

Tevens per fax: 020 - 541 27 45

Aan de Rechtbank Amsterdam
Sector Bestuursrecht
Postbus 75850
1070 AW AMSTERDAM

datum 20 november 2007
dossier 20060024B
e-mail evdbogaard@westhoff.nl
uw referentie 06 / 5066 WAO
inzake **A. / UWV**

Aanvullend beroepschrift

- 1 Inleiding
- 2 Beoordeling van psychische problematiek: strijd met artikel 3:2 en 3:4 Awb, artikel 18 lid 1 WAO en artikel 4 lid 1 en 2 van het Schattingsbesluit
 - 2.1 Inleiding
 - 2.2 De groeiende kloof tussen verzekeringsgeneeskunde en "evidence-based" geneeskunde
 - 2.3 Wettelijk kader
 - 2.4 Klinische blik bij psychische klachten: berucht onbetrouwbaar
 - 2.5 Algemeen aanvaarde verzekeringsgeneeskundige onderzoekmethoden
 - 2.6 Bijzondere problemen bij - retrospectieve - toetsing in bezwaar en beroep
- 3 Wanneer dient de rechter een deskundige in te schakelen?
 - 3.1 Twijfel
 - 3.2 Wat geeft aanleiding tot twijfel?
 - 3.3 Materiële inhoud
 - 3.4 Wijze van totstandkoming
- 4 Bijzonderheden in de zaak A.
 - 4.1 Consistentie van het medisch oordeel in de loop der tijd
 - 4.2 Sociale omstandigheden
- 5 Conclusie

1 Inleiding

- 1.1 In het inleidend beroepschrift van 17 oktober 2006 is al aangegeven dat aanleiding bestaat om een expertise te doen verrichten. In dit aanvullend beroepschrift wordt dit verzoek nader onderbouwd. Daarbij zal worden ingegaan op verschillende aspecten van de zorgvuldigheid van de verzekeringsgeneeskundige beoordeling.
- 1.2 In paragraaf 2 wordt in meer algemene zin ingegaan op de bijzondere problemen zich voordoen bij de beoordeling van psychische problematiek. Aan de hand van internationale wetenschappelijke literatuur, die gemakshalve ook als bijlage bij dit aanvullend beroepschrift wordt overgelegd, zal worden aangetoond dat de wijze waarop verweerder omgaat met psychische problematiek van verzekerenden in strijd is met de eisen van zorgvuldigheid en belangenafweging van artikel 3:2 en 3:4 Awb, het objectiviteitsvereiste van artikel 18 lid 1 WAO, het reproduceerbaarheidsvereiste van artikel 4 lid 1 onder b van het Schattingsbesluit en met het vereiste van algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethoden van artikel 4 lid 2 van het Schattingsbesluit.
- 1.3 In paragraaf 3 wordt ingegaan op de vraag in welke situaties de rechter tot raadpleging van een deskundige moet worden overgegaan.
- 1.4 In paragraaf 4 wordt ingegaan op de bijzonderheden van de zaak van de heer A..

2 Beoordeling van psychische problematiek: strijd met artikel 3:2 en 3:4 Awb, artikel 18 lid 1 WAO en artikel 4 lid 1 en 2 van het Schattingsbesluit

2.1 Inleiding

- 2.1.1 In verzekeringsgeneeskundige rapportages waarin psychische problematiek een rol speelt, zijn vaak passages te lezen in de trant van "bij oriënterend psychiatrisch onderzoek geen (evidente) tekenen van (grove) psychopathologie". Ook is vaak een summiere beschrijving opgenomen van de "status praesens mentalis" zoals die op de verzekeringsarts is overgekomen. Van vragenlijsten wordt doorgaans geen gebruik gemaakt.

In zijn algemeenheid kan worden gezegd dat het onderzoek naar (de gevolgen van) psychische problematiek door verweerders verzekeringsartsen (zeer) beperkt van opzet en inhoud is. Indien de psychische problematiek de hoofdklacht is, is dit niet anders - hetgeen toch eigenlijk bevreedt.

- 2.1.2 In de onderhavige zaak is de rapportage van de primaire arts A.K. van B. van 2 mei 2006 een mooi voorbeeld van een uiterst summier psychisch onderzoek:

“Psychische toestand:

Helder bewustzijn, georiënteerd in tijd en plaats. De stemming is normofoor met een normaal modulerend affect. Er is normaal oogcontact. Er is een normale inhoud en tempo van denken en spreken.” (B35.2)

- 2.1.3 Het gevolg hiervan is dat de beoordeling van psychische problematiek door verweerders verzekeringsartsen niet voldoet aan de eisen die daaraan blijkens internationale wetenschappelijke literatuur behoren te worden gesteld. Daardoor handelt verweerder in strijd met artikel 3:2 en 3:4 Awb en in het bijzonder met hetgeen in artikel 18 lid 1 WAO en artikel 4 leden 1 en 2 van het Schattingsbesluit is bepaald.
- 2.1.4 Bij de uitwerking van deze stelling in het navolgende wordt in paragraaf 2.2 eerst ingegaan op het probleem dat de verzekeringsgeneeskundige structureel niet meegaat met de ontwikkelingen binnen de “evidence-based” geneeskunde. Vervolgens wordt in paragraaf 2.3 het wettelijk kader geschetst aan de hand waarvan de gevolgen hiervan dienen te worden beoordeeld. In paragraaf 2.4 wordt tegen de achtergrond van dit wettelijke kader ingegaan op de wetenschappelijke kritiek op de “klinische blik” als zelfstandig beoordelingsinstrument. In paragraaf 2.5 wordt hierbij de in het Schattingsbesluit geformuleerde eis van algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethoden betrokken. In paragraaf 2.6 wordt tot slot nog ingegaan op een bijzonder probleem bij de beoordeling van psychische problematiek in het kader van de bewaar- en beroepsprocedure.

2.2 De groeiende kloof tussen verzekeringsgeneeskundige en “evidence-based” geneeskunde

- 2.2.1 In zijn advies “Beoordelen, behandelen, begeleiden; Medisch handelen bij ziekteverzuim en arbeidsongeschiktheid” van 22 juli 2005, nr. 2005/10, blz. 27, heeft de Gezondheidsraad geconstateerd dat uit recente adviezen van de Raad voor Gezondheidsonderzoek over bedrijfsgeneeskunde en verzekeringsgeneeskunde blijkt dat veel van het medisch handelen bij ziekteverzuim en arbeidsongeschiktheid niet wetenschappelijk is te onderbouwen. De verzekeringsgeneeskunde blijkt erg op zichzelf te staan:

“Verzekeringsgeneeskundigen zijn, ook in de recent op gang gekomen multi-disciplinaire samenwerkingsprojecten, vrijwel niet betrokken bij de ontwikkeling van curatieve en bedrijfsgeneeskundige richtlijnen.

(...)

Omgekeerd lijkt de evidence based richtlijnontwikkeling aan de kwaliteitsbevordering in de verzekeringsgeneeskunde tot op heden voorbij te gaan." (blz. 42/43)

- 2.2.2 Ook binnen de verzekeringsgeneeskunde is men zich bewust van de groeiende kloof tussen de medische wetenschap en de verzekeringsgeneeskundige praktijk. Zo waarschuwen Van Dijk et al in het Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde van december 2003 ervoor dat het achterblijven van de kwaliteit van de dienstverlening door bedrijfs- en verzekeringsartsen niet acceptabel is:

"Arbodiensten en UWV, en niet in het minst de professionals zelf als verantwoordelijke deskundigen zullen alle zeilen bij moeten zetten om een inhaalslag te maken. Het verlenen van zorg en dienstverlening die niet up-to-date is, maakt arbodiensten en UWV buitengewoon kwetsbaar voor klachten en procedures over de geleverde kwaliteit, van de zijde van bedrijven, werknemers, verzekeraars, overheid en politiek, vooral als duidelijk is dat het veel beter kan.

(...)

De kloof met de wetenschappelijke wereld en met de technische mogelijkheden wordt snel groter. Arbodiensten en UWV lopen hierbij een groot risico van klachten en procedures betreffende de kwaliteit van de zorg en diensten."¹

- 2.2.3 De Gezondheidsraad wijst er in zijn advies echter ook op dat het medisch handelen in relatie tot arbeid zijn houvast ontleent aan dezelfde wetenschappelijke kennis die het medisch handelen in het algemeen ten dienste staat – met name die over ziektemechanismen, over diagnostiek en over de effecten van behandelingen (blz. 27) en dat de verzekeringsarts die beschikbare kennis in zijn beoordeling zal moeten betrekken (blz. 34). Ook al is de wetenschappelijke basis waarop de verzekeringsarts kan steunen beperkt, hij kan steun ondervinden van de instrumenten en modellen voor bewijsvoering, informatieverzameling en outputspecificatie die in de loop der tijd ontwikkeld zijn (blz. 35).

- 2.2.4 Op grond van dit advies van de Gezondheidsraad dient te worden gesteld dat, indien met betrekking tot de te beoordelen medische problematiek evidence-based instrumenten bestaan ten behoeve van diagnostiek en behandeling, de verzekeringsarts ook gehouden is - zowel vanuit medische optiek als vanuit juridische optiek - om van die instrumenten gebruik te maken.

De consequentie hiervan is dat indien de verzekeringsarts geen gebruik maakt van wel beschikbare instrumenten, hij onzorgvuldig - want in strijd met de medische standaard - handelt en dat een besluit dat op een dergelijk gebrekkig onderzoek berust, gebrekkig is.

2.3 Wettelijk kader

- 2.3.1 Artikel 3:2 Awb bepaalt dat het bestuursorgaan bij de voorbereiding van een besluit de nodige kennis vergaart omtrent de relevante feiten en de af te wegen belangen.
- 2.3.2 Artikel 3:4 Awb bepaalt in lid 1 dat het bestuursorgaan de rechtstreeks bij het besluit betrokken belangen afweegt en in lid 2 dat de voor één of meer belanghebbenden nadelige gevolgen van een besluit niet onevenredig mogen zijn in verhouding tot de met het besluit te dienen doelen.
- 2.3.3 Artikel 18 lid 1 WAO bepaalt dat de arbeidsongeschiktheid als gevolg van ziekte, gebreken, zwangerschap of bevalling objectief medisch dient te worden vastgesteld. De Centrale Raad van Beroep vertaalt dit in vaste jurisprudentie met de formulering: "op medische gronden naar objectieve maatstaven gemeten".
- 2.3.4 Artikel 4 lid 1 onder b van het Schattingsbesluit arbeidsongeschiktheidswetten bepaalt dat een door een andere verzekeringsarts uitgevoerd verzekeringsgeneeskundig onderzoek tot dezelfde bevindingen en conclusies zal kunnen leiden.
- 2.3.5 Artikel 4 lid 2 van het Schattingsbesluit arbeidsongeschiktheidswetten bepaalt dat het verzekeringsgeneeskundige onderzoek dient te geschieden aan de hand van algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethoden.

2.4 Klinische blik bij psychische klachten: berucht onbetrouwbaar

- 2.4.1 De kern van het probleem bij de beoordeling van psychische problematiek schuilt daarin dat de verzekeringsgeneeskundige beoordeling van de psychische belastbaarheid doorgaans uitsluitend is gebaseerd op de "klinische blik" van de verzekeringsarts. Aanvullend onderzoek vindt meestal niet plaats.
- 2.4.2 Uit verschillende, als bijlagen bij dit aanvullend beroepschrift overgelegde, internationale publicaties blijkt echter dat de betrouwbaarheid van de klinische blik al meer dan 40 jaar voorwerp van onderzoek is en dat uit die onderzoeken consequent naar voren komt dat deze onbetrouwbaar is.^{2, 3, 4}

2.4.3 Op 4 en 5 oktober 2007 heb ik aan de Maastrichtse hoogleraar psychologie prof. dr. H.L.G.J. Merckelbach dit probleem voorgelegd. Zijn reactie en de daarbij meegezonden literatuur treft u ook als bijlagen. Zijn opvatting:

“ik denk dat de huidige procedure die VA’s volgen zeer aanvechtbaar is. Het is aantoonbaar riskant om op basis van een gesprek uit te maken of de klachten van een kandidaat uitkeringsgerechtigde al dan niet authentiek zijn. In de VS wordt het verzuim om in zulke situaties tests te gebruiken als een ernstige kunstfout gezien. En dat is terecht. Punt is namelijk dat er goede tests voor handen zijn waarmee de kans op vals positieven en vals negatieven valt te reduceren.”⁵

2.4.4 Het eerste probleem is dat de klinische blik per definitie subjectief is, zodat persoonlijke belangstelling van de verzekeringsarts en de chemie tussen arts en onderzochte een grote rol spelen. Dit wordt met zoveel woorden erkend door Laitinen-Krispijn en Nicolas in het Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde van april 2006:

“In de praktijk is de kwaliteit van de verzekeringsgeneeskundige beoordeling sterk afhankelijk van ervaring en belangstelling van de verzekeringsarts voor psychische klachten.

...

‘Het hangt in grote mate af van welke verzekeringsarts je spreekt wat de uitkomst is.’”⁶

Dit leidt ertoe dat:

- niet wordt voldaan aan de in artikel 18 lid 1 WAO gestelde eis dat de gevolgen van ziekte, gebreken, zwangerschap of bevalling objectief medisch worden vastgesteld;
- zeer beperkt wordt voldaan aan de in artikel 4 lid 1 onder b van het Schattingsbesluit gestelde eis van reproduceerbaarheid;
- de willekeur op de loer ligt en daarmee strijd met artikel 3:4 Awb.

2.4.5 Verder blijkt uit de literatuur dat de klinische blik wordt vertroebeld door de volgende processen (zie de artikelen van Faust & Ziskin (1988)² en Meyer et al (2001)³):

- zoeken naar bevestiging van de eerste indruk, waarbij aan tegenbewijs (falsificatie) minder of geen betekenis wordt gehecht;
- overmatig vertrouwen in eigen kunnen;
- het is zonder specifiek gericht onderzoek niet mogelijk vast te stellen of iemand simuleert (de overgelegde literatuur laat hiervan sterke voorbeelden zien)⁷;
- er bestaat bij psychische problematiek geen eenduidige relatie tussen diagnose en functionele mogelijkheden⁶.

- 2.4.6 Uit het artikel van Meyer et al (blz. 150/151)³ blijkt hoe groot de verschillen zijn tussen de resultaten van mono-methodisch onderzoek (klinische blik) en multi-methodisch onderzoek. De verschillen zijn zo groot dat het de moeite waard is de gehele paragraaf te citeren:

“Method Disparities and Errors in Practice

Current knowledge about the substantial disagreements between methods of information gathering has important implications for health care. The data indicate that even though it may be less expensive at the outset, a single clinician using a single method (e.g., interview) to obtain information from a patient will develop an incomplete or biased understanding of that patient. To the extent that such impressions guide diagnostic and treatment decisions, patients will be misunderstood, mischaracterized, misdiagnosed, and less than optimally treated. Over the long term, this should increase health care costs.

These issues are not trivial. The evidence indicates that clinicians who use a single method to obtain patient information regularly draw faulty conclusions. For instance, Fennig, Craig, Tanenberg-Karant, and Bromet (1994) reviewed the diagnoses assigned to 223 patients as part of usual hospital practice. Clinical diagnoses were then compared with diagnoses derived from a comprehensive multimethod assessment that consisted of a semistructured patient interview, a review of the patient's medical record, a semistructured interview with the treating clinician, and an interview with the patient's significant other, all of which were then reviewed and synthesized by two clinicians to derive final diagnoses from the multimethod assessment.

Even though Fennig, Craig, Tanenberg-Karant, et al. (1994) used very liberal criteria to define diagnostic agreement (e.g., major depression with psychotic features was treated as equivalent to dysthymia), the diagnoses assigned during the course of typical clinical practice had poor agreement with the diagnostic formulations derived from the more extensive synthesis of multiple assessment methods. Overall, after discounting chance agreement, the clinical diagnoses agreed with the multimethod conclusions only about 45-50% of the time. This was true for a range of disorders on the schizophrenic, bipolar, and depressive spectrums. Because these conditions are treated in decidedly different ways, such frequent misdiagnoses in typical practice suggest that many patients erroneously receive antipsychotic, antimanic, and antidepressant medications.

Another example involves fully structured interviews like the Composite International Diagnostic Interview (CIDI), which have a format that makes them essentially equivalent to an oral self-report instrument. A salient question concerns the extent to which diagnoses from CIDI-type scales agree with those derived from clinicians who also rely on their impression of the patient (e.g., from semistructured interviews, from clinical consensus after following the patient over time). Although diagnoses from the CIDI and diagnoses derived from semistructured interviews suffer from criterion contamination because both the predictor and criterion rely on the patient's report as a primary source of information (see, e.g., Malgady, Rogler, & Tryon, 1992), Table 3 indicates that across 33 samples and 5,990 patients, the correspondence between CIDI-type diagnoses and clinician diagnoses was quite

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWV**
pagina 8

modest ($K = .34$; Table 3, Entry 34; see Meyer, in press). Similar findings have been observed when Axis I diagnoses from the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders were compared with clinician diagnoses (mean $K = .26$, $N = 100$; Steiner, Tebes, Sledge, & Walker, 1995), suggesting again that the source of information for diagnostic inferences exerts a prominent influence over final classifications (see, e.g., Offord et al., 1996).

Although the above disagreements are pronounced, even more drastic errors have been found for personality disorders. Perry (1992) and Pilkonis et al. (1995) compared diagnoses derived from a semistructured clinical interview with diagnoses based on more extensive and complex assessments using multiple methods of gathering patient information. Across studies, there was a meager correspondence between the diagnoses derived from a single clinician using the single method of assessment and the diagnoses derived from the multimethod evaluations ($K = .28$; $N = 218$; see Entry 48 in Table 3). In fact, after correcting for agreements due to chance, about 70% of the interview-based diagnoses were in error.

The evidence also indicates that personality disorder diagnoses diverge substantially across other sources of information. For instance, Table 3 shows that diagnoses derived from self-report bear little resemblance to those derived from clinicians ($K = .18$, $N = 2,859$; Table 3, Entry 33) and that diagnoses from semistructured patient interviews bear little resemblance to those based on semistructured interviews with significant others in the patient's life ($K = .12$, $N = 768$; Table 3, Entry 30).

Though the latter results are sobering, they are open to interpretation about which perspective is more correct. The most relevant evidence is that which compared interviews with the multimethod synthesis of information. These data clearly demonstrate how conclusions derived from a typical evaluation using a single method of assessment had little correspondence with those derived from a more comprehensive evaluation. By necessity then, the research findings indicate that many patients may be misunderstood or improperly treated when they do not receive thorough assessments. Errors of misappraisal and mistreatment are most likely when administrative efforts to save money restrict clinicians to very brief and circumscribed evaluations." (onderstrepingen EvdB)

Dat klinische (mono-methodisch gestelde) diagnoses in niet meer dan 45-50% van de gevallen, en bij persoonlijkheidsstoornissen zelfs maar in ongeveer 30% van de gevallen, door multi-methodisch onderzoek worden bevestigd - met als gevolg dat in 50-70% van de gevallen een verkeerde behandeling is ingezet -, is ronduit schokkend.

Vertaald naar de verzekeringsgeneeskunde betekent dit dat het aannemelijk is dat in 50% of meer van de beoordelingen sprake is van een onjuiste vaststelling van de psychische belastbaarheid.

2.4.7 Deze onbetrouwbaarheid van interview-gebaseerde beoordelingen komt ook naar voren in de veel breder opgezette meta-analyse van 136 onderzoeken door William M. Grove et al ⁴:

“On average, mechanical-prediction techniques were about 10% more accurate than clinical predictions. Depending on the specific analysis, mechanical prediction substantially outperformed clinical prediction in 33%-47% of studies examined. Although clinical predictions were often as accurate as mechanical predictions, in only a few studies (6%-16%) were they substantially more accurate. Superiority for mechanical-prediction techniques was consistent, regardless of the judgment task, type of judges, judges' amounts of experience, or the types of data being combined. Clinical predictions performed relatively less well when predictors included clinical interview data.” (onderstreping EvdB)

2.4.8 Waartoe overmatig zelfvertrouwen bij diagnosticerende artsen kan leiden, blijkt uit een onderzoek uit 1987 over het vermogen van adolescenten om bij het invullen van vragenlijsten aannemelijke neuropsychologische defecten te simuleren:

“Three subjects between the ages of 9 and 12 years were recruited through friends of the laboratory support staff. The subjects had no history of learning difficulties, serious medical illness, or head injury, and each was an above-average student. We informed subjects and their parents that we wished to examine faking on psychological tests. The subjects were told to perform less well than usual but not to be so obvious in doing so that the person testing them would know they were faking. No other guidance was provided.

(...)

Of the 42 case respondents (15 for Case I, 11 for Case 2, and 16 for Case 3), 39 (93%) judged the protocols as abnormal and 3 (7%) as normal. Of the former 39 respondents, 34 (87%) attributed the abnormality to conical dysfunction, 5 (13%) to functional factors, and 0 to malingering. Thus, almost all of the respondents were convinced that abnormality was present, and no respondent detected faking or malingering.

(...)

About three fourths of the judges indicated moderate or greater confidence, and almost two fifths indicated high or very high confidence. We had planned to analyze relations between diagnostic accuracy and a number of variables, including confidence. However, significant findings were precluded because accuracy did not vary (i.e., error rate was 100%). It seems appropriate to call most of the judges overconfident because the majority indicated moderate or greater confidence on a task in which no judge made an accurate identification.” ⁸

De meeste beoordelaars hadden een zeer groot vertrouwen in de juistheid van hun bevindingen, terwijl die bevindingen in 90-100% van de gevallen onjuist waren.

Uit dit onderzoek blijkt dat het bij gezonde personen die opzettelijk simuleren nauwelijks tot niet mogelijk is om objectief en betrouwbaar vast te stellen of en in hoeverre sprake is reële neuropsychologische problematiek of van simulatie.

Uit ander onderzoek blijkt echter dat simulatie wel betrouwbaar is vast te stellen met behulp van specifieke daarop gerichte testen.^{8,9}

- 2.4.9 De verzekeringsarts moet op grond van een gesprek met een verzekerde, die (in beginsel) niet simuleert, de geclaimde klachten en beperkingen beoordelen. Hij moet daarbij onderscheiden tussen wat reëel is en als gevolg van ziekte of gebrek kan worden aangemerkt en datgene wat in de presentatie mogelijk te sterk wordt aangezet of niet als gevolg van ziekte kan worden aangemerkt.

Als het maken van dat onderscheid al niet mogelijk is bij gezonde personen die bewust simuleren, dan is dat bij zieke personen wier klachtenbeleving mogelijk sterker is dan medische verklaarbaar, maar voor een (substantieel) deel wel reëel, nog veel minder mogelijk.

Er valt dan ook niet te ontkomen aan de conclusie dat een klinische beoordeling van psychische klachten op basis van niet meer dan een beoordelingsgesprek ontoereikend is en met andere vormen van onderzoek moet worden aangevuld om tot een betrouwbaar en reproduceerbaar resultaat te komen.

Het belang hiervan wordt onderstreept door het onderzoek van Greiffenstein en Baker waaruit naar voren komt dat in letselschadezaken ("compensable context") de kans op simulatie groter is naar mate het objectieve letsel beperkter is.¹⁰

Niet vergeten mag worden dat het fenomeen "simulatie/maligneren" twee zijden heeft: uiteraard dient voorkomen te worden dat claims worden gehonoreerd c.q. beperkingen worden aangenomen die er de facto niet zijn, maar evenzeer dient te worden voorkomen dat op grond van een onvoldoende onderbouwd vermoeden van maligneren claims c.q. beperkingen worden afgewezen die er wel degelijk zijn.

- 2.4.10 Uit de literatuur blijkt dat er voldoende betrouwbare tests beschikbaar zijn om bijv.:
- de klachten goed in beeld te brengen (SCL-90¹¹ en 4DKL¹²), zodat gericht nader onderzoek kan plaatsvinden;
 - een angststoornis of depressie uit te sluiten (HADS en DASS)¹³, zodat van nader onderzoek naar die stoornissen kan worden afgezien of (indien die diagnoses niet worden uitgesloten) gericht verder onderzoek kan plaatsvinden;
 - simulatie vast te stellen (ASTM en vooral SIMS)⁹.

- 2.4.11 Uit de literatuur komen geen argumenten naar voren op grond waarvan bij een verzekeringsgeneeskundige beoordeling van het gebruik van dergelijke instrumenten zou moeten worden afgezien.

Wel komen uit die literatuur sterke argumenten naar voren die het (aanvullende) gebruik van testonderzoek ondersteunen. Zie Meyer et al:

“Data from more than 125 meta-analyses on test validity and 800 samples examining multimethod assessment suggest 4 general conclusions:

- (a) Psychological test validity is strong and compelling,
- (b) psychological test validity is comparable to medical test validity,
- (c) distinct assessment methods provide unique sources of information, and
- (d) clinicians who rely exclusively on interviews are prone to incomplete understandings.

Following principles for optimal nomothetic research, the authors suggest that a multimethod assessment battery provides a structured means for skilled clinicians to maximize the validity of individualized assessments.”³

- 2.4.12 Op basis van de overgelegde wetenschappelijke literatuur kan dan ook niet anders dan worden geconcludeerd dat een verzekeringsgeneeskundige beoordeling die is gebaseerd op alleen de “klinische blik” van de verzekeringsarts, niet als ‘state of the art’ kan worden aangemerkt, maar dient te worden aangevuld met en getoetst aan de resultaten van testonderzoek. Zie ook C. de Ruiter, Evidence-based psychodiagnostiek.¹⁴

Verder dient te worden vastgesteld dat er voldoende valide en betrouwbare testmethodieken zijn die ook in de verzekeringsgeneeskunde kunnen worden toegepast, zodat eigenlijk onbegrijpelijk is waarom daarvan niet meer gebruik wordt gemaakt; zie de waarschuwing in het hiervoor al geciteerde artikel van Van Dijk et al¹.

Tot slot verdient opmerking dat in de bedrijfsgeneeskunde, die toch sterk aan de verzekeringsgeneeskunde is gelieerd, wel in toenemende mate van testonderzoek gebruik wordt gemaakt (zie de NVAB-richtlijn “Handelen van de bedrijfsarts bij werknemers met psychische klachten” van 28 januari 2000).

2.5 Algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethoden

- 2.5.1 In artikel 4 lid 2 van het Schattingsbesluit wordt de eis gesteld dat het verzekeringsgeneeskundige onderzoek geschiedt aan de hand van algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethoden.

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWW**
pagina 12

2.5.2 In de Nota van Toelichting bij het Schattingsbesluit is ter toelichting bij de introductie van deze eis het volgende te lezen:

“In zijn brief aan de Staatssecretaris van Sociale Zaken en Werkgelegenheid van 1 september 1997 over de richtlijn MAOC wijst het Ctsv er op dat «uit hetgeen de wetgever en de rechter naar voren brengen kan worden opgemaakt, dat de uitvoerder van de AAW/WAO bij zijn medische beoordeling niet over een nacht ijs mag gaan. Een gedegen medisch onderzoek is altijd een vereiste en hoe moeilijker het is een medische oorzaak vast te stellen, hoe grondiger en uitvoeriger het te verrichten onderzoek zal moeten zijn». Het besluitvormingsproces dient derhalve voldoende waarborgen te bevatten voor een gefundeerd medisch oordeel.

...

Om te kunnen spreken van een kwalitatief goede claimbeoordeling dient aan de volgende eisen te worden voldaan. Ten eerste is daar de vaststelbaarheid van de stoornissen, beperkingen en handicaps.

...

De tweede voorwaarde is dat het vaststellen van de ongeschiktheid medisch dient te zijn, hetgeen wil zeggen op in de reguliere gezondheidszorg aanvaarde wijze. Dit vereist dat onderzoeksmethoden toegepast dienen te worden die in de verzekeringsgeneeskunde onder artsen en paramedische deskundigen algemeen geaccepteerd zijn.

...

Een derde voorwaarde is dat het oordeel van de verzekeringsarts objectief moet zijn. Objectief in dit verband wil zeggen gebaseerd op de feiten, niet beïnvloed door vooroordelen.” (Stb 2000, 307, blz. 15/16)

Er dient dus sprake te zijn van een gefundeerd medisch oordeel dat op een in de reguliere gezondheidszorg aanvaarde wijze wordt vastgesteld en objectief is.

2.5.3 Gelet op de uit de wetenschappelijke literatuur daartegen blijvende bezwaren, kan de loutere klinische blik niet worden aangemerkt als een “algemeen aanvaarde verzekeringsgeneeskundige onderzoeksmethode” als bedoeld in artikel 4 lid 2 van het Schattingsbesluit.

2.5.4 Blijkens het Handboek psychodiagnostiek onder redactie van Corine de Ruiter & Martin Hildebrand uit 2006, blz. 11-17, dient psychodiagnostiek zoveel mogelijk evidence-based te zijn, hetgeen inhoudt:

- dat er gebruik wordt gemaakt van gestandaardiseerde en wetenschappelijk onderbouwde instrumenten;
- er sprake is van transparantie (expliciet maken van de keuze voor bepaalde meetinstrumenten en van de relatie tussen ruwe testgegevens en de interpretatie daarvan);
- er multimethodisch wordt gewerkt, d.w.z. dat gebruik wordt gemaakt van meer testmethoden en meer informatiebronnen.¹⁴

De eis van artikel 4 lid 2 van het Schattingsbesluit leidt er dan ook toe dat de verzekeringsarts zijn oordeel over psychische problematiek zal dienen te baseren op de combinatie van testonderzoek en klinische beoordeling.

Het testonderzoek kan daarbij zowel de richting wijzen voor nader onderzoek als een kritische toets zijn voor het eigen klinisch oordeel, bijv. doordat simulatie objectief wordt aangetoond, een diagnose wordt uitgesloten of een eigen hypothese wordt gefalsificeerd of juist bevestigd.

- 2.5.5 De vraag tot slot of de verzekeringsarts in het gebruik van dit soort instrumenten is geschoold, is niet relevant: hij kan zich ofwel daarin alsnog scholen ofwel gebruik maken van externe expertise.

Dat aan externe expertise kosten zijn verbonden, mag niet ten nadele van de verzekerde worden gebracht. Artikel 3:4 lid 2 Awb staat daaraan in de weg.

2.6 Bijzondere problemen bij - retrospectieve - toetsing in bezwaar en beroep

- 2.6.1 Hetgeen hiervoor is beschreven, doet zich voor bij de beoordeling van de actuele belastbaarheid. Indien eenmaal een besluit met betrekking tot het recht op arbeidsongeschiktheidsuitkering is genomen, dan staat in bezwaar of beroep niet meer de actuele belastbaarheid ter beoordeling, maar die op de datum in geding. Zoals de Centrale Raad van Beroep in zijn uitspraak van 30 januari 2007, LJN AZ7600, terecht overweegt, hebben dergelijke zaken altijd een retrospectief karakter en zal dan steeds een beoordeling van een situatie in het verleden moeten plaatsvinden.
- 2.6.2 Prof. dr. P.J. van Koppen wijst er echter in zijn artikel "Weg van de toerekeningsvatbaarheid: Over rapportages over de verdachte" in *Trema*, 27, 221-228, op dat er helemaal geen tests bestaan waarmee iemands toestand in het verleden kan worden vastgesteld (blz. 224 l.k.) en dat er vele factoren zijn waardoor een retrospectieve beoordeling wordt verstoord.
- 2.6.3 Dat een beoordeling achteraf niet rechtstreeks mogelijk is en met bijzondere moeilijkheden heeft te kampen, leidt ertoe dat aan de eisen die aan een actuele beoordeling dienen te worden gesteld, strikt de hand moet worden gehouden.
- 2.6.4 Tweede consequentie daarvan is dat gebreken in de beoordeling zo kort mogelijk na de datum in geding moeten worden hersteld. De inschatting van de situatie per datum in geding op basis van een onderzoek ongeveer 1½ jaar na dato (in het kader van de procedure bij de rechtbank, zoals in casu) is veel moeilijker dan op basis van een onderzoek rondom tot enkele maanden na die datum (zoals in bezwaar mogelijk

is). Indien het multimethodische onderzoek pas na 4 tot 5 jaar plaatsvindt (zoals bij onderzoeken in opdracht van de Centrale Raad van Beroep het geval is) geldt dat nog meer.

- 2.6.5 Een en ander leidt ertoe dat, indien de rechtbank de medische grondslag ontoereikend acht, hij zelf opdracht tot nader onderzoek moet geven en niet mag volstaan met een vernietiging en terugwijzing naar het UWV. Met dat laatste is immers meer tijd gemoeid.

3 Wanneer dient de rechter een deskundige in te schakelen?

3.1 Twijfel

- 3.1.1 Volgens vaste jurisprudentie is er in het algemeen slechts aanleiding tot het inschakelen van een medisch deskundige, indien de gedingstukken dan wel de aangedragen medische gegevens reden geven voor twijfel aan de juistheid van de voor betrokkene vastgestelde beperkingen. Zie bijv. CRvB 04-08-1998, USZ 1998/273.

Voorts kan daartoe aanleiding bestaan indien sprake is van een situatie waarin de standpunten van de verzekeringsarts en de behandelend arts tegenover elkaar staan zonder dat bij voorbaat kan worden gezegd dat een van die standpunten volstrekt onvoldoende gemotiveerd of anderszins volstrekt onhoudbaar is. Zie CRvB 11-01-2000, USZ 2000/79. In CRvB 28-03-2007, LJN BA2093, wordt een vergelijkbaar oordeel gegeven, maar dan met verwijzing naar onderdeel 2.3.5. van de Standaard "Onderzoeksmethoden bij psychische stoornissen".

3.2 Wat geeft aanleiding tot twijfel?

- 3.2.1 Twijfel kan veroorzaakt worden door zowel de wijze van totstandkoming als de materiële inhoud van de rapportage van de verzekeringsarts. Indien de wijze van totstandkoming al tot twijfel aanleiding geeft, wordt aan de materiële inhoud eigenlijk niet meer toegekomen.
- 3.2.2 Dit betekent dat voor het onderbouwen van twijfel het niet altijd noodzakelijk is nieuwe medische stukken over te leggen. Dat volgt ook uit de aangehaalde uitspraak van de Centrale Raad van Beroep (USZ 1998/273), waarin de reden tot twijfel wordt gevonden in "de gedingstukken dan wel de aangedragen medische gegevens" (onderstreping EvdB).

3.3 Materiële inhoud

- 3.3.1 Materiële twijfel is het gemakkelijkst en vaak ook het meest in het oog springend. Verschil van inzicht tussen verzekeringsarts en behandelaar over de arbeidsmogelijkheden, zoals in de hiervoor aangehaalde uitspraken, is een voorbeeld hiervan. Ook kan sprake zijn van verschillen in gestelde diagnose of van informatie die voorheen nog niet bekend was.
- 3.3.2 Materiële twijfel kan ontstaan door het overleggen van nieuwe medische stukken, maar ook vanuit het reeds bestaande dossier. In de jurisprudentie over de consistentie van het medisch oordeel in de tijd (zie hierna in paragraaf 4.1) komt de twijfel steevast voort uit discrepanties tussen de bestreden beoordeling en eerdere beoordelingen zonder dat sprake is van nieuw overgelegde medische stukken.

3.4 Wijze van totstandkoming

- 3.4.1 Er mag echter niet aan voorbij worden gegaan dat ook de wijze van totstandkoming van het advies van de verzekeringsarts aanleiding kan geven tot twijfel.
- 3.4.2 Indien het door de verzekeringsarts verrichte onderzoek niet voldoet aan de eisen van:
- zorgvuldig onderzoek (artikel 3:2 en 3:4 Awb),
 - objectiviteit (18 lid 1 WAO),
 - reproduceerbaarheid (artikel 4 lid 1 onder b Schattingsbesluit),
 - consistentie (artikel 4 lid 1 onder c Schattingsbesluit) en
 - onderzoeksmethode (artikel 4 lid 2 Schattingsbesluit)

dan kan niet zonder meer worden uitgegaan van de juistheid van het resultaat van die beoordeling en moet er rekening mee worden gehouden dat er een reële kans bestaat dat een onderzoek dat wel aan alle eisen voldoet, tot andere uitkomsten leidt.

In deze situatie is weliswaar minder evident dat het advies van de verzekeringsarts mogelijk onjuist is dan bij een duidelijk verschil van inzicht met de behandelaar, maar de mogelijkheid dat kwalitatief beter onderzoek tot andere resultaten leidt, mag niet voor risico en rekening van de betrokken verzekerde worden gelaten.

- 3.4.3 Van een dergelijke situatie is sprake indien de verzekeringsarts heeft verzuimd gebruik te maken van onderzoeksinstrumenten die wel beschikbaar zijn.

- 3.4.4 Verder dient bij deze toetsing te worden betrokken het in hoofdstuk 3 (blz. 31 e.v.) van zijn advies van 4 juni 2007, nr. 2007/14, "Verzekeringsgeneeskundige mediprudentie", door de Gezondheidsraad voorgestelde toetsingskader.

Indien de verzekeringsgeneeskundige rapportage niet voldoet aan de in dit toetsingskader geformuleerde eisen, dan is niet voldaan aan de daaraan te stellen eisen van zorgvuldig onderzoek en kan niet zonder meer van de juistheid van de bevindingen en conclusies worden uitgegaan.

- 3.4.5 Tot slot moet er rekening mee worden gehouden dat een verzekerde aanzienlijk minder gemakkelijk toegang heeft tot (onafhankelijke) medische expertise dan het UWV, dat die expertise immers in dienst heeft.

Volgens het rapport "Ben ik overtuigd?" behoort in verband hiermee het beginsel van de ongelijkheidscompensatie een rol te spelen bij de beslissing een deskundige te benoemen (Werkgroep Medische Toetsing, Ben ik overtuigd? Notitie van een werkgroep voorzitters sectoren bestuursrecht, mr. A.H.J. Lennaerts, Arrondissementsrechtbank Assen, mei 1995).

4 Bijzonderheden in de zaak A.

4.1 Consistentie van het medisch oordeel in de loop der tijd

- 4.1.1 Artikel 4 lid 1 van het Schattingsbesluit arbeidsongeschiktheidswetten bepaalt dat het verzekeringsgeneeskundig onderzoek dient te voldoen aan de volgende vereisten:
- a. (...)
 - b. een door een andere verzekeringsarts uitgevoerd verzekeringsgeneeskundig onderzoek zal tot dezelfde bevindingen en conclusies kunnen leiden;
 - c. de redeneringen en conclusies van het verzekeringsgeneeskundig onderzoek zijn vrij van innerlijke tegenspraak.

Deze eisen kunnen kort worden aangeduid als eisen van reproductiviteit (onderdeel b) en consistentie (onderdeel c).

De Centrale Raad van Beroep heeft in een reeks van uitspraken de lijn neergelegd dat een wijziging van de belastbaarheid in beginsel dient te kunnen worden teruggevoerd op een verandering van de medische situatie. Indien sprake is van een lichtere inschatting van de beperkingen bij een overigens ongewijzigd medisch beeld, dient dat verschil in inschatting overtuigend te worden gemotiveerd. Zie CRvB 08-03-2005, USZ 2005/168; CRvB 08-03-2005, USZ 2005/168; CRvB 03-01-2007, LJN AZ6490; CRvB 14-02-2007, LJN AZ8558. De overwegingen in CRvB 11-05-

2007, LJN BA5749, met betrekking tot het laten vervallen van een urenbeperking zijn zeer illustratief:

“Bij de thans in geding zijnde schatting heeft het Uwv, anders dan bij een eerdere beoordeling, geen urenbeperking meer geïndiceerd geacht voor appellante. Naar het oordeel van de Raad ontbreekt echter een voldoende overtuigende onderbouwing voor het laten vervallen van deze urenbeperking. De primaire verzekeringsarts is in zijn rapport d.d. 15 april 2003 op grond van het dagverhaal van appellante tot de conclusie gekomen dat zij voldoende rust krijgt en er geen indicatie meer is voor een urenbeperking en de bezwaarverzekeringsarts heeft zich, zonder nader in te gaan op dit aspect, achter dit medische oordeel van de primaire verzekeringsarts geschaard. Een dergelijke motivering acht de Raad in het licht van de eerdere medische rapportage waarin wel een urenbeperking is aangenomen echter onvoldoende, te meer daar appellante heeft gesteld dat in haar gezondheidstoestand geen verbetering is gekomen.”

Hoewel de Centrale Raad in bovengenoemde uitspraken niet met zoveel woorden verwijst naar artikel 4 lid 1 van het Schattingsbesluit, dient deze jurisprudentie wel in het licht van deze bepaling te worden gezien. Waar artikel 4 lid 1 Schattingsbesluit met name lijkt te zien op het verzekeringsgeneeskundig onderzoek met betrekking tot alleen de datum in geding, ziet de aangehaalde jurisprudentie van de Centrale Raad met name op de reproduceerbaarheid en consistentie van de beoordeling in de loop der tijd.

- 4.1.2 Bezwaar-verzekeringsarts W.M. K. schrijft in haar rapportage van 4 oktober 2006 (B58) dat de ernst van de psychische problematiek ten opzichte van het vorige beoordelingsmoment (22 mei 1996, B24 en 25/26) is afgenomen.

Toch is dat niet zo evident: A.'s huisarts spreekt over een depressieve stoornis met paniekaanvallen, waar de primair oordelende arts A.K. van B. spreekt van spanningsklachten (B35.2) en de bezwaar-verzekeringsarts van een aanpassingsstoornis, recidiverend met depressieve stemming. De stelling van de bezwaar-verzekeringsarts in haar rapportage van 6 december 2006 (B66) dat men het eigenlijk over alles eens is, is dan ook niet juist. Er bestaat wel degelijk verschil van mening over zowel de diagnose als (vooral) de ernst daarvan.

Hier wreekt zich direct dat verweerder geen gebruik maakt van vragenlijsten als de 4DKL of SCL-90. Indien wel van dit soort (screenende) vragenlijsten gebruik zou zijn gemaakt, zouden de lijsten van 1996 en heden met elkaar vergeleken kunnen worden en een meer objectief beeld kunnen worden verkregen van de verschillen in de door A. ervaren c.q. gerapporteerde klachten tussen beide data.

Verder ligt een opklaren van de depressieve stoornis niet zo voor de hand, nu de thuissituatie, die daarbij blijkt het dossier een aanzienlijke onderhoudende rol

speelt, niet wezenlijk is veranderd. Ook vanuit de wetenschappelijke literatuur kan worden onderbouwd dat veeleer sprake is van een juist ongunstige prognose; zie W. van der Hoek, Prognose van depressie in het kader van de claimbeoordeling voor de nieuwe WAO in TBV 2006, nr. 4.¹⁵

De in het dossier te vinden waarschuwingen van psychiater Y. B. (B18) en psychiater A. K. (B24) tot slot ondersteunen de vermeende opklaring ook niet.

- 4.1.3 Kortom, de bevindingen die ten grondslag liggen aan het bestreden besluit wijken zodanig af van de overige in het dossier aanwezige bevindingen, dat serieuze twijfel is gerechtvaardigd over de juistheid van de medische onderbouwing van het bestreden besluit.

4.2 Sociale omstandigheden

- 4.2.1 Uit het dossier komt naar voren dat verweerders besluit vooral is ingegeven door de overweging dat de belastende thuissituatie van A. bij de beoordeling van zijn arbeidsmogelijkheden geen rol mag spelen. Zoals de bezwaar-verzekeringarts het in haar rapportage van 6 december 2006 (B66) formuleert: verweerder meent dat A. een keus moeten maken tussen het verzorgen van zijn zieke vrouw en het verrichten van arbeid.

- 4.2.2 Dit standpunt is niet alleen buitengewoon cynisch, maar ook onjuist. In het Verzekeringsgeneeskundig protocol Algemene inleiding van 19 december 2006, nr. 2006/22, blz. 15/16, schrijft de Gezondheidsraad:

“Het concept “ongeschiktheid tot werken ten gevolge van ziekte” in het Schattingsbesluit wordt in de bijbehorende Nota van toelichting uitgewerkt aan de hand van drie begrippen die ontleend zijn aan de in 1980 voor het eerst verschenen International Classification of Impairments, Disabilities and Handicaps (ICIDH) van de WHO: de begrippen stoornis, beperking en handicap.
(...)

De ICIDH werd in 2001 opgevolgd door de International Classification of Functioning, Disability and Health (ICF): (...) Centraal in de ICF staat een model van het menselijk functioneren (figuur 1), waarin de onderlinge relaties tussen stoornissen, beperkingen en handicaps (in de ICF-terminologie: participatieproblemen) in beeld worden gebracht. De drie niveaus waarop ziekte zich manifesteert worden omschreven als

- het niveau van de functies en anatomische eigenschappen, waarop zich stoornissen kunnen voordoen,
- het niveau van de activiteiten, de onderdelen van iemands handelen, waarop beperkingen aanwezig kunnen zijn, en
- het niveau van de participatie, waarop sprake kan zijn van participatieproblemen, i.e.: problemen met het deelnemen aan het maatschappelijk leven.

De drie niveaus van het menselijk functioneren hangen niet alleen in onderlinge wisselwerking met elkaar samen, maar zijn ook verbonden met externe factoren (fysieke en sociale) en met persoonlijke factoren (bijvoorbeeld leeftijd, persoonlijkheid levensstijl), die zowel op het functioneren van invloed zijn als erdoor beïnvloed worden.

De verzekeringsarts maakt bij de arbeidsongeschiktheidsbeoordeling van dit samenhangende, dynamische systeem van het menselijk organisme en zijn omgeving een momentopname. Is er op dit moment sprake van een consistente en plausibele samenhang van stoornissen, beperkingen en participatieproblemen? Is dat het geval, dan kan er sprake zijn van "ongeschiktheid tot werken ten gevolge van ziekte" in de zin van het Schattingsbesluit.

Om na te kunnen gaan of bij iemand een dergelijke consistente en plausibele samenhang aan de orde is, is het van belang zicht te hebben op de externe en persoonlijke factoren die met het functioneren van de persoon verbonden zijn. De stoornissen, beperkingen en participatieproblemen doen zich altijd voor in een bepaalde persoon in een bepaalde omgeving. Al naar gelang de eigenschappen van die persoon en omgeving kan de samenhang van stoornissen, beperkingen en participatieproblemen verschillende gedaanten aannemen. Mede daarom is de arbeidsongeschiktheidsbeoordeling altijd een individuele beoordeling. Soms zal blijken dat in het ontstaan of voortbestaan van een samenhangend geheel van stoornissen, beperkingen en participatieproblemen externe en persoonlijke factoren, zoals gebrek aan sociale steun, inactiviteit, persoonlijkheidskenmerken, ambities, copingstijlen enzovoort, een belangrijke rol spelen. Dat is geen reden om, wanneer is vastgesteld dat van een dergelijke samenhang sprake is, de omvang van de beperkingen als geringer te beoordelen dan wanneer niet of nauwelijks aanwijsbare persoonlijke of externe factoren in het geding zouden zijn geweest." (onderstreping EvdB)

4.2.3 Het geheel buiten beschouwing laten van de invloed van de thuissituatie is dan ook in strijd met huidig wetenschappelijk medisch inzicht.

4.2.4 Hieraan wordt toegevoegd dat blijkens hetgeen psychiater Y. B. (B18) en psychiater A. K. (B24) schrijven, het verstoren van het wankele evenwicht thuis een relevant risico op decompensatie met zich meebrengt. Of zoals prof. dr. H.J. van A. het in zijn brief van 17 september 1996 (B23) schrijft:

"Niet zijn arbeidsongeschiktheid wordt sociaal maatschappelijk bepaald, maar zijn psychische ziekte wordt sterk sociaal/maatschappelijk beïnvloed."

Verweerder gaat hier ten onrechte aan voorbij.

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWV**
pagina 20

5 Conclusie

- 5.1 Een en ander leidt tot de conclusie dat de verzekeringsgeneeskundige grondslag van het bestreden besluit niet voldoet aan de eisen die daaraan op grond van artikel 3:2 en 3:4 Awb, artikel 18 lid 1 WAO en artikel 4 lid 2 van het Schattingsbesluit dienen te worden gesteld.
- 5.2 Deze gebreken aan de beoordeling leiden ertoe dat ernstig dient te worden getwijfeld aan de juistheid van de uit die gebrekkige beoordeling getrokken conclusies. Dat dient te leiden tot benoeming van een onafhankelijk deskundige.
- 5.3 Voorgesteld wordt daartoe een psycholoog te benoemen en in de te verstrekken opdracht de instructie op te nemen in het kader van zijn onderzoek tevens een psychiatrisch consult te doen plaatsvinden.
- 5.4 Indien u overgaat tot de benoeming van een deskundige, verzoek ik u mij in de gelegenheid te stellen mij uit te laten over de te formuleren vraagstelling.

Hoogachtend,

E. van den Bogaard

Bijlagen

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWV**
pagina 21

Bijlagen:

(om de reden van overlegging van de betreffende artikelen te illustreren, heb ik bij een aantal artikelen enkele in het oog springende passages geciteerd)

1. F. van Dijk, G. Reinders, F. Schaafsma, D. Spreeuwers, C. Hulshof, M. Abspoel, F. Vlek, J. Verbeek, Een virtuele bibliotheek voor arbo en sociale zekerheid, *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde*, 11 (december 2003), p. 368-372
2. David Faust and Jay Ziskin (1988), The Expert Witness in Psychology and Psychiatry, *Science*, Vol. 241, 1 July 1988, p. 32

“Studies that compare clinicians' predictions against objectively determinable, hard data commonly show that error rate exceeds accuracy rate (3). In one study, for example, a series of military recruits was retained in service despite psychiatrists' recommendations that they be discharged for severe psychiatric liabilities (11). After 2 years, most of these individuals had remained on active duty and their overall rate of success and adjustment was not substantially different from that of matched controls initially judged to be free of pathology.”

“In another study, lay interviewers using standardized questions produced information of equal or greater validity than psychiatrists conducting interviews in their preferred manner.”

“In a study on the detection of malingering, most clinicians expressed extreme confidence on a diagnostic task in which error rate ranged from 90 to 100%.”

“Clinicians commonly propose that their conclusions rest on a careful weighting of many variables, whereas objective analysis typically shows that only a few variables, perhaps two or three, exert a significant impact (39-40). Clinicians also assert that complex configural analysis or data integration is necessary to reach accurate conclusions—that one never considers datum in isolation but rather the "whole" or overall pattern of results. However, numerous studies suggest that no clinician, or human being for that matter, can begin to manage such complex cognitive operations (41-43). The attempt to grasp interactions among even two or three variables can outstrip human cognitive capacities.”

3. Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001), Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56, 128–165

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWV**
pagina 22

4. Grove, W.M., Zald, D.H., Hallberg, A.M., Lebow, B., Snitz, E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19–30

Het artikel van Meyer et al is weliswaar veel directer toegespitst op de beoordeling van psychische problematiek, maar het onderzoek van Grove et al geldt als klassieker in de “statisch versus klinisch” discussie. Er komt uit naar voren dat de gebrekkigheid van het klinisch oordeel zeer algemeen is en zich in versterkte mate voordoet zodra dat oordeel is gebaseerd op gegevens die zijn verkregen uit een interview. Dat laatste is relevant aangezien een beoordelingsgesprek met een verzekeringsarts een interview is.

“The process of making judgments and decisions requires a method for combining data. To compare the accuracy of clinical and mechanical (formal, statistical) data-combination techniques, we performed a meta-analysis on studies of human health and behavior. On average, mechanical-prediction techniques were about 10% more accurate than clinical predictions. Depending on the specific analysis, mechanical prediction substantially outperformed clinical prediction in 33%-47% of studies examined. Although clinical predictions were often as accurate as mechanical predictions, in only a few studies (6%-16%) were they substantially more accurate. Superiority for mechanical-prediction techniques was consistent, regardless of the judgment task, type of judges, judges' amounts of experience, or the types of data being combined. Clinical predictions performed relatively less well when predictors included clinical interview data. These data indicate that mechanical predictions of human behaviors are equal or superior to clinical prediction methods for a wide range of circumstances.”

“Why do we obtain these results? Humans are susceptible to many errors in clinical judgment (Garb, 1998; Kahneman, Slovic, & Tversky, 1982). These include ignoring base rates, assigning nonoptimal weights to cues, failure to take into account regression toward the mean, and failure to properly assess covariation. Heuristics such as representativeness (which leads to belief in the law of small numbers) or availability (leading to over-weighting vivid data) can similarly reduce clinicians' accuracy. Also, clinicians often do not receive adequate feedback on the accuracy of their judgments (Einhorn & Hogarth, 1978), which gives them scant opportunity to change maladaptive judgment habits. In this regard, it is notable that experienced psychologists frequently show little improvement in the accuracy of their clinical judgments relative to the clinical judgments of psychology graduate students (Garb, 1989, 1998).

5. E-mailwisseling prof. dr. H.L.G.J. Merckelbach d.d. 4 en 5 oktober 2007
6. S.M. Laitinen-Krispijn, L.C. Nicolai, Verzekeringsgeneeskundige beoordeling bij psychische problematiek, *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde*, 4 (2006), p. 167-170

7. Faust, David; Hart, Kathleen J.; Guilmette, Thomas J., Pediatric malingering: The capacity of children to fake believable deficits on neuropsychological testing, *Journal of Consulting and Clinical Psychology*, 1988 Aug Vol 56(4) 578-582
8. M. Jelacic, H. Merckelbach, M. Cima, Over het simuleren van cognitieve stoornissen, *Tijdschrift voor Psychiatrie* 45 (2003) 11, p. 687-696

“Zodra clinici het juridisch domein betreden om daar cognitieve klachten en symptomen van patiënten te taxeren, moeten zij beducht zijn voor de mogelijkheid dat afhankelijk van de context, klachten en symptomen worden gemaximaliseerd of juist geminimaliseerd. Het is voor psychiaters en neurologen goed om er rekening mee te houden dat simuleren vooral kan voorkomen bij patiënten die verward zijn in de afhandeling van letselschade, bij verdachten van misdrijven en bij forensische patiënten. Deze categorieën van patiënten kunnen voordeel hebben bij het aandikken of juist onderrapporteren van cognitieve stoornissen. Nagaan of patiënten met schadeclaims cognitieve stoornissen simuleren, kan het best geschieden met gespecialiseerde tests als de AKTG en vragenlijsten als de SIMS. Om bij verdachten onderscheid te maken tussen organische en gesimuleerde amnesie voor een misdrijf, moet men zich afvragen of hij of zij de specifieke kenmerken van hersenbeschadiging heeft. Daarnaast kan in bepaalde gevallen worden overwogen om een SVT in te zetten. De SS is een veelbelovend instrument om forensische patiënten te onderzoeken op hun neiging tot *fake good*.

Mocht de clinicus bij patiënten met letselschade, verdachten van een misdrijf of forensische patiënten onvoldoende rekening houden met simuleren, dan is sprake van een gemankeerd onderzoek.”

9. Elicic, M., Merckelbach, H., Candel, I., & Geraerts, E. (in press). Detection of feigned cognitive dysfunction using special malingering tests: A simulation study in naïve and coached malingerers. *International Journal of Neuroscience*, Volume 117, Issue 8 August 2007, pages 1185 - 1192

The aim of the present study was to compare the accuracy of the Amsterdam Short Term Memory (ASTM) test with that of the Structured Inventory of the Malingered Symptomatology (SIMS) in detecting feigning of cognitive dysfunction in naïve and coached participants. Ninety undergraduate students were administered the ASTM and the SIMS and asked to respond honestly (controls; n = 30), or instructed to malingering cognitive dysfunction due to head injury. Before the both instruments were administered, naïve malingerers received no further information (n = 30), whereas coached malingerers were given some information about brain injury and a warning not to exaggerate symptoms (n = 30). Both tests correctly classified 90% of the naïve malingerers. The ASTM detected 70% of the coached malingerers, whereas the SIMS continued to detect 90% of them. The findings suggest that coaching undermines the diagnostic accuracy of the ASTM, but does not seem to influence the accuracy of the SIMS.

datum 20 november 2007
dossier 20060024B
aan Rechtbank Amsterdam, Sector Bestuursrecht
inzake **A. / UWV**
pagina 24

10. Greiffenstein, Manfred F.; Baker, W. John, Miller was (mostly) right: Head injury severity inversely related to simulation, *Legal and Criminological Psychology*, Volume 11, Number 1, February 2006 , pp. 131-145

“We found robust inverse dose-response relationships in two of three functional areas. Poorer performance on simple memory and motor tasks in persons with minor craniocervical injury. Examination of base rates for three types of simulated deficit (memory, motor, psychiatric) demonstrated that suspected simulation was the rule rather than the exception. Those reporting the most symptoms (‘floridity’ in Miller’s terms) showed the highest rates of egregious simulation, although not universally so.”

“We strongly urge formal effort testing in all compensation seekers.”

11. H. Scharft, C.J.F. Kemperman, M.P. Steenhuis, De bruikbaarheid van de nieuwe normen van de SCL-90 bij het opsporen van psychische klachten in een arbeids(on)geschiktheidsprocedure, *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde*, 2004 (12), 7, 204-207
12. Verzekeringsgeneeskundig protocol Angststoornissen, Gezondheidsraad 2007/5, blz. 33; NVAB-richtlijn “Handelen van de bedrijfsarts bij werknemers met psychische klachten”, blz. 80 (beiden niet overgelegd)
13. E.M. de Croon, K. Nieuwenhuijsen, N.I.R. Hugenholtz, F.J.H. van Dijk, Drie vragenlijsten voor diagnostiek van depressie en angststoornissen, *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde*, 13 (2005), p. 98-103

14. C. de Ruiter, Evidence-based psychodiagnostiek, in C. de Ruiter en M. Hilderbrand (red.), Handboek psychodiagnostiek, van testmethode naar toepassing, Harcourt 2006, blz. 11-18

“Ten slotte houdt evidence-based psychodiagnostiek in dat er multimethodisch gewerkt wordt. Dit wil zeggen dat diagnostische uitspraken over personen zoveel mogelijk gebaseerd dienen te zijn op gegevens die verkregen zijn uit meerdere testmethoden en van meerdere informatiebronnen. Wetenschappelijk onderzoek laat zien dat monomethodisch onderzoek (bijvoorbeeld: louter zelfrapportagetests, of louter een onderzoeksgesprek met de cliënt) veelal leidt tot onjuiste interpretaties (Meyer e.a., 2001). Wanneer eenzelfde persoonlijkheidskenmerk, bijvoorbeeld impulsiviteit, met behulp van verschillende testmethoden wordt onderzocht (bijvoorbeeld: met een experimentele cognitieve computertaak, niet een zelfrapportagevragenlijst als de Barratt Impulsiveness Scale en door heteroanamnestische gesprekken met bekenden van de onderzochte over dit aspect van zijn persoonlijkheid, ontstaat pas een volledig beeld. Wanneer de resultaten verkregen met verschillende testmethoden convergeren, kan de psychodiagnosticus zijn diagnostische interpretaties met meer stelligheid formuleren dan wanneer dat niet het geval is. Soms roepen divergerende testresultaten aanvullende vragen op die alleen met nader diagnostisch onderzoek of door andere informatiebronnen aan te boren, beantwoord kunnen worden. Dit proces verhoogt uiteraard de kans op een juiste en nauwkeurige diagnose. Bij multimethodisch werken behoort uiteraard ook de integratie van contextuele factoren bij de interpretatie van diagnostische gegevens. Contextuele factoren kunnen immers een rol spelen bij de totstandkoming van de uiteindelijke testcores, denk bijvoorbeeld aan de rol van vermoeidheid, aandacht en medewerking bij de scores op neuropsychologische tests.”

15. W. van der Hoek, Prognose van depressie in het kader van de claimbeoordeling voor de nieuwe WAO, *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde*, 2006 (14), 4, 161-166

“Voor de verzekeringsarts is van belang dat bij een cliënt die 2 jaar verzuimt wegens een depressie en die een einde-wachttijdbeoordeling voor de WIA ondergaat, gesproken kan worden van een chronische depressie met ongunstige prognose. Zo al (gedeeltelijk) herstel optreedt dan nog is er een grote kans op terugval, verhoogd verzuim en verminderde productiviteit. Dit betekent dat in de belastbaarheid vastgesteld bij einde wachttijd in het kader van de WIA vaak geen belangrijke verbetering mag worden verwacht.”