

Vragen bij de kwaliteit van rechtspsychologisch onderzoek over verhoortechnieken

MARC BOCKSTAELE^a

^a (Ere)hoofdcommissaris Federale Gerechtelijke Politie Gent; Inhoudelijk coördinator van de nationale cursussen verhoortechnieken (Corresp.: marc.bockstaele@skynet.be).

Het verhoren van slachtoffers, getuigen en vooral van verdachten is bijzonder veelzijdig en complex. Er bestaat geen 'standaardverhoor'; elk verhoor, zelfs over gelijkaardige feiten en met dezelfde bewijzen of aanwijzingen, is anders omdat de betrokkenen met hun eigen persoonlijkheid en karakter verschillen. Elke verhoorde vraagt een gepersonaliseerde aanpak en match met de verhoorder. Verhoorders ontwikkelen hun *verhoorvaardigheid* hoofdzakelijk door observatie van collega's en door ervaring. Veel minder door opleiding en bijna niet door consultatie van wetenschappelijke onderzoeken en studies, gepubliceerd in internationale tijdschriften.

Jaarlijks worden wereldwijd een paar honderd wetenschappelijke onderzoeken en studies gepubliceerd, vooral door (rechts)psychologen, in verband met alle mogelijke verhoorfacetten. Alhoewel die auteurs daarvan het misschien anders zouden denken of wensen, bereiken de meeste publicaties de verhoorders en rechtspractici niet. Het hangt bijna volledig van de lezers verhoortechnieken af, in de mate dat zij hun lessen actualiseren, of (rechts)psychologisch onderzoek de weg vindt naar de politiepraktijk. Bovendien zijn er geen echt vernieuwende onderzoeken of studies meer, evenmin nieuwe verhoortechnieken. Rechtspsychologisch onderzoek beperkt zich al een decennium tot het bevestigen, benoemen en een nieuwe naam geven aan bestaande verhoortechnieken: oude wijn in nieuwe vaten gieten. Maar het is positief dat academici verhoortechnieken opvolgen. Op die wijze wordt de weg naar al bestaande verhoortechnieken wetenschappelijk onderbouwd.

Rechtspsychologisch onderzoek over verhoren blijft dus vooral academisch en dat is maar best zo. De conclusies zijn immers niet altijd eenduidig en niet zelden onjuist door een onderzoeksdesign dat te ver af staat van de realiteit (BOCKSTAELE, 2015). HUYBRECHTS (2011) schreef ook al dat de keuze die de waarnemer maakt bij het opzetten van zijn experiment in belangrijke mate de uitkomst daarvan bepaalt. Dat de kwaliteit van rechtspsychologisch onderzoek inzake verhoren niet altijd deugdelijk is illustreren we met een analyse van enkele wetenschappelijke onderzoeken over de SCAN-verhoortechniek.

1. WAT IS SCIENTIFIC CONTENT ANALYSIS (SCAN)?

De SCAN bestaat uit zes fases (SAPIR, 2005, nr. 195-198):

- de initiële fase: het bekomen van de persoonlijke gegevens;
- het bekomen van de 'zuivere versie' (pure statement);
- de analyse van de tekst, het middels regels (rules) opsporen van kritische punten (hot-spots);
- het gedetailleerd verhoor waarin de resultaten van de analyse zijn verwerkt;
- de fase na het verhoor;
- het bekomen van de bekentenis, als de verhoorder uit voorgaande fases meent dat de verhoorde misleidend is.

Essentieel voor een goede SCAN is het bekomen van de 'zuivere versie', waarbij de betrokkene, aan de hand van *één open vraag*, wordt uitgenodigd zelf zijn verklaring over een inci-

dent waarbij hij zelf aanwezig was te (laten) noteren, zonder enige vraag of sturing van de onderzoeker. Klassieke verhoren zijn niet zelden beïnvloed door de visie van de onderzoeker, die op basis van de politionele, technische en tactische vaststellingen suggestieve vragen zal stellen. De SCAN sluit dit nadeel van klassieke verhoren uit, en is daarom de *meest zuivere verhoormethode*.

Een analyse van een verklaring, de derde fase van de SCAN, kan men vergelijken met het onderzoek ter plaatse bij een moordzaak. Een ongetrainde politieman kan op een plaats delict (PD) rondwandelen en absoluut niets zien wat van waarde kan zijn voor het onderzoek. Een onderzoeker die getraind is in het systematisch doorlopen en doorzoeken van diezelfde PD zal heel wat bruikbare elementen waarnemen. Het resultaat van een SCAN geeft de kritische punten in een verklaring weer die verder bevraagd of onderzocht moeten worden. Uit de praktijk blijkt dat verhoorders die een cursus SCAN volgden, in een verklaring méér te bevragen elementen voor de waarheidsvinding zien dan ‘niet SCAN-getrainde’ verhoorders. Geregeld verschijnen in vaktijdschriften artikelen over de meerwaarde die de analyse van schriftelijke verklaringen, zoals SCAN, in de verhoorpraktijk biedt (bv. BURKE, 2013; MATSUMOTO, HWANG & FRANK, 2012; SCHAFER, 2008; SANDOVAL, 2008; ADAMS & HARPSTER, 2008; WELLS, 2008; BREWER, SAMPAIO & BARLOW, 2005; ADAMS & JARVIS, 2004; SMITH & SHUY, 2002).

Een goed verhoorder vraagt zich bij elk verhoor af of de verhoorde de waarheid spreekt, misleidt of zich gewoon vergist. Hij beschikt niet altijd over externe controlepunten en laat zich daarom heel dikwijls leiden door zijn ‘buikgevoel’, een ander woord voor verhoorer-ervaring. Een verhoorder die geen signalen van misleiding kent en niet voortdurend alert is voor misleiding, is naïef en niet goed bezig. Een SCAN kan dit aanvoelen ondersteunen en leiden tot minder verkeerde inschattingen en meer gerichte vraagstellingen met meer onderzoekshandelingen voor de waarheidsvinding tot gevolg.

2. HET ONDERZOEK VAN BOGAARD, MEIJER, VRIJ, BROERS EN MERCKELBACH (2014)

De conclusies van dit onderzoek waren dat SCAN geen goed gedefinieerde lijst van *criteria* heeft en een *gestandaardiseerd scoresysteem* mist. Omwille van de *lage interbeoordelaars-overeenstemming* is SCAN onvoldoende ontwikkeld als forensisch instrument en daarom moet het gebruik van SCAN in de praktijk *ontmoedigd* worden. Om tot deze conclusies te komen analyseerden zij 82 verklaringen aan de Nederlandse politie van slachtoffers, verdachten en getuigen waarin SCAN werd toegepast. Twee van de vijf auteurs evalueerden, als onafhankelijke codeurs, of de ‘criteria’ die de politie Amsterdam gebruikte, aanwezig waren in de verklaringen. Ze analyseerden niet de verklaringen zelf, maar onderzochten slechts de codeerschema’s die de SCAN-analisten produceerden voor elke verklaring. De resultaten toonden ook aan dat SCAN voornamelijk zou gedreven worden door twaalf ‘criteria’, die als uitgangspunt kunnen dienen voor toekomstig onderzoek op hun psychometrische eigenschappen.

2.1. Basis van hun onderzoek

De basis van hun onderzoek is een ‘pilot’ over het gebruik van de SCAN, die de politie van Amsterdam in 2006 uitvoerde op 118 geschreven verhoren en documenten in zedendisdrijven (BETHLEM, OOSTING & TEERNSTRA, 2006). Aan vermeende slachtoffers, getuigen en verdachten werd bij aankomst op het politiebureau – voordat een ondervraging gebeurde – gevraagd op te schrijven wat hun versie van het gebeurde was. Deze verklaringen werden dan voor een SCAN ingediend waarna de resultaten van de analyse konden verwerkt worden in een navolgend verhoor.

Het doel van deze ‘pilot’ was te onderzoeken of een verbale analysemethode zoals SCAN kan bijdragen aan de *waarheidsvinding* in opsporingsonderzoeken en zo ja, op welke wijze (BETHLEM, OOSTING & TEERNSTRA, 2006: 5). De auteurs van het wetenschappelijk onderzoek schrijven dat het doel van de ‘pilot’ was te onderzoeken in hoeverre SCAN zou kunnen helpen in het *determineren van de geloofwaardigheid* van slachtoffers, getuigen en verdachten. Determineren van *geloofwaardigheid* is veel beperkter dan bijdragen aan *waarheidsvinding*, waar het niet noodzakelijk om geloofwaardigheid of misleiding gaat. Waarheidsvinding kan bijvoorbeeld gaan over het gedetailleerd bevragen van iemand die opening van zaken geeft betreffende alle elementen die met de feiten te maken kunnen hebben. De auteurs veranderen de doelstelling van een uitgevoerd onderzoek in de zin van hun eigen onderzoek.

Het ‘pilot-team’ bestond uit vier SCAN-analisten die een paar maanden voordien de 3-daagse cursus SCAN voor beginners bij Avinoam SAPIR hadden voltooid. Bij de start van het project had geen van hen ervaring met het toepassen van SCAN. Eén van de SCAN-analisten had ook de 2-daagse cursus voor gevorderden gevolgd. Een volledige opleiding SCAN-analist bestaat inderdaad uit vijf volledige dagen. Alhoewel de auteurs dit moesten weten – een van hen volgde ook de 3-daagse opleiding voor beginners – vermeldden ze dit niet. Vermits geen van de vier SCAN-analisten enige ervaring had met het toepassen van SCAN, is dit verschil in opleiding nochtans van groot belang. Een analyse van een onervaren *gedeelte*lijk opgeleide scanner zal er anders uitzien dan van een *volledig* opgeleide onervaren scanner, laat staan van een ervaren en volledig opgeleide scanner. Dit maakt dat de resultaten van de interbeoordelaars-overeenstemming alleen al op dit punt niet juist kunnen zijn. De ‘pilot’ vermeldt dat ‘in principe’ elk document door *minimaal twee* scanners, onafhankelijk van elkaar, werd geanalyseerd. Sommige documenten zijn dus door *meerdere* scanners geanalyseerd (BETHLEM, OOSTING & TEERNSTRA, 2006: 6). De auteurs schrijven affirmatief dat alle verklaringen van de ‘pilot’ door *twee van de vier* scanners werden geanalyseerd, zodat men de vraag kan stellen op welke criteria zij slechts twee analyses weerhielden in de gevallen waar er meer waren.

Overeenkomstig de SCAN-richtlijnen die een ‘zuivere versie’ benadrukken als uitgangspunt voor analyse, en om de set van verklaringen homogeen te maken, namen de auteurs 82 gevallen in hun analyse op. De ‘pilot’ heeft het over slechts 77 bruikbare ‘zuivere versies’ (BETHLEM, OOSTING & TEERNSTRA, 2006: 11). Misschien hebben de auteurs in hun onderzoek vijf onbruikbare versies meegenomen.

Twee van de vijf auteurs hebben de analyses van de ‘pilot’ gecodeerd. Eén codeur volgde de 3-daagse SCAN cursus voor beginners, de andere codeur las de SCAN-cursus en was getrouwd met de SCAN literatuur en het ‘codeerschema’ (sic), waarvan de auteurs later schrijven dat het in SCAN niet aanwezig is. De andere drie auteurs van een artikel over SCAN, kennen de SCAN dus niet. De resultaten van de twee codeurs voor de aan of aanwezigheid van SCAN-kenmerken hadden vaak een aanzienlijke afwijking van 50%. Dit duidt volgens de auteurs op een zeer onregelmatige dataset. Er werd niet onderzocht of de reden van de grote verschillen niet kan liggen in het verschil in competenties tussen de twee codeurs. SCAN is immers een *vaardigheid*, zoals men niet kan tennissen door een boek erover te lezen kan men ook niet scannen door enkel het (beperkt) cursusboek te lezen. Zo scoorden de auteurs *enkel* het belangrijk werkwoord ‘leaving’ (verlaten) van een plaats, terwijl ook synoniemen van dit werkwoord (zoals weggaan, vertrekken, ...) moesten gescoord worden (SAPIR, 2005, nr. 273-275). Een 5-daagse SCAN-opleiding is zeer intensief, met talrijke oefeningen, toepassingen en zelfs met thuiswerk.

2.2. Afwezigheid van goed gedefinieerde, gestandaardiseerde ‘criteria’

De auteurs gebruiken de term ‘criterium’ (Engels: criterion), die in de cursustekst van SAPIR (2005) nergens voorkomt. SAPIR heeft het telkens over ‘regels (rules)’¹, die veel ruimer kunnen geïnterpreteerd worden dan criteria². SCAN pretendeert nergens te bestaan uit ‘gestandaardiseerde criteria’ op basis waarvan *een beslissing* wordt genomen. Integendeel heeft de cursustekst het doorlopend over *verhoren*, hoe goede vragen te stellen. Vooral, maar niet uitsluitend, als gevolg van de gevonden hotspots (kritische punten) bij de analyse van de zuivere eigenhandig geschreven versie. Zoals reeds aangegeven bestaat de SCAN uit zes stappen. De auteurs hebben het slechts over één stap daarvan: de analyse. De resultaten van een SCAN moeten gezien worden als een aanvulling op – en niet als een substituuut – van bevindingen uit een verhoor. De resultaten van een SCAN zijn dus geen ‘direct bewijs’ noch een ‘steunbewijs’ in de rechtszaal. SCAN is enkel een *hulpmiddel* om verhoren beter, vollediger, te maken, geen zelfstandig opsporingsmiddel. In die zin zijn gedefinieerde gestandaardiseerde ‘criteria’ overbodig. De aandachtspunten die de verhoorder bij de voorbereiding van een traditioneel verhoor weerhoudt, kunnen ook onmogelijk gestandaardiseerd zijn.

2.3. Lage interbeoordelaars-overeenstemming

Om de betrouwbaarheid van ‘SCAN-criteria’, zoals de auteurs ze noemen, te evalueren berekenden ze de exacte frequentie van elk ‘criterium’. Hoewel de scanners *getraind* waren in het gebruik van SCAN, was hun interbeoordelaars-overeenstemming teleurstellend laag schrijven ze en *daardoor* is de SCAN-methode onvoldoende ontwikkeld als forensische tool.

Zoals wetsartikelen *regels* zijn die de rechter moet toepassen op concrete situaties, zijn de SCAN-regels dat ook. Het is dus normaal dat er verschillen in beoordeling zijn, afhankelijk van de ervaring van de politiemans als verhoorder en als scanner. De scanners van de politie Amsterdam hadden *geen* ervaring met de toepassing van de SCAN en hadden slechts een paar maand voordien de 3-daagse cursus SCAN voor beginners gevolgd. De resultaten zouden helemaal anders geweest zijn indien ervaren scanners, met een volledige 5-daagse opleiding, de analyses hadden gedaan. Niet bevorderend voor de objectiviteit van de publicatie is dat men de vier scanners in de discussie plots *getrainde analisten* in het gebruik van SCAN noemt, terwijl ze eerder in hetzelfde artikel beschreven worden als onervaren scanners.

De CBCA/SVA³, een andere schriftelijke analysemethode door psychologen ontworpen met het doel ware uit onware verhalen te onderscheiden, heeft *wel* gedefinieerde, gestandaardiseerde criteria (RASSIN, MERCKELBACH & CROMBAG, 1997). Problematisch bij CBCA is echter dat er geen vast scoringsstelsel bestaat, dat er geen cut-off score voorhanden is en dat de scores moeten geïnterpreteerd worden in het licht van de Validity Checklist (SVA) (VERSCHEURE, MEIJER & VRIJ, 2010: 712). Uit de praktijk komt inderdaad naar voor dat de werkwijze van de deskundigen die gebruik maken van deze methode onderling sterk verschillend is en dat nog zal moeten blijken dat verschillende beoordelaars de criteria ook op een consistente wijze gebruiken (NIEROP, VAN DEN ESHOF & BRANDT, 2006). Ook volgens MERCKELBACH, één van de auteurs in een ander artikel, heeft de CBCA slechts voor enkele criteria een voldoende interbeoordelaars-overeenstemming (RASSIN, MERCKELBACH & CROMBAG, 1997; ANSON, GOLDING

¹ VAN DALE, Engels-Nederlands. Rule: (1) voorschrift; (2) gebruik.

² VAN DALE, Engels-Nederlands. Criterion: toetssteen, maatstaf, standaard, doorslaggevend/bepalend/beslissend kenmerk/argument.

³ CBCA: Criteria Based Content Analysis. SVA: Statement Validity Assessment (KÖHNKEN, 2004; BOCKSTAELE, RAEYMAEKERS & ENHUS, 2007: 209-241).

& GULLY, 1993). Het was objectief geweest dit te vermelden want geen enkele schriftelijke analysemethode – ook niet deze ontworpen door psychologen – voldoet aan de ‘eis’ van betrouwbare, gedefinieerde gestandaardiseerde criteria zoals de auteurs dit vooropstellen.

Wetenschappelijk onderzoek over de CBCA/SVA gaat ook meestal over de CBCA-criteria en niet over de meeromvattende SVA, waarvan de CBCA-criteria slechts een onderdeel uitmaken. Academics maken er zich van af door te vermelden dat SVA buiten beschouwing wordt gelaten omdat: ‘dit onderdeel nauwelijks gesystematiseerd zou zijn’ (RASSIN, MERCKELBACH & CROMBAG, 1997). Een techniek op slechts een onderdeel beoordelen, zoals ook het wetenschappelijk onderzoek over de SCAN, is niet correct.

De conclusie van de ‘pilot’ was dat er na het uitvoeren van 118 SCANS *weinig onderlinge* verschillen bleken tussen de scanners die dezelfde tekst hadden geanalyseerd. De auteurs geven geen uitleg hoe dit verschil met *hun* conclusies kan verklaard worden.

2.4. Aantal ‘criteria’

De auteurs bekritisieren dat teksten over SCAN niet hetzelfde aantal ‘criteria’ vermelden. Dat is te verklaren omdat de oorspronkelijke SCAN-cursustekst het niet heeft over *criteria*, maar over *regels* betreffende de analyse van verklaringen en ook over *regels* betreffende de vragen die moeten gesteld worden op basis van de gemaakte analyse. Voor de ene verhoorder heeft een regel betrekking op de analyse, voor de andere op het verhoor, naargelang zijn eigen verhoor of SCAN-ervaring en het misdrijf waarover verklaard wordt. De analyse van een verklaring en het verhoor daarover, ingebed in een traditioneel verhoor, vormen een geheel. De auteurs bespreken de meerwaarde van het geheel niet.

De auteurs vermelden:

- ‘plaats van emotie’ als een ‘SCAN-criterium’. Dit ‘criterium’ komt *niet* voor in de SCAN-basistekst (SAPIR, 2005);
- dat ‘spontane correcties’ een ‘SCAN-criterium’ is, vergelijkbaar met criterium 14 van de CBCA. ‘Spontane correcties’ volgens CBCA zou wijzen op waarachtigheid, terwijl dit volgens de SCAN op bedrog zou wijzen. Dit staat niet in de SCAN-basis cursus (SAPIR, 2005), wél dat als de *rechercheur* een fout maakt en de persoon ze niet verbeterde, dit wijst op misleiding (SAPIR, 2005, nr. 210). Verandering van voornaamwoorden in *dezelfde zin* (bv. ‘ik, euh wij’) kan ook wijzen op misleiding in een verklaring (SAPIR, 2005, nr. 267a). Deze nuances leggen de auteurs niet. Meer nog, het CBCA-criterium ‘spontane correcties’ bleek volgens een van de auteurs in een ander artikel (WINKEL & VRIJ, 1995) *ondeugdelijk* om het verschil tussen waar en gefingeerd te maken (RASSIN, MERCKELBACH & CROMBAG, 1997);
- dat ‘gebrek aan overtuiging of het geheugen’ een ‘SCAN-criterium’ is dat *dicht* bij CBCA criterium 15 met dezelfde naam ligt en dat zij tegengestelde conclusies zouden hebben. Deze vermelding is fout, de auteurs gaan er aan voorbij dat een SCAN betrekking heeft op een eigenhandig geschreven verklaring waarbij de verhoorder geen enkele vraag mag stellen, terwijl het verzamelen van CBCA-criteria het gevolg is van een interview (VERSCHEURE, MEIJER & VRIJ, 2010: 710). Dus is het eerste geval komt de verklaring *spontaan* uit het geheugen – hoe kan iemand zich een handeling *niet* herinneren als hem er niet naar gevraagd wordt? (SAPIR, 2005, nr. 313-314) – en in het andere geval zijn het *antwoorden op vragen*. Minstens twee van de auteurs weten dat, maar vermelden het niet, terwijl dat punt een groot verschil in beoordeling uitmaakt. Bovendien, het CBCA-criterium ‘toegeven zich iets niet meer te herinneren’, bleek ook volgens een van de auteurs in een ander

artikel (WINKEL & VRIJ, 1995) *ondeugdelijk* om het verschil tussen waar en gefingeerd te maken (RASSIN, MERCKELBACH & CROMBAG, 1997).

- dat de ‘regel van de zuinigheid’ (The rule of thrift, SAPIR, 2005, nr. 162m en 458-460), door de auteurs ‘out of sequence information’ genoemd, tegengesteld is aan het overeenstemmende CBCA-criterium 9, ‘overbodige details’. Dat is fout, beide regels wijzen in dezelfde richting: een oprecht verhaal zal meer ‘onnodige verwijzingen’ hebben. Bovendien is eerder CBCA-criterium 7 (onverwachte complicaties), in plaats van criterium 2 (ongestructureerde reproductie) te vergelijken met de SCAN-regel.

2.5. De meest voorkomende ‘criteria’

De studie onderzocht welke ‘SCAN-criteria’ *het meest* worden gebruikt *in het veld*. Dit moet eigenlijk zijn: welke SCAN-criteria *in deze ‘pilot’* het meest werden gebruikt door vier *onervaren en niet volledig opgeleide* SCAN-analisten. Zij kwamen tot een lijst van 12 criteria die werden gebruikt in meer dan 25% van de gevallen. Het is normaal dat rechercheurs een voorkeur hebben voor bepaalde kritische punten, naargelang het misdrijf waarover het verhoor gaat en naargelang hun verhoorervaring en ervaring met SCAN. De auteurs vermelden niet de lengte van de geschreven teksten. De langste tekst bestond uit 244 regels, de kortste uit 3. De gemiddelde lengte van de verklaringen was 50 regels. Het ligt voor de hand dat in korte verklaringen niet zoveel ‘criteria’ kunnen voorkomen dan in lange. Nochtans kan in de kortste verklaring van 3 regels één hotspot voorkomen die zeer belangrijk is en die niet in andere verklaringen voorkomt. De auteurs hebben geen rekening gehouden met de verhouding tussen het aantal regels van elke verklaring en het aantal ‘criteria’ daarin.

2.6. Foute conclusie

Onverdedigbaar is de conclusie van de auteurs dat, omwille van een te lage interbeoordelaars-overeenstemming, het gebruik van SCAN in de praktijk moet ontmoedigd worden. De resultaten van de ‘pilot’ waren (BETHLEM, OOSTING & TEERNSTRA, 2006: 13):

- in 33 gevallen leidde SCAN tot het stellen van aanvullende vragen in het vervolghoort, waardoor de verklaring vollediger werd;
- in 7 gevallen paste de betrokkene na SCAN-vragen zijn verklaring aan;
- SCAN gaf in 64 zaken meer inzicht in de betrokken persoon en de onderlinge relaties met de overige betrokkenen in een zaak;
- 68 keer werden de vermoedens van rechercheurs in een zaak door de SCAN-rapportage bevestigd. De meerwaarde daarvan is dat SCAN vanuit een heel andere invalshoek, puur taaltechnisch, tot deze bevindingen kwam;
- naast bevestiging, leidde de SCAN-rapportage in 25 gevallen tot nieuwe gezichtspunten in het opsporingsonderzoek, omdat de bevindingen vanuit de analyse niet, of niet volledig overeen kwamen met de bevindingen van de rechercheur;
- SCAN leidde op basis van de voornoemde nieuwe gezichtspunten in 5 gevallen tot een wijziging in de opsporingsrichting;
- in 7 zaken heeft de SCAN geen meerwaarde gehad;
- bij 5 documenten is geen analyse uitgevoerd;
- in 1 zaak droeg SCAN bij aan de herkenning van een vrijwillige valse bekentenis.

Het is indrukwekkend dat scanners, net na een cursus SCAN voor beginners, en zonder ervaring ermee in de praktijk, tot deze meerwaarde voor de waarheidsvinding komen.

3. DRIE BELGISCHE STUDIES

Ook in België worden twijfelachtige ‘wetenschappelijke’ onderzoeken geproduceerd. Zo werden op het 4^{de} internationaal symposium over verhoortechnieken van 28 juni t/m 1 juli 2010 te Brussel drie studies inzake de betrouwbaarheid en de validiteit van SCAN gepresenteerd (VANDERHALLEN, JASPAERT & VERVAEKE, 2010).

De **eerste studie** (VANDERHALLEN, JASPAERT & VERVAEKE, 2010) onderzocht de accuraatheid van SCAN als *leugendetectie-instrument*, waarvoor het niet ontworpen is en in België door de politie in die zin niet gebruikt wordt. Er werd in de studie een vergelijking gemaakt tussen studenten rechten, politiemensen zonder SCAN-training en politiemensen met SCAN-training. De drie groepen dienden vier verklaringen (2 leugenachtig en 2 waarheidsgetrouw) over de betrokkenheid bij een *verkeersongeval* te kwalificeren als waarheidsgetrouw of leugenachtig⁴. De studie toonde aan dat politiemensen met SCAN-training beter scoren in de detectie van leugenachtige verklaringen in vergelijking met studenten. Echter, studenten en politiemensen zonder SCAN-training bleken beter in het detecteren van de waarheid. Het onderzoek maakte geen onderscheid tussen liegen en misleiding, wat voor SAPIR nochtans essentieel is omdat liegen veel meer spanning veroorzaakt (SAPIR, 2005: 8 en 51).

Ook hier is de studie beperkt tot slechts één fase van SCAN: de analyse. De respondenten werden geconfronteerd met vier teksten waarvan zij niet wisten hoe die tot stand waren gekomen, bv. welke instructies gegeven werden aan de schrijvers. De analyse is slechts een *tussenstap* ter voorbereiding van het er opvolgend verhoor. In een ‘wetenschappelijk’ onderzoek aan respondenten vragen⁵, op basis van *die ene fase*, zich uit te spreken of een verklaring leugenachtig of waarheidsgetrouw is (onbeslist was zelfs geen keuzemogelijkheid!) is niet verantwoord en geeft een verkeerd beeld van de SCAN-techniek.

De **tweede studie** (VAN GEEST, 2007-2008), onderzocht de mogelijkheid om SCAN te gebruiken als een middel bij de voorbereiding van een verhoor. Uit deze studie zou blijken dat SCAN-verhoorders niet meer of betere vragen stellen dan andere verhoorders. De onderzoeksgroep bestond uit slechts twaalf verhoorders zonder SCAN-opleiding, wiens analysere-sultaten vergeleken werden met de analyse van één (1) volwaardig opgeleide SCAN-analist en een reserve-analist. Dit laag aantal is in strijd met methodologische handleidingen over wetenschappelijke onderzoeken (LEARY, 2004). De resultaten van één analist als maatstaf nemen voor alle anderen maakt het onderzoek waardeloos. Slechts de enige referentiepersoon volgde de volledige SCAN-opleiding⁶. Vijf van de twaalf respondenten volgden de cursus *rechercheverhoortechnieken* (RVT) aan de Nationale Rechercheschool, waarin destijds ook gedurende *een halve dag* de SCAN-techniek als *kennismaking* werd uitgelegd. Het onderzoek maakt evenmin onderscheid tussen liegen en misleiding.

De onderzoeksvraag luidde: *‘In welke mate stelt men met behulp van de SCAN-methode meer en betere vragen voor verder verhoor’*. De onderzoeksvraag werd onderverdeeld in drie subvragen:

1° *Detecteren niet-scanners dezelfde hotspots?* Alle twaalf ervaren verhoorders-respondenten hebben de hen toegestuurde twee teksten afkomstig uit reële zaken *in het algemeen* geanalyseerd, zoals het rechercheurs met gemiddeld 22 jaar verhoorervaring betaamt.

⁴ Verklaring 1: verkeersongeval met auto zonder gekwetsten op het rondpunt te Wijnegem; Verklaring 2: paardrijdster wordt aangereden door auto; Verklaring 3: verkeersongeval met auto zonder gekwetsten aan afrit autostrade; Verklaring 4: verkeersongeval met auto zonder gekwetsten i.v.m. voorrang van rechts.

⁵ Cf. brief d.d. 2 november 2007 van Miet VANDERHALLEN aan de respondenten met SCAN-opleiding.

⁶ Drie dagen voor de basisopleiding en twee dagen voor de voortgezette opleiding.

Uit het onderzoek blijken voor verklaring 1: 79 hotspots en voor verklaring 2: 42 hotspots, zonder onderscheid te maken tussen *algemene hotspots* bekomen na een globale analyse en hotspots bekomen na toepassen van de SCAN-regels. Nochtans blijkt uit de studie duidelijk dat er in beide verklaringen belangrijke hotspots waren die *uitsluitend* door politiemensen met SCAN-kennis werden gedetecteerd. De onderzoekster merkt dit op en schrijft: *'Dit bevestigt het vermoeden dat zij veel beter weten naar waar ze moeten kijken in een verklaring'* (VAN GEEST, 2007-2008: 7). Wat niet bekend is wordt niet uitgeplozen en blijft dus onbekend voor de waarheidsvinding. Dit nochtans cruciaal gegeven werd niet verder uitgewerkt.

- 2° *Stellen SCAN-analisten méér vragen dan niet-SCAN-analisten?* Om dit te onderzoeken werd gebruik gemaakt van vijf van de eeuwenoude 7-W vragen (Wie, Wat, Waar, Wanneer en Waarom). Dit heeft niets met SCAN te maken, de 7-W vragen zijn zodanig ingeburgerd dat elke verhoorder er overweg mee kan. Het zijn vragen waarover de verhoorder *in het algemeen* moet doorvragen, niet uitsluitend over de hotspots die middels de SCAN-techniek worden gevonden. Het is dus normaal dat er geen verschillen gevonden werden tussen collega's met en zonder kennis van de SCAN-techniek.
- 3° *Stellen SCAN-analisten betere vragen dan niet-SCAN-analisten?* Met 'betere vragen' worden in het onderzoek open, niet suggestieve vragen bedoeld (VAN GEEST, 2007-2008: 59 en 86). Dat gesloten vragen ook zinvol kunnen zijn komt niet aan de orde, evenmin als niet voor de hand liggende vragen. De respondenten kregen als opdracht om voor elke hotspot, ook deze als gevolg van een algemene analyse, één of meerdere vragen als voorbereiding op het verhoor te formuleren. In de studie werd dan onderzocht of die vragen open en niet suggestief waren. De meeste verhoorders maken hun voorbereiding op een verhoor in telegramstijl, zo ook de suggesties voor vragen als gevolg van een SCAN-analyse. Dit betekent dus niet dat de verhoorder die vragen zal stellen onder de vorm dat ze zijn neergeschreven. Omdat deze subvraag niets met SCAN te maken heeft is het ook hier normaal dat er geen verschillen gevonden werden tussen collega's met en zonder kennis van de SCAN-techniek. De eindconclusie van dit onderzoek dat *SCAN geen meerwaarde biedt om meer of betere vragen te stellen*, is derhalve fout.

De **derde studie** (FLEMMINGS, 2008-2009) poneert dat, als analisten de criteria verschillend hanteren dit kan leiden tot potentieel verschillende interpretaties van verklaringen (FLEMMINGS, 2008-2009: 38). De studie vermeldt: *'Het gaat hier om een instrument op basis waarvan belangrijke beslissingen over individuen genomen kunnen worden. Het mag daarom voor de resultaten van een SCAN niet uitmaken of de methode door analist A of B wordt uitgevoerd'* (FLEMMINGS, 2008-2009: 42). Dit uitgangspunt is verkeerd. Op basis van een SCAN-analyse worden geen belangrijke beslissingen genomen, maar worden integendeel vragen gegeneereerd die bijdragen tot de waarheidsvinding tijdens een eropvolgend verhoor.

De onderzoeksvraag was: *'wat is de interbeoordelaarsbetrouwbaarheid van elk SCAN-criterium binnen een groep ervaren scanners, een groep onervaren scanners, een groep ervaren rechercheurs en binnen een groep politiemensen in opleiding (novicen).'* Elk van die vier groepen bestond uit amper vier (4) respondenten! Voor de groep van de ervaren scanners gold als selectiecriterium dat ze een cursus SCAN moesten gevolgd hebben bij SAPIR of door een lesgever die bij SAPIR de cursus volgde. Dit betekent dat er bij de groep van de vier ervaren scanners misschien geen enkele politieambtenaar was die de volledige cursus bij SAPIR volgde (3 + 2 dagen) en slechts kennis heeft van de module SCAN in de cursus RVT, die slechts een halve dag duurde. Op basis daarvan mag men de mogelijkheden van SCAN niet afrekenen. Bovendien, met amper vier respondenten per groep heeft deze studie ook methodologisch geen waarde.

Bovenstaande ondeugdelijke onderzoeken werden op internationale symposia over verhoortechnieken de wereld ingestuurd⁷. Politie en rechtspractici doen geen wetenschappelijke onderzoeken en publiceren amper over deze onderwerpen, zodat enkel de mening van academici in die materie wordt gelezen. Tekortkomingen worden zelden rechtgezet in andere publicaties of op andere symposia, de conclusies gaan een eigen leven leiden en dienen ten slotte als (foute) argumentatie in verhoorkamers en opleidingen. Verhoorders en rechtspractici, waaronder advocaten die op verhoren toezien, moeten zich daarvan bewust zijn.

4. CONCLUSIES

Bij het lezen en in rekening brengen van publicaties over rechtspsychologisch onderzoek moeten politie en rechtspractici zeer voorzichtig zijn en, zoals uit bovenstaande blijkt, absoluut rekening houden met subjectieve, onvolledige, onzorgvuldige en zelfs foutieve rapportages (SIMMONS, NELSON & SIMONSOHN, 2011). Meer nog, BAKKER en WICHERTS (2011) onderzochten 281 artikels uit psychologische tijdschriften, zowel low als high-impact, en vonden dat uit de erin besproken *correcte statistieken* liefst 15% *verkeerde conclusies* getrokken waren!

Sommige rechtspsychologische publicaties inzake verhoren hebben alleen de meerwaarde dat ze kunnen meetellen voor de quota die academici moeten halen. Rechtspsycholoog professor MERCKELBACH (2011) schreef dat politieagenten en magistraten kritischer mogen zijn tegen zijn vakgenoten, een oproep waaraan we nu gevolg geven. Het is goed dat de meeste van de paar honderd academische publicaties over verhoren per jaar niet bekend raken bij verhoorders en rechtspractici, sommige ontnemen de verhoorders bruikbare tools. Zelfs zeer ervaren verhoorders zijn na amper een halve dag uitleg over de SCAN, verbaasd dat zij in een verklaring meer te bevragen punten en gespreksonderwerpen ontdekken. Daar is het SCAN om te doen: met het oog op waarheidsvinding, na toekenning van de 'Salduzrechten', op een legale wijze zoveel mogelijk bruikbare informatie halen uit een verhoorde.

REFERENTIES

- ADAMS, S. & JARVIS, J. (2004). Are you telling me the truth? – indicators of veracity in written statements. *FBI Law Enforcement Bulletin*, oktober, 7-12.
- ADAMS, S.H. & HARPSTER, T. (2008). 911 Homicide calls and statement analysis. Is the caller the killer? *FBI Law Enforcement Bulletin*, juni, 22-30.
- ANSON, D.A., GOLDING, S.L. & GULLY, K.J. (1993). Child sexual abuse allegations: Reliability of Criteria Based Content Analysis. *Law and Human Behavior*, 17, 331-341.
- BAKKER, M. & WICHERTS, J.M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods*, 43, 666-678.
- BETHLEM, T., OOSTING, J. & TEERNSTRA, M. (augustus 2006). *Gezocht: De waarheid. Eindrapportage waarheidsvinding via woordanalyse*. Niet gepubliceerd.
- BOCKSTAELE, M. (2015). Over het 'gezag' van (rechts)psychologen in verband met verhoren. *Tijdschrift voor Strafrecht*, 1, 2-13.
- BOCKSTAELE, M., RAEYMAEKERS, S. & ENHUS, E. (2007). *Leugens en hun detectie*. Antwerpen, Maklu.
- BOGAARD, G., MEIJER, E.H., VRIJ, A., BROERS, N.J. & MERCKELBACH, H. (2014). SCAN is largely driven by 12 criteria: results from sexual abuse statements. *Psychology, Crime & Law*, 20/5, 430-449.

⁷ Symposium te Sorrento van 2 t/m 5 september 2009 en te Brussel van 28 juni t/m 1 juli 2010.

- BREWER, W.F., SAMPAIO, C. & BARLOW, M.R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory and Language*, 52, 618-627.
- BURKE, S. (2013). Vetting confidential human source intelligence through investigative statement analysis. *FBI Law Enforcement Bulletin*, December, 1-3.
- FLEMINGS, I. (2008-2009). *De interbeoordelingsbetrouwbaarheid van de SCAN-criteria*. Leuven: KUL.
- HUYBRECHTS, L. (2011). Wat is een verhoor vraagt U? In M. BOCKSTAELE, E. DEVROE & P. PONSAAERS (eds.), *Salduz – Bijstand van advocaten bij verhoren*. Antwerpen, Maklu, 53-66.
- KÖHNKEN, G. (2004). Statement validity analysis and the detection of the truth. In P.A. GRANHAG & L.A. STRÖMWALL (eds.), *The detection of deception in forensic contexts*, Cambridge, Cambridge University Press, 41-63.
- LEARY, M.R. (2004). *Introduction to behavioral research methods*. London: Pearson.
- MATSUMOTO, D., HWANG, H.S. & FRANK, M.G. (2012). The role of emotion in predicting violence. *FBI Law Enforcement Bulletin*, januari, 1-7.
- MERCKELBACH, H. (2011). Er zijn veel simulanten. *Blauw*, 17, 12-15.
- NIEROP, N.M., VAN DEN ESHOF, P. & BRANDT C. (2006). De beoordeling van geloofwaardigheid in zedenzaken. *Nederlands Juristenblad*, 43, 2456-2464.
- RASSIN, E., MERCKELBACH, H. & CROMBAG, H. (1997). De Criteria Based Content Analysis (CBCA) als instrument om de geloofwaardigheid van getuigenverklaringen te betalen. *Nederlands Juristenblad*, 42, 1923-1929.
- SANDOVAL, V.A. (2008). Interview clues. Words that leave an investigative trail. *FBI Law Enforcement Bulletin*, januari, 1-9.
- SAPIR, A. (2005). *The course on Scientific Content Analysis (SCAN)*, Laboratory for scientific interrogation. Phoenix, Arizona.
- SCHAFER, J.R. (2008). Text bridges and the micro-action interview. *FBI Law Enforcement Bulletin*, januari, 20-24.
- SIMMONS, J.P., NELSON, L.D. & SIMONSOHN, U. (2011). False-Positive Psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11) 1359-366.
- SMITH, S.S. & SHUY, R.W. (2002). Forensic psycholinguistics. Using language analysis for identifying and assessing offenders. *FBI Law Enforcement Bulletin*, april, 16-21.
- VAN GEEST, E. (2007-2008). *Misleiding in verklaringen: De SCAN-techniek*. Leuven, KUL.
- VANDERHALLEN, M., JASPAERT, E. & VERVAEKE, G. (2010). Het verhaal van (misleidende) verdachten in geschreven verklaringen. Brussel, *Symposium verhoortechnieken*, 28 juni t/m 1 juli 2010.
- VERSCHUERE, B., MEIJER, E.H. & VRIJ, A. (2010). Verbale leugendetectie. In P.J. VAN KOPPEN, H. MERCKELBACH, M. JELICIC & J.W. DE KEUSER (eds.), *Reizen met mijn rechter*, Deventer, Kluwer, 705-722.
- WELLS, R.C. (2008). The art of investigative interviewing. Countering the lie of omission. *FBI Law Enforcement Bulletin*, januari, 10-12
- WINKEL, F.W. & VRIJ, A. (1995). Verklaringen van kinderen in interviews: Een experimenteel onderzoek naar de diagnostische waarde van de Criteria Based Content Analysis (CBCA). *Tijdschrift voor ontwikkelingspsychologie*, 22, 61-74.