuld ontkent 'ik heb het niet gedaan' duidt dit op onschuld.
- (2) **Sociale introductie:** refereert aan hoe personen in de verklaring geïntroduceerd worden. Mensen die een rol spelen in de verklaring zouden idealiter met naam en rol beschreven moeten worden (vb. Mijn man, Jan). Afwijkingen van deze vorm duiden op misleiding.
- (3) **Structuur van de verklaring:** Een ware verklaring bestaat doorgaans uit 20% proloog, 50% hoofdzaak en 30% epiloog. Afwijkingen van deze structuur duiden op misleiding.
- (4) **Emoties:** Dit criterium onderzoekt waar emoties beschreven worden in de verklaringen. Idealiter worden emoties beschreven in de epiloog. Wanneer emoties al beschreven worden in de proloog duidt dit op misleiding.
- (5) **Objectieve en subjectieve tijd:** Objectieve tijd refereert aan de werkelijke tijd dat een gebeurtenis duurt, terwijl subjectieve tijd refereert naar het aantal regels dat gependeed wordt aan het beschrijven van de gebeurtenis. Gemiddeld genomen is de schrijfsnelheid 3 regels per objectief uur. Afwijkingen van deze schrijfsnelheid duiden op misleiding.
- (6) **Eerste persoon enkelvoud, verleden tijd:** Refereert aan de tijd waarin een verklaring geschreven wordt. Een ware verklaring zal doorgaans geschreven worden in eerste persoon enkelvoud, verleden tijd (vb. Ik zag haar liggen in de badkamer). Afwijkingen duiden op een gebrek aan betrokkenheid, en dus misleiding.
- (7) **Voornaamwoorden:** Dit criterium onderzoekt het gebruik van persoonlijke voornaamwoorden ('mijn', 'ik', 'wij'). Wanneer persoonlijke voornaamwoorden missen in een verklaring (vb. Het huis i.p.v. 'ons huis') indiceert dit dat een schrijver afstand neemt van zijn/haar verklaring en duidt dus op misleiding.
- (8) **Verandering van taalgebruik:** Refereert aan een verandering van terminologie en dus een verandering in de realiteit van de schrijver. Met name woorden die binnen de categorieën wapens, familieleden, communicatie, mensen of transport te maken hebben, dienen hierbij nagekeken te worden. Wanneer een verandering in taalgebruik aanwezig is (vb. Mes naar zakmes) maar hiervoor geen rechtvaardiging kan worden gevonden in de verklaring, duidt dit op misleiding.
- (9) **Spontane correcties:** Refereert aan alle correcties die gemaakt worden binnen een verklaring. Voor de schrijver aan zijn/haar verklaring begint krijgen ze de instructie om

niets te schrappen. Wanneer men zich niet houdt aan deze instructie duidt dit op misleiding.

- (10) **Gebrek aan herinneringen of betrokkenheid met de gebeurtenis:** Dit criterium onderzoekt of er sprake is van vaag taalgebruik binnen een verklaring (vb. Ik denk, ongeveer, zoiets). Ook onderzoekt het of de schrijver toegeeft dat hij of zij iets niet meer weet. Vaag taalgebruik en toegegeven dat men iets niet meer weet duiden op misleiding.
- (11) **Disproportionele of buitencontextuele vermeldingen:** Onderzoekt of er uitleg in de verklaring staat die spontaan gegeven wordt (vb. Ik ging naar de badkamer omdat ik mijn tanden wilde poetsen), zonder dat deze uitleg belangrijk is voor het begrijpen van de verklaring. Het gaat hierbij dus over irrelevante informatie. Op deze manier wordt de lezer afgeleid van belangrijkere zaken. De aanwezigheid van deze informatie duidt daarom ook op misleiding.
- (12) **Missing information:** Refereert aan specifieke woorden die aanduiden dat bepaalde informatie wordt weggelaten binnen een verklaring. Bijvoorbeeld, kort daarna, een uur later, de volgende dag, etc. Dit is met name relevant binnen de beschrijving van de hoofdzaak. Weglaten van informatie duidt op misleiding.