



Universiteit Utrecht

**Psychologische Diagnostiek is Diagnostiek van de  
Psychologie**

***Jan J.F. ter Laak Faculteit  
Sociale Wetenschappen  
Afdeling Ontwikkelingspsychologie  
Universiteit Utrecht***

© Jan ter Laak

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier.

*Werken, om 't werk niet, maar om tegen 't zwijgen  
En twijf'len argumenten te verkrijgen*

Martinus Nijhoff: 'Levensloop'. Uit *Vormen*, 1<sup>e</sup> druk, Bussum: Van Dishoeck, 1924.

Voor Marie-Louise en Mose

## Inhoud

<b>Dankwoord</b>	5
<b>I Inleiding</b>	6
1. Psychologie studeren en beoefenen	7
2. Inspiratie	12
3. Standpunten	15
4. Vooruitblik	17
<i>Referenties</i>	21
<b>II Wat is psychologische diagnostiek?</b>	24
1. Het begin: Psychologie en de vraag, het probleem van de cliënt	25
2. Het diagnostisch proces	34
3. Wat is gedrag?	37
4. Recht doen aan de cliënt	44
5. Diagnostiseren is gedrag	48
6. De weg van een gedragsconstruct, bijvoorbeeld van de Big Five	58
7. Het einde: Ware beweringen en valide diagnoses	75
<i>Referenties</i>	87
<b>III Dynamiek van het kennisbestand</b>	94
1. Herzien: vervangen, opruimen, erosie, vergeten, muteren, terugkeren	95
2. Aanvullen: constructen toevoegen, nuanceren, herinterpreteren	98
3. Samenvatten: meta-studies	109
4. Toepassen: ons lot verbeteren, of gaan we vanzelf vooruit?	111
5. Toepassen: Diagnose Behandel Combinaties	113
<i>Referenties</i>	128
<b>IV Omgeving van Diagnostiek</b>	134
1. Expliciete theorie tussen impliciete leken- en alternatieve theorie	135
2. Psychometrie en Diagnostiek	140
<i>Referenties</i>	146
<b>V Hoe kom je vooruit?</b>	148
1. Vragen stellen	149
2. Baanbrekend onderzoek	154
3. Zo kom je vooruit	158
<i>Referenties</i>	161
<b>VI Slot</b>	163
Terugdenkend aan psychologische diagnostiek	163
<i>Referenties</i>	175



## Dankwoord

Na het beëindigen van mijn werkzaamheden als docent diagnostiek bij de vakgroep Ontwikkelingspsychologie boden dr. Frank Jan van Dijk, directeur van de Faculteit Sociale Wetenschappen en drs. Mirande van Leuven, directeur van Psychologie mij de gelegenheid om van de faciliteiten van de faculteit gebruik te blijven maken. Ik ben Dr. Annemarie Houkes, directeur van de afdeling Ontwikkelingspsychologie erkentelijk voor haar hulp om dit boek gedrukt te krijgen en te verspreiden. Ik heb baat gehad bij de toegang tot de bibliotheek en van de hulp van het secretariaat.

Van mijn collega's uit Enschede, Nijmegen, Tilburg, Breda, Leiden en Utrecht heb ik veel geleerd. Aan de leden van het hoofdbestuur, de Cotan en de sectie Kinder en Jeugd (K & J) van het Nederlands Instituut van Psychologen (NIP) bewaar ik goede herinneringen.

In Utrecht heb ik me na mijn actieve periode welkom gevoeld bij mijn kamergenoten Jan Boom en Herman van Boxtel en aan het eind van de gang Liesbeth Aleva. De Aio's en junior docenten waren er altijd. Ze gaven thuis als ik om hulp vroeg. Je mist, ouder wordend, de ontwikkelingen met de computer en de talentvolle jeugd springt bij.

De discussies met Jan Boom over wat ontwikkeling en modelleren van ontwikkeling nu echt zijn hebben me geïnspireerd. Met Herman van Boxtel hoef je maar één minuut te praten om een substantieel vraagstuk bij de horens te vatten, ook buiten de psychologie. Ze boden mij bovendien de gelegenheid te vertellen over wat ik nu weer gelezen had en dat was niet alleen psychologie. Poëzie en muziek zijn het mooiste wat er is, maar zonder psychologie ging het - althans tot nu - niet. Hierna wellicht wel: '*...laat mij uit dit land vandaan, laat mij zingend en zonder herinnering het volgende land ingaan...*' (vrij naar 'Het derde land' van Nijhoff).

Bieuwe van der Meulen en zijn collega Daphne van Loon ben ik erkentelijk voor de tijd die ze genomen hebben om de eerste druk van mijn *Elementair Begrip van de Diagnostiek* te lezen en van commentaar te voorzien. Bieuwe stimuleerde me door te gaan met een tweede druk en gaf me waardevolle suggesties. Hij nam de moeite om ook deze tekst te lezen en me op inconsistenties te wijzen. Ik ben verantwoordelijk voor de overgebleven inconsistenties.

Ik dank Arne Evers voor zijn nuchtere kijk op het vak en zijn bemoediging dit soort teksten te maken. Hij is een collega met wie het plezierig is om het oneens te zijn. Peter Heymans gaf mij subtiele hints die ik graag benut heb. Herman van Boxtel wees me op wat ik geschreven had, en voor wie.

*Jan ter Laak*

*Utrecht, zomer, 2015*

## I Inleiding

*Dit geschrift is bedoeld als een afsluitende reflectie op mijn werk als docent diagnostiek bij de afdeling Ontwikkelingspsychologie. In deze Inleiding vermeld ik mijn Werdegang, dat wil zeggen personen en boeken die mij inspireerden, mijn standpunt ten opzichte van psychologie en diagnostiek en ik blik vooruit op de thema's van dit werk.*

## 1. Psychologie studeren en beoefenen

Tot de jaren 60 kwam je via kranten en tijdschriften nauwelijks in aanraking met psychologie. Mijn kennismaking met psychologie was personeelsselectie bij een overheidsinstelling. Waren sollicitanten geschikt voor een baan bij gemeentelijke bedrijven en diensten? Psychologie was het onderzoeken van een sollicitant met behulp van tests, observatie en anamnese. De kwaliteit van het instrumentarium, vooral predictieve kwaliteit, werd bij die instelling onderzocht door experts van de UVA. Bijklussen is van alle tijden. Ik herinner me de eerste predictieve validiteitscoëfficiënt van  $r = .52$ . Deze was met de hand berekend op een groot databestand met onder meer de Differentiële Aanleg Test (DAT) en een intelligentietest als predictoren en met de personeelsbeoordeling als criterium. Een tijdelijk aangestelde ambtenaar/rekenaar deed er bijna twee maanden over. In 1998 kwamen Schmidt en Hunter<sup>1</sup> in een meta-studie naar voorspellende waarde van tests voor werkprestaties tot een gemiddelde waarde van  $r = .50$  na opschoning van de data. Ze vatten in die studie 85 jaar onderzoek samen.

Het selectie-onderzoek mondde uit in een rapport waarbij één geschikte kandidaat werd aanbevolen uit drie sollicitanten. Er was een voorselectie door personeelsfunctionarissen van de bedrijven of diensten. Er was nooit een ex aequo van de kandidaten. Omdat de functie- en opleidingseisen en ervaring goed omschreven waren, verwonderde me dit. Zou het dan in de persoonlijkheid zitten? Ging het er om of iemand 'aardig' was? En, wie kon dat zien? Of, zou het zijn als nu bij functies voor de overheid, waarbij naast opleiding en ervaring 'gevoel voor ambtelijke verhoudingen' een voordeel is? Wat zouden ze daar mee bedoelen? Leek en lijkt me moeilijk te observeren of met behulp van testen te achterhalen.

Voor ik aan de psychologiestudie begon - met wat filosofiestudie en enige praktijkervaring in de personeelsselectie - las ik Duijker et al.'s *Leerboek der Psychologie*<sup>2</sup>. Er stond behoorlijk wat geschiedenis (Palland) en methodologie (De Groot) in. Vuyks ontwikkelingspsychologische bijdrage was voor mij herkenbaar als beschrijvend en fenomenologisch. Later las ik Linschotens<sup>3</sup>, Kouwers<sup>4</sup> en De Groot's<sup>5</sup> magna opera. Deze boeken las ik zonder er tentamen over te hoeven doen. Het is een voordeel zo te mogen lezen en wat aantekeningen maken van zaken die je opvallen en interesseren. Vooral Kouwers *Spel van de persoonlijkheid* sprak me aan.

Ik begon de studie psychologie in Nijmegen belast of bewapend met een mengelmoes van filosofie: haar geschiedenis, Thomistische psychologie, Heidegger<sup>6</sup> en Merleau-Ponty<sup>7</sup>. Verder een Inleiding in de psychologie en een Persoonsleer die dicht bij de existentiële filosofie stond. Methodologie was in mijn ogen nuttig en er was ruimte voor diagnostiek, selectie, testen en voor de *context of discovery*. De subtitel zei dat het boek ging over het *denken* in de sociale wetenschappen. Ik was vooringenomen door Heidegger en Merleau-Ponty die al gauw de term sciëntisme lieten vallen als wetenschap ter sprake kwam en dat was psychologie toch ook. Het stond in ieder geval in de boeken. Je kunt niet aan alles blijven twijfelen. Dat deed ik ook niet, of minder toen ik assistent voor het experimenteel practicum bij Gerrit van Vliet was en daarna bij ontwikkelingspsychologie: zo moest het, zo moest je werkstukken beoordelen, groepen begeleiden en *multiple choice* vragen maken.

In de doctoraalfase werd aandacht besteed aan het afnemen en interpreteren van tests en rapportage. Testtheorie, attitudeschaling en multivariate analyse werden als losstaande vakken onderwezen. Dat kwam me vreemd voor want de Guttmanschaal leek me geschikt voor de ordening van testcores en voor ontwikkelingsschalen. En, hoge ladingen van items op de eerste factor zeiden toch ook iets over de Cronbach alfa van die items. De vakgroep Methoden & Technieken, in Nijmegen Mathematische Psychologie, verzorgde het onderwijs in de statistiek. Naast een boek met een Inleiding in de beschrijvende en toetsende statistiek van Edwards<sup>8</sup> werd variantie- en factoranalyse behandeld. Een computerprogramma, Factor80, werd door een ingenieur onbegrijpelijk uitgelegd. Tachtig variabelen meer kunnen er niet in! Testtheorie werd aan de afzonderlijke vakgroepen overgelaten. Sociale psychologie verzorgde bijvoorbeeld *Techniques of Attitude Scale Construction* van dezelfde Edwards. Ik heb de eerste Edwards geleend aan een student uit Thorn (L.). Nooit teruggezien. De tweede bewaar ik en dat zou ik ook met de eerste gedaan hebben. Deze tekst bevatte een deel van statistiek en testtheorie. In Drenth's<sup>9</sup> *Testtheorie* werd schaling niet apart behandeld. Diagnostiek werd meer met testen, dat wil zeggen onderzoek met behulp van tests dan met meten en schalen verbonden.

Stages waren voor de meeste studenten en voor een deel ook voor mij een klinische of een onderwijspsychologische. Methoden & Techniekstudenten deden de stage meestal 'op' de vakgroep, op een enkele uitzondering na. Vrijwel alle praktijkstages bevatten diagnostiek. Daar leerde je dat de soep niet zo heet gegeten wordt als die wordt opgediend. Betrouwbaarheid en validiteit waren belangrijk maar de tests werden er niet op geselecteerd. Projectietests afnemen mocht eigenlijk niet: subjectief, niet betrouwbaar en hermeneutisch. Het laatste was een eufemisme voor: je verzint iets. De praktisch psycholoog gebruikte ze echter regelmatig. Je deed gewoon mee met de Boomtest ('Teken een droom-, een fantasieboom, een mens...'), Zinaanvultest, *Wartegg* en de *Thematic Apperception Test*. Je mocht soms meekijken bij de afname en interpretatie van de Rorschachtest. Bij een veelkleurige plaat werd het antwoord: 'Parijs bij nacht' gescoord als een creatief antwoord. Wie was er in die tijd in Parijs geweest of had er belangstelling voor? Het kan verkeren. Na het Exner scoringssysteem is de betrouwbaarheid van de Rorschach op orde en meta-studies laten een predictieve validiteit zien die gelijk is aan die van de bekende vragenlijsten met r-waarden rond de .30 (Meyer et al.<sup>10</sup>) voor een aantal specifieke criteria. Er zijn nu auteurs die de Rorschach ten voorbeeld stellen aan andere wijzen van diagnosticeren van de persoonlijkheid, emotie en cognitie van de cliënt. Het kan in de testpsychologie alle kanten op.

Aan het werk met psychologie als studie/vak/beroep betekende in eerste instantie ontwikkelingspsychologie beoefenen. Ik las Duitse teksten over het bestaan en de aard van ontwikkelingsstadia. Mönks legde dit uit. Aandacht voor stadia verdween daarna en vlamde even op toen de chaostheorie kort voet aan de grond kreeg. Nu is de belangstelling weer verdwenen. Het bestaan van stadia is een formeel modelleringsvraagstuk geworden. Of je er stadia in ziet is jouw interpretatie en die doet er niet zoveel toe. Het gaat er om of het model aanvaardbaar is. Daarnaast was er onderwijspsychologie: een amalgaam van Rogeriaans *Leren in Vrijheid*, *Testtheorie* van Drenth (en Sijtsma), Onderwijsdoelstellingen (Bloom et al.<sup>11</sup>) meestal hiërarch-

chisch geordend en de opbouw en inhoud van curricula. Daarna volgde een confrontatie met de klassieke testtheorie en de persoonsleer. Magnussons<sup>12</sup> *Test Theory* bevatte een passage die me verwonderde: complete validering met behulp van constructvaliditeitsonderzoek. Een poging die Messick<sup>13</sup> herhaalde in een moeilijk hoofdstuk in Lynn et al. (1989). Dat kan niet, dacht ik, want dan zou de wetenschap afgerond zijn. Bij persoonsleer lag de nadruk op individuele verschillen en niet op ontwikkeling waar ik me voorheen mee bezighield. Het persoon-situatie debat stelde de situatie centraal. 'Trekken (traits) bestaan niet' luidde de slagzin. Ondertussen gingen we gewoon door met trekvragenlijsten.

Door herinnering aan de colleges van Strasser<sup>14</sup> over filosofie en empirische menskunde kwam ik op een driedeling in theorievorming: Individuele verschillen, Ontwikkeling en (sociale) Context. Daar passen alle theorieën in inclusief niet *mainstream* alternatieve. Strasser sprak over een niet te ontkennen oer-evidentie: ik bevind me hier en nu samen met anderen. Wat een 'on-zin' zei een Leidse collega toen ik dat opschreef. Hij bedoelde dat - hoop ik - redactioneel: een niet goed lopende zin dus. Niettemin paste ik de 'on-zin' toe en kwam tot een indeling in drie theoretische oriëntaties.

Van de testtheorie en persoonsleer via de praktijk van het schooladvieswerk - echt testen en leraren bijstaan om lastige jongens en meisjes in het gareel te houden - weer terug naar de ontwikkelingspsychologie. Die was in de jaren '70 tot '90 toegepast en ze moest nuttig zijn voor de mensen. Ondanks de vrolijkheid van die tijd heb ik me zelden aan de gedachte kunnen onttrekken dat de uitbuiting van de mens door de mens het motief van veel handelen was, is en zal blijven. Barendregt<sup>15</sup> had met zijn Theophrastes hertaling op geestige wijze laten zien dat psychologen nogal gewichtig deden. Bovendien wees hij er op dat toegankelijkheid van psychologie en therapie bedenkelijke kanten hadden. Zijn boek *De Zielenmarkt* gaat daar over. De subtitel luidt: *Over psychotherapie in alle ernst*.

Mijn 'terug tot'<sup>16</sup> ontwikkelingspsychologie was een tocht naar bezet gebied. Alles was verdeeld en diagnostiek was in handen van een goeroe, leermeester. Methodologie bleef over. Wat maak je daarvan, als psychologie nuttig en toegepast moet zijn? *Nobody cared*, dus maar iets gedaan: Experimentele en quasi-experimentele designs in samenwerking met Leo van der Kamp van Methoden & Technieken. Cook en Campbell<sup>17</sup> probeerden *social experimentation* binnen het causaal empirisch-analytische kader te houden. Het ademde de geest van het verbeteren van ons lot. Hulpverlening stond centraal en dat is meestal N = 1 onderzoek met eigen regels. Bij gedragsmodificatie was dat uitgewerkt, maar naast *eye balling* - verdedigd door Cohen<sup>18</sup> van *The world is round; p < .05* - kwamen er ingewikkelde technieken aan zoals regressie-discontinuïteits designs en *AutoRegressive Integrated Moving Average* modellen van Glass et al.<sup>19</sup>.

Een ontwikkelingspsycholoog leert observeren. Het practicum was een corvee maar nuttig voor docenten en studenten. De laatsten maakten werkstukken in gewaagde omgevingen. Speel maar eens een droogstaande drugsverslaafde op het CS in Leiden en maak een observatieschema voor wat voorbijgangers zeggen en doen. Observatie van gedrag in het echt is uitdagend. Dat geldt zowel voor het observeren zelf als voor de analyse. Ik gebruikte een tekst van Lamb et al.<sup>20</sup> waarin dat naar mijn oordeel goed uitgewerkt werd. Analyse van sequentiële afhankelijkheden was lastig (instabiele resultaten; het hing er van af, waar je begon) en leverde inhoudelijk en theoretisch niet zoveel op. Ja, alles hangt met alles een

beetje samen maar daar zitten wetenschap en praktijk niet op te wachten. Het gaat om robuuste effecten en verbanden.

Van 1985 tot 2000 was de trend sociaalwetenschappelijke afdelingen te organiseren rond maatschappelijk relevante thema's. De werkwijze zou interdisciplinair onderzoek bevorderen en de samenleving had er ook nog wat aan. Ik kwam terecht bij de Utrechtse Kinder- studies, een combinatie van Ontwikkelingspsychologie en Pedagogiek, nu voor het eerst met diagnostiek als opdracht in het kandidaats/bachelor onderwijs. Er werd gedacht dat epidemiologisch onderzoek naar stoornissen bij kinderen erbij hoorde en uitvoerbaar was. Dat was veel te groot, veel te weinig mensen om de samenleving te laten weten hoe het met het lijden in de kindertijd gesteld is, met dat in de volwassenheid en in de ouderdom: die *kleine, stille wildernis...*, waar lang vervlogen wel en wee zijn intrek neemt als het avond is.

Het ontbreken van een eigen onderzoeksdomein leidde tot een poging er een te claimen. Zo gaat dat bij wetenschappelijke bedrijven. Liefst een inhoudelijk terrein dat voor een tijdje in de belangstelling staat, zoiets als de *Big Five*. De keuze voor persoonsontwikkeling was een ongelukkige. Er was en is geen *undergraduate* handboek en syllabi, *readers* mochten niet meer van de uitgevers. Het APA boek '*Can Personality Change*'<sup>21</sup> ging nauwelijks over persoonsontwikkeling, eerder over stabiliteit. Er werden twee veranderingen beschreven: een *quantum leap*, de druppel die de emmer doet overlopen en een omslag in de vorm van een en bekering. Dat was een EO-achtige verandering, bijvoorbeeld van de drugs naar God. Het omgekeerde komt ook voor maar daar had de EO het niet over. Een themanummer van het *European Journal of Personality* van 2015 over persoonsontwikkeling gaat *McCroftiaans* vooral over stabiliteit. Erikson en Piaget zijn niet populair. Erikson wordt weliswaar een gigant (Denissen<sup>22</sup>) genoemd, maar beiden komen in het persoonsontwikkelingsverhaal nauwelijks voor. Asendorpf<sup>23</sup> behandelt Erikson niet in zijn persoonsleerboek. De aandacht voor Piaget kwijnt weg. Een recent handboek over cognitieve ontwikkeling van Goswami<sup>24</sup> vermeldt van meet af aan dat Piaget er naast zit.

Als een onderzoeker ontwikkeling van gedrag bestudeert, betekent dat nu het passen van een model. Een dergelijk model verklaart doorgaans evenals de *r*- en *d*-waarden bij correlatieel en experimenteel onderzoek: tussen de 20% en 40% van de variantie. Als gedragsontwikkeling inhoudelijk aan de orde is, bijvoorbeeld bij een criminele loopbaan, is de interpretatie vaak voordehandliggend: het wordt erger, het neemt af, of het blijft gelijk. Studies en essays laten zien dat leken en ontwikkelingspsychologen veranderingen over tijd variabel lezen. Adolph et al.<sup>25</sup> (p. 528) vatten ze samen: lineair, versnellend, asymptotisch, stapsgewijs (stadia), *S*-vormig, variabel en onsystematisch, *stair-climbing*, *U*-vormig en omgekeerd *U*-vormig. Deze veranderingen worden gewaardeerd als goed of slecht, passend of niet passend en als normaal of afwijkend.

Je zit niet altijd binnen. Er zijn congressen en buitenlandse collega's. Op congressen heb ik me niet thuis gevoeld. Psychologen zijn met velen. Probeer daar maar eens bij te horen en gehoord te worden. De organisatie laat vooraf weten wie iets te melden heeft door *key note speakers* aan te wijzen. En, leerde je er inhoudelijk iets dat je niet in de boeken en tijdschriften kon vinden? Je hoorde soms een verhaal dat je een half jaar eerder gelezen had, ook al werd geclaimd, dat het een studie met *fresh data* was. En, dat verse associeerde ik recalcitrant met

een opmerking van Heidegger over *Sein und Zeit*. Het ging er niet om of zijn boek nieuws bevatte, *...sondern wenn es alt genug währe*. Een eigen verhaal<sup>26</sup>, een observatiestudie van Piet Maas en mij onder begeleiding van Bert Hermans over de ontwikkeling van prestatiemotivatie en faalangst bij 9- tot 12-jarigen werd weliswaar in het *Annual Review of Psychology* van 1974 besproken maar er werd bij gezegd dat de resultaten al bekend waren in de VS.

Buitenlandse collega's waren waardevol. Ze deden me inzien dat psychologie een westerse aangelegenheid is, gedomineerd door de VS. De bezoeken lieten zien dat veel mensen iets anders aan hun hoofd hebben als psychologie. Ze is in veel landen een elitesport. Nu ziet het er naar uit dat, na een periode van aandacht voor de VS psychologie, de Zuid-Amerikaanse, Russische, Chinese en Indiase psychologen langzaam hun eigen weg gaan. Die samenlevingen zien er anders uit en het vak heeft daar een andere positie. De economische winnaar van de toekomst doet misschien de dominantie van de VS afnemen. Het tegenovergestelde kan zich ook voordoen. In dat geval wordt de psychologie regionaler. Eigenlijk is ze dat al, gezien de Amerikaanse dominantie. Het is aandoenlijk te lezen hoezeer onderzoekers uit vele landen van de Filippijnen tot België hun best doen om op de Amerikaanse normatieve structuur van de *Big Five* te lijken. Groningen ligt een beetje dwars met het AB5C model en twee factoren extra. Daarnaast is er een ander Gronings model met één factor minder en een andere interpretatie van de factoren<sup>27</sup>. Van Kampen<sup>28</sup> van de VU legt zich eveneens niet neer bij de hegemonie van de *Big Five*. Hij gaat terug naar klassieke psychologen die een theorie als basis formuleerden voor de Vijf plus of min Twee á Drie.

## 2. Inspiratie

*With a little help of my friends* zongen de Beatles, populair in mijn studententijd, op zoek naar onze *Jude* (m/v). Iedere beroepsuitoefenaar staat op de schouders van anderen. De wetenschapsfilosoof Hacking<sup>29</sup> zei dat originele wetenschappers, zelfs Nobelprijswinnaars, regelmatig de bron niet citeerden die hen aan hun inzichten hielp. Watson en Crick vergaten hun assistente in het laboratorium te noemen die de helixstructuur (spiraalvorm) van het DNA als eerste zag. Binet nam items van Blin over. Pythagoras heeft 'zijn' stelling niet zelf ontdekt.

Ieder kent professionals en situaties die helpen je thuis te voelen in een vak en je aansporen en de vrijheid geven om er iets van te maken. Het waren geen *friends* - ik heb ze op twee na nooit ontmoet - maar het was inspirerend ze te lezen. Andere zijn vergeten bij mijn onzorgvuldige *récherche* naar de *temps perdu*. Elk geheugen is immers leugenachtig, alleen al door te selecteren en door zaken mooier, of als het zo uitkomt, lelijker voor te stellen dan ze zijn. Inspiratie voor uitwerking van psychologische diagnostiek en diagnostiek van de psychologie zijn werkzaamheden in praktijk (cliënten en sollicitanten: kinderen en volwassenen), onderwijs, kennismaking met de geschiedenis van de filosofie, het werk van James<sup>30</sup> (*Principles of Psychology*) dat Rutten in Nijmegen aanbeval. 'William is de broer van Henry', zei hij er bij. Hij was behalve in de psychologie gepromoveerd in de letteren. Later las ik James's<sup>31</sup> boek over *Religious Experience* naar aanleiding van colleges Cultuur- en Godsdienstpsychologie van Fortmann en Kempen, opnieuw Merleau-Ponty (*Phénoménologie de la Perception*), Heidegger (*Sein und Zeit*) - beide niet helemaal: veel te moeilijk Frans en Duits - en monografieën van Strasser, Linschoten, Kouwer en De Groot uit de jaren zestig. Deze denkers en psychologen/pioniers namen een standpunt in en gaven de psychologie een eigen vorm en inhoud. Ze zijn lang genoeg geschiedenis om met afstand naar hun opvattingen te kijken. Je kunt er met rede en een gerust hart een draai aan geven. Je hoeft niet blijven steken in de *sixties*.

Strasser wilde een bijdrage leveren aan een ideaal van wetenschap in zijn *Fenomenologie en Empirische Menskunde* van 1964. Een ideaal is iets dat men verwezenlijkt hoopt te zien. Het kan voortdurend veranderen door oude en nieuwe inzichten. Je kunt er een contrafactisch streven aan ontlenuen om de vraag van de cliënt te beantwoorden: dat wil zeggen een definitief antwoord komt er nooit, maar we blijven ernaar streven om te komen tot een rijk, genuanceerd en voldoende begrip of verklaring voor/van zijn gedrag. 'Contrafactisch' komt vaker voor in deze tekst. Het begrip hangt samen met objectiviteit als regulatief principe. Objectiviteit verwijst naar wat je onderzoekt en hoe je dat moet doen, dat wil zeggen zonder vertekeningen, vooroordelen en eigenbelang. Het lukt je nooit het object van onderzoek geheel en al recht te doen. Een regulatief principe is daarom contrafactisch: een streven dat je volhoudt 'tegen de feiten in'.

Strassers<sup>32</sup> idee van een waarheidsgeweten legde nadruk op een relatie tussen methode en inhoud. Het Streng Methodologische Geweten is mooi, maar niet genoeg. Dit geweten had in de Nijmeegse empirisch-analytische benadering van de jaren 60 en 70 bijna het karakter van het tegengaan van de zonde.

Linschotens *Idolen van de Psycholoog* van 1964 zorgde er voor dat psychologie en haar beoefe-

naars een zelfreinigend vermogen verkregen. Hij zegt het zo: *'Als de psycholoog wetenschap wil beoefenen dient hij zich daar naar te gedragen. De regels van het spel in acht nemen. Meten en registreren. Verborgen vooronderstellingen opsporen, eigen vooroordelen ontmaskeren. Alledaagse opvattingen met het grootste wantrouwen te bekijken. De mensbeelden die ons in wijsgerige en religieuze tradities worden aangereikt, worden tot idolen wanneer de wetenschappelijke onderzoeker ze onkritisch aanvaardt en er zijn theorieën op bouwt'*. Linschoten keerde op zijn fenomenologische schreden terug. In zijn streven naar een wetenschappelijke psychologie werd de methode centraal voor het borgen van objectiviteit. Diagnostiek wordt daardoor - strikt doorgedacht - het werk van een klerk, een robot. Hij spreekt letterlijk over een robot. Individuele verschillen tussen diagnostici zijn echter onontkoombaar, ze ontwikkelen zich en ondergaan de invloed van hun sociale context. Daar is ook empirisch onderzoek naar verricht. Vooroordelen, of wat leken denken dat de oorzaak van gedrag is, dat wil zeggen hun impliciete theorieën vooraf en hun rechtvaardigingen achteraf worden door de diagnosticus serieus genomen. Ze hebben hun eigen opvattingen over hun gedrag. Het gaat er niet om of die juist, valide zijn in de empirisch-analytische zin. De kunst is er achter te komen welke rol ze vervullen bij de vraag/het probleem. Het heeft meestal geen zin tegen de cliënt te zeggen dat zijn opvatting over en verklaring van zijn problematisch gedrag niet deugen in het licht van het kennisbestand van de psychologie.

Kouwer liet in zijn *Spel van de Persoonlijkheid* zien dat men voor theorie over de persoon uit veel bronnen - van belletrie tot schilderkunst - mag putten, al werden zijn belangstelling en eruditie bestraft door de psychiater Kuiper<sup>33</sup>. Deze zei over het *Spel van de Persoonlijkheid* dat Kouwers belezenheid wel tot wat meer had mogen leiden. Die straf moet je voor lief nemen. En, voor mij relevanter, bij Kouwer mag je onzeker zijn over wat een persoon is: iemand met een eigen verhaal, een onzekere identiteit en die terugpraat als een wetenschapper of wie dan ook hem karakteriseert. Hij laat de lezer in de laatste drie mooie hoofdstukken van zijn *Spel* in verwarring achter. Het stichten van verwarring daagt uit om minimale orde te scheppen. Dat heb ik gedaan door het formuleren van drie elementen, drie bronnen en drie theoretische oriëntaties die elkaar volledig kruisen. Kouwer hielp de schroom te overwinnen om filosofie en literatuur te benutten zonder vakfilosoof of schrijver/essayist te zijn. Het is niet verboden filosofie en belletrie te lezen al valt het omgevallen-boekenkast-verwijt gauw als je deze in denken, lezen en schrijven een plaats geeft.

De Groot heeft een *Methodologie* geschreven waarin regels en voorschriften voor het denken over- en verrichten van onderzoek staan zonder de *context of discovery* te veronachtzamen. Deze krijgt volgens sommige lezers, bijvoorbeeld Smaling<sup>34</sup>, te beperkt aandacht. Ik sluit me daar bij aan. Hij heeft een wetenschappelijke psychologie/diagnostiek willen grondvesten: 'grondslagen' staat er in de subtitel. Evenals Kouwer die een boek over de psychologische tests schreef, was hij op de hoogte van de praktijk van diagnostiek. Zijn *Methodologie* was in eerste instantie bedoeld als handboek voor diagnostiek. Objectiviteit is door hem voornamelijk methodologisch uitgewerkt. Daardoor zijn proefleiders bij adequaat ingericht experimenteel onderzoek inwisselbaar. Je kunt er aan toevoegen dat diagnostiek door een klerk bedreven kan worden als hij beschikt over valide instrumenten en diagnostische protocollen

stipt uitvoert. De Groot spreekt daarnaast over objectiviteit als recht doen aan het object van onderzoek. Deze uitdrukking wordt hier gekoesterd. Objectiviteit is echter een veelkoppig, meerduidig begrip. Het kan verwijzen naar verschillende uiteenzettingen over de relatie tussen de uitspraak over gedrag en het gedrag zelf. Hier wordt naast de empirisch-analytische ook een hermeneutische en kritische uitleg toegestaan. Deze worden van belang geacht voor een rijke en voldoende beschrijving van de vraag van de cliënt. Objectiviteit is ook als houding, attitude uit te leggen. Ik kom erop terug.

### 3. Standpunten

Van Parreren zei eens in de wandelgangen dat Pedagogiek en Ontwikkelingspsychologie bij elkaar hoorden. De Utrechtse keuze voor 'Kinderstudies' was volgens hem juist. Pedagogiek is een handelingswetenschap. Ontwikkelingspsychologie is dat mijns inziens *niet*. Ze is een vak met een eigen formeel object en geen beperking wat betreft de populatie, bijvoorbeeld alleen kinderen, en gedragsdomein, bijvoorbeeld disfuncties. Ik stem daarom niet in met de omschrijving van diagnostiek door Witteman et al.<sup>35</sup>. De twee disciplines gingen niet door één deur en werden na enige tijd ontvlochten. Als vaak ligt dat eerder aan mensen dan aan disciplines. Dat resulteerde voor mij in *diagnostiek als een activiteit zonder formeel object (geen bepaald perspectief) en zonder eigen materieel object (geen specifiek gedragsdomein of populatie) die geldige uitspraken wil doen over het gedrag van één cliënt met behulp van het kennisbestand van de psychologie*. Dit standpunt werk ik verderop uit. Diagnostiek is zo beschouwd net zo zwak/sterk als het kennisbestand van de psychologie. Het Kennisbestand, ga er maar aan staan. Of dat gelukt is in mijn tekstboek *Elementair begrip van de Psychologische Diagnostiek*<sup>36</sup> moet de lezer beoordelen. De tekst zegt niet wat de diagnosticus moet doen en er wordt van alles *bijgehaald*. Hij sluit *niet* aan bij de APA en EFPA standaarden voor onderwijsstoetsen en psychologische tests. Deze leggen nadruk op regels hoe te diagnosticeren en omzeilen de vraag wat diagnostiek is.

Psychologiestudenten, -docenten en practici moeten zelf een weg zoeken in hun opleiding en daarna in hun beroepsuitoefening. Misschien geldt dat voor veel disciplines. Voor mij kwamen enkele kenmerken naar voren die ik op diagnostiek geprojecteerd heb: ze is geen handelingswetenschap, er is geen sprake van *agein*, maar van toegepaste wetenschap zonder een specifiek domein, perspectief en populatie. Wetenschapstheoretisch is diagnostiek een hybride: naast de logisch positivistische, empirisch-analytische bewijsvoering spelen hermeneutisch interpreteren en kritisch analyseren een rol. Psychologisch theoretisch en epistemologisch houdt diagnostiek alle ballen in de lucht: theorie over individuele verschillen, ontwikkeling en de sociale context en de empirisch-analytische, hermeneutisch/constructivistische en kritische stromingen. Voor de laatste stroming denk ik aan de Kritische Psychologie. Ik vatte dat op als een Duits antwoord op de hegemonie en eenzijdigheid van het behaviorisme: 'metafysica bedreven met normproefpersonen'. Duitsland had in en na WO II zijn beste psychologen verloren: vermoord (ondermeer Selz), geëmigreerd (ondermeer Brunswik en Lewin) en de in Duitsland achtergeblevenen zwegen na de misdaden van het Derde Rijk. Verder kon ik in de psychologie geen Popperiaanse cruciale experimenten ontdekken die een theorie omverwierpen als ongeldig, onwaar. Er is daarom mijns inziens *geen* hiërarchie tussen theorieën. Ze kunnen slechts *contrafactisch* naar overeenstemming streven. Daarom blijven wij met alle theorieën zitten en onze tekstboeken blijven ze herhalen. De reactie van studenten luidt dat al die theorieën iets anders zeggen: 'Het is zo onduidelijk'. Ze vragen als ware kinderen van Descartes om heldere en onderscheiden (*claires et distinctes*) ideeën. Er is geen argument om de hermeneutische en kritische stromingen te verwerpen ook al wordt gesuggereerd dat ze niets zinvols te vertellen hebben. Bunge<sup>37</sup> en Bèrmùdez<sup>38</sup> zijn filosofen

die over en voor psychologie schrijven. Ze komen weliswaar dicht bij de logisch-positivistisch, empirisch-analytische stroming uit, maar sluiten niets uit. Chalmers<sup>39</sup> maakt geen keuze en legt enkele mogelijkheden voor. Het verwerven van kennis over ons gedrag en denken is een *persistent problem situation*, zegt hij.

De cliënt mag rekenen op een 'rijke en voldoende beschrijving' van zijn vraag/probleem die hem 'recht doet', zeg ik Simon<sup>39</sup> en De Groot na. Dat is evenwichtskunst, vergt maat houden, inhoudelijk relevant zijn en openstaan voor de vele lagen van gedrag, zowel wat zijn oorzaken c.q. redenen als wat zijn inhoud betreft. Er zijn in de werkelijkheid, die de cliënt is, niet alleen responsies maar ook intenties en betekenisverleningen. De mens is zijn gedrag en gedrag is ongeveer alles: *anima est quodammodo omnia* zei Thomas van Aquino. Deze regel komt uit de Scholastische lijst met filosofische axioma's (derde rubriek: 'Psychologia', de negende van de 16). In geen psychologieboek terug te vinden; er staan aardige tussen. Mijn leraar Latijn vertelde dat we *anima mea* met 'ik' konden vertalen.

#### 4. Vooruitblik

Deze tekst is niet gepland. Enkele zaken die me inhoudelijk en/of methodologisch interesseerden zijn bijeengezet vanuit gedachten of stemmingen zoals vind ik dit een aardig onderwerp, moet ik die methodologie er wel of niet in zetten, bijvoorbeeld het gebed zonder eind over nulhypotesetoetsing, de hegemonie van Rasch<sup>40</sup> en de Amerikaanse psychologie? Een 'raar maar waar rubriek' maken over IQ en persoonlijkheid? Is psychologie wel zo internationaal als ze pretendeert te zijn? Hoe heeft de psychologie haar geschiedenis vorm gegeven? Zal ik iets proberen te zeggen over de druk waaronder de psychologie en haar beoefenaren staan naar aanleiding van fraude en de oordelen van studenten en publiek? Iets over de geschiedenis van het werken op de universiteit tussen de jaren 60-80 en nu: van 'alles mag en niets moet', naar 'alles moet en bijna niets mag'? Of, opmerkingen plaatsen bij het veranderende denken over onderwijs en de waardering van studenten van onderwijs? Een pas afgestudeerde psycholoog - het zijn meestal mannen die commentaar geven - zei in het Utrechts Universiteitsblad, verwijzend naar een hoorcollege: 'Iemand die 8 x 2 uur loopt te schreeuwen'. En, 'Je hoeft alleen uittreksels te lezen voor een voldoende', 'De meeste studenten zijn ongemotiveerd'. Geen docent zou dat over zijn kant moeten laten gaan.

Al kwam deze tekst niet diagnostisch-proces-achtig getrap tot stand, *achteraf* kun je orde maken. Evenals Nijhoff<sup>41</sup> '*...weet ik in iedren oogopslag een ernst te liegen*'. Ik houd me in deze tekst niet aan de regels, of zo u wilt, het schrikbewind van het *APA Manual* voor publicaties. Het '*...hebt gij dit van uzelf, of hebben anderen u dit verteld*', kan nagegaan worden in referenties. Ze zijn bedoeld als illustratie en als erkenning dat ik het allemaal 'niet van mij zelve heb'.

Ik begin met bronnen voor mijn vormgeving van diagnostiek. Dat zijn mensen met stevige schouders. Ze rusten al lang - jong gestorven - of sinds kort. Ze kunnen me dus niet meer vertellen of het wel kan, wat ik zeg. Ze zouden het ook niet doen: belangrijker zaken aan hun hoofden. En, *quod licet lovi non licet bovi*, zo hield een van mijn eerste van de VU afkomstige collega's me voor, als ik het met een autoriteit niet eens was.

Na de inleiding werk ik een viertal thema's uit. Ze zijn gemerkt met Romeinse cijfers. De secties van de thema's zijn Arabisch genummerd. Het eerste thema bevat een standpunt over diagnostiek en psychologie. Het start met de vraag/het probleem van de cliënt en eindigt met wat een ware bewering over en een valide diagnose van het gedrag van de cliënt is. De *eerste* sectie gaat over wat diagnostiek en psychologie zijn en voor elkaar betekenen. Dat is een filosofische en een gewetensvraag. Vermijden dus, maar je moet er toch iets mee. Diagnostiek mag dan geen zelfstandige subdiscipline zijn, ze wordt ondertussen onderwezen bij vrijwel iedere afdeling. Er is geen onderzoekslijn met 'Diagnostiek' in de titel. Ook daar geen houvast. De Groot's *Methodologie* van 1961 verbond de twee nog maar dat is verdwenen. Proefschriften over testconstructie zoals in de jaren 60 tot 90 zijn nu zeldzaam. En, testconstructie en testen zijn niet hetzelfde als diagnostiek. Deze stand van zaken maakt de weg vrij om een omschrijving te verzinnen, om er iets van te maken: *poein*. Deze omschrijving moet diagnostiek van de cliënt een vorm en een referentiekader geven. Ze moet in staat stellen het relevante uit het kennisbestand van de psychologie te benutten. Dat kennisbestand is een

*mer à boire* en moet bepaald worden door de diagnosticus, gegeven de vraag van de cliënt. De *tweede* sectie gaat het over wat de diagnosticus met die kennis doet en hoe hij zich verhoudt ten opzichte van de cliënt. Dit is het diagnostisch proces. Daarvoor zijn richtlijnen, maar uiteindelijk is hij degene die inhoud, verloop en verhouding bepaalt.

Diagnostiek gaat over het gedrag van één cliënt, maar wat is dat- gedrag? In de *derde* sectie ga ik daar op in. Uit de boeken leer je genoeg te nemen met een werkdefinitie van gedrag, maar wat hoort er bij en wat niet? Kun je gedrag naar believen inperken (waarneembare responsen) of uitbreiden (onderbewuste motieven)? En hoe zit het met de veroorzakers, factoren, doeners, daders en redenen van en voor gedrag? Kun je die ook naar believen inperken (stimulus) en uitbreiden (van biologische processen tot cultuur)? Kan de psychologie daarnaast naar willekeur te rade gaan bij andere wetenschappen van moleculair (DNA, neurologie, genen) tot molair (sociologie en culturele antropologie)? In onderzoek kun je, mag je en moet je je beperken tot één omschreven vraag naar een effect of samenhang, maar kan dat bij de vraag, het probleem van een cliënt? De slagzin *Scientia non est Individuorum* zegt dat onderzoek bij één cliënt geen wetenschap is. Moet je je daar iets van aantrekken?

Gedrag willen we objectief en valide beschrijven. Objectief wordt hier aansluitend bij De Groot opgevat als recht doen aan de cliënt. Dat is het onderwerp van de *vierde* sectie.

Gedrag is ongeveer alles en dat betekent dat diagnosticeren zelf ook gedrag is. Dat houdt in dat diagnostici individueel verschillen, zich ontwikkelen en beïnvloed worden door hun omstandigheden en sociale context. Dat werk ik uit in de *vijfde* sectie. Het product, de diagnose staat niet los van de diagnosticus.

In de *zesde* sectie wordt gekozen voor de gedragseenheid: het construct, het gedragsbegrip. Enkele psychometrici vinden dat een taalspel: *in the eye of the beholder*. Roskam zei eens dat meelesen van dissertaties in de psychologie journalistiek was of nederiger, het werk van een corrector. In Wittgensteins *Philosophical Investigations*<sup>42</sup> ligt de inhoud van een construct niet vast. Dat is ook het geval in psychologie en taalkunde. In constructvalideringsonderzoek doen we ons best om een construct te grijpen. We leggen het op een Procrustesbed. Denk aan Cronbach en Meehls<sup>43</sup> en Campbell en Fiskes<sup>44</sup> werk van de jaren 50 en 60. Als je dat niet doet, kun je zoals Messick (1989) verdwaald raken in interpretaties. Hoe is de levensloop van een gedragsconstruct, bijvoorbeeld van de Big Five? Waarheen leidt de weg die populaire gedragsconstructen moeten gaan?

Het einde van de diagnostische activiteit is in een ware bewering over, een valide diagnose van de cliënt. Dat is het onderwerp van de *zevende* sectie. Zo is het eerste thema rondgebred: psychologie komt uit de filosofie voort en gaat er naar terug als ze ware beweringen wil doen over een cliënt.

Het eerste thema gaat over de psycholoog/diagnosticus die naar zijn eigen activiteiten kijkt en zich afvraagt wat het is dat hij doet en of het tot een geldige uitspraak over de cliënt kan leiden. Hij kan er niet omheen dat psychologie en diagnostiek zelfbetrokkenheid activiteiten zijn. Het Cartesiaanse onderscheid tussen het Ik (de diagnosticus) en zijn object het Niet-Ik (de cliënt) schuurt in de diagnostiek en ook in de psychologie met inbegrip van de functieleer. Dat geldt voor meer, zo niet alle wetenschappen, we verzinnen die immers zelf.

Aan zelfbetrokkenheid wordt gauw het epitheton 'subjectief' gekoppeld: 'U doet subjectieve, niet betrouwbare en niet valide uitspraken over het gedrag van uw cliënt'. Mogelijk speelt daarom *methodologie* zo'n prominente rol in de psychologie. In een poging aan te sluiten bij de exacte, objectieve wetenschappen is gezocht naar gedragswetten. In Nijmegen werd Hull, de Newton van de psychologie, behandeld door Kremers. Hij nam enige afstand van Hull. Tolman was zijn favoriet. Een tweede reden om aan te sluiten bij de exacte wetenschap is serieus genomen willen worden. Die is immers objectief op objecten en verschijnselen buiten jezelf gericht. Het streven naar objectiviteit leidde bij Cronbach<sup>45</sup> in *Essentials of Psychological Testing* tot de diagnosticus als *clerk*. Linschoten sprak het over de onderzoeker als *robot* en De Groot over een goed geïnstrueerde *klerk*. Het hypothese toetsend model (HTM) vooronderstelt dat diagnostici inwisselbaar zijn, als ze het HTM, een objectieve procedure strikt volgen. Kan of doet een diagnosticus dat?

Het tweede thema gaat over het kennisbestand van de psychologie. Zo neem ik enige afstand van het diagnosticeren sec en kijk naar de psychologie als leverancier van kennis over het gedrag van de cliënt. Wat is de dynamiek van dat bestand? Wordt het regelmatig herzien (sectie 1)? Worden nieuwe constructen toegevoegd (sectie 2)? Hoe wordt de stroom aan empirisch onderzoek in de hand gehouden? Wat leveren meta-studies op (sectie 3)? Het kennisbestand wordt toegepast. Zijn er toepassingen die ons lot verbeteren, of is dat naïef, of gaat het vanzelf (sectie 4)? Een specifieke combinatie van diagnostische en psychologische kennis wordt gevormd in de Diagnose Behandel Combinaties (DBC's: sectie 5). Hoe sterk zijn Diagnoses en Behandelingen en hoe sterk hun Combinaties?

Het derde thema betreft de omgeving waar diagnostiek zich afspeelt, weer een stapje verder weg van het diagnosticeren sec. Ik kijk naar de burens en zoals we bij 'De Rijdende Rechter' zien, gaat dat er soms ruw aan toe. Zijn de subdisciplines goede burens, werken ze samen of is het een jungle? 'Wat is de wet van de jungle' vroeg Bagheera: 'Eerst slaan en dan waarschuwen'. Dit zinnetje komt uit het jungle boek van Kipling, waar de zwarte tijger Bagheera Mowgli naar de mensenwereld brengt. Het verhaal ken ik van de Disney film. Een enkele keer gaat iemand er met gestrekt been in. Een voorbeeld: naar aanleiding van de aanval op *priming* experimenten zijn 28 commentaren geschreven op een kritisch artikel van Newell en Shanks<sup>46</sup>. Eén is van Dijksterhuis et al.<sup>47</sup>: *It - het artikel van Newell en Shanks - is neither a theoretical article, as it lacks a theory, nor it is a good review article, because it is biased and selective en verder the degree of cherry picking is too extreme* (p. 25). De auteurs blijven de rol van *Unconscious Thought* (UT) benadrukken en vinden steun in studies waar *Unconscious Thought Theory* (UTT) verbonden wordt met de *fuzzy trace theory* en met evidentie voor onbewust denken uit MRI scans. Ze besluiten met: *Although we surely agree that the road to progress in this field is rocky, focusing on consciousness without understanding its unconscious precursors is a dead end*. De burens in de wetenschap bijten elkaar af en toe.

De omgeving, de buurt is hier bepaald als impliciete alledaagse kennis, inclusief het oordeel van het publiek en alternatieve theorie (sectie 1). Het publiek kan het kennisbestand als gepraat, geklets afdoen. De polemist Dautzenberg<sup>48</sup> vindt sociale en economische wetenschappen saai en bedrieglijk. Op zijn hoogst bieden ze een enkel handvat, maar dan heb je

het ook wel gehad. In sectie 2 komt de psychometrie aan de orde. Ze speelt vooral een rol omdat de diagnosticus tests gebruikt. Je kunt psychometrie opvatten als een minitheorie (bijvoorbeeld een item respons functie van een bepaalde vorm) over een minigedrag (het antwoord op een item). Zo beschouwd staat ze neutraal ten opzichte van diagnostiek. Dat neutrale ervaart de diagnosticus doorgaans niet. De psychometricus kijkt over zijn schouder mee naar de kwaliteit van zijn instrumentarium. De psalmdichter<sup>49</sup> (no. 3, vers 1) beschrijft de burenen zo: *'Heer, hoe talrijk zijn mijn belagers, velen vallen mij aan, velen zeggen van mij: God zal hem niet redden'*. En in psalm no. 22 vers 5: *'Een troep stieren staat om mij heen, buffels van Basam omsingelen mij, roofzuchtige, brullende leeuwen sperren hun muil naar mij open'*. Zo erg is het niet, maar alledaagse kennis, alternatieve theorieën en psychometrie morrelen aan de basis van de geprotocolleerde psychologische diagnostiek en dagen uit duidelijk te zijn over wat een diagnosticus vermag.

Het vierde thema staat weer iets verder van diagnosticeren sec af. Het gaat over de toekomst: Hoe kom je vooruit in de wetenschap waaronder in de psychologie? Het lijkt eenvoudig, want vragen stellen (sectie 1) kan iedereen, net als sterven. Was het maar zo simpel. De minister van Onderwijs wil het weldenkend publiek inschakelen om vragen boven water te krijgen: een vorm van *mental crowdfunding*. De nieuwe voorzitter van de KNAW schreef in NRC van 18-04-2015 dat het stellen van *goede* vragen moeilijk is. Wetenschappers zijn er goed in omdat ze vragen afbreken in deelvragen. In het interview staan geen voorbeelden van goede vragen. Ze wijst ook op het stellen van onderzoekbare vragen. Wat bedoelt ze daar precies mee? Mijn voorbeelden aan studenten waren dooddoeners: 'Is er leven na de dood'? Wat niet onderzoekbaar is, hangt af van epistemologische *partis pris*. Daarna bespreek ik suggesties van vier Nederlandse bètahoogleraren over hoe je baanbrekend onderzoek doet (sectie 2). Ze vatten dat op als het publiceren van artikelen in hoge *impact* Engelstalige tijdschriften. Wat moet de (jonge) wetenschapper ermee? Hoe komt hij vooruit (sectie 3)?

Het Slot bestaat uit *oneliners* over diagnostiek. Terugdenkend aan diagnostiek constateer, beweer en vind ik van alles, dat ik verder niet verdedig. Ze kunnen als een los uit de pols gemaakte samenvatting of conclusie van de thema's gelezen worden. Ik kom tot de slotsom dat een diagnosticus zijn werk *zelf* moet maken met met behulp van enkele voorschriften en protocollen die hem aangereikt en/of opgelegd worden. Diagnostiek moet je doen: *poiein*. Er is geen definitief protocol, geen veilig *'... bestaan in een lied, een eind'lijke vrouw'* (Bloem<sup>50</sup>) dat de diagnosticus leidt, bij de hand neemt, rust geeft. Tegelijkertijd mag hij die hand niet overspelen. Hij wordt immers regelmatig een brokkenmaker genoemd. De voorschriften gaan overwegend over het hóe en terzijde, bijna binnensmonds over het wát. Voor dat laatste kan de diagnosticus aankloppen bij het kennisbestand van de psychologie in zoverre dat relevant is voor de diagnose van de cliënt. De diagnosticus moet tegen beter weten in 'alles weten' om tot een rijke en voldoende beschrijving van het gedrag van de cliënt te komen.

Achter de Inleiding, de thema's en het Slot worden Referenties vermeld.

Telkens als er 'hij' of 'hem' staat, dient 'zij' en 'haar' gelezen te worden.

## Referenties

1. Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 2, 262-274.
2. Duijker, H.J.C., Palland, B.G. & Vuyk, R. *Leerboek der Psychologie* (1960, 2<sup>de</sup> druk) Groningen: Wolters.
- Duijker, H.J.C. (1959). Nomenclatuur en systematiek der psychologie. *Nederlands Tijdschrift voor de Psychologie*, 14, 89-116.
3. Linschoten, J. (1964). *Idolen van de Psycholoog*. Bijleveld: Utrecht.
4. Kouwer, B.J. (1963.) *Het spel van de persoonlijkheid*. Bijleveld: Utrecht.
5. De Groot, A.D. (1961). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen* (12<sup>de</sup> druk, 1994). Den Haag: Mouton.
6. Heidegger, M. (1927; 6<sup>e</sup> druk 1962). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag. Nederlandse vertaling SUN Nijmegen door Marc Wildschut: *Zijn en Tijd*.
7. Merleau-Ponty, M. (1945). *Phénoménologie de la Perception*. Paris Gallimard: vertaald door Colin Smith: *Phenomenology of Perception* (1958). London: Routledge Classics. Nederlandse vertaling beschikbaar: *Fenomenologie van de Waarneming*.
8. Edwards, A.L. *Statistical Methods of the Behavioral Sciences* (1954). New York: Holt, Rinehart & Winston.
- Edwards, A.L. (1957). *Techniques of Attitude Scale Construction*. New York: Appleton-Century-Crofts.
9. Drenth, P.J.D. (1967). *Testtheorie: Inleiding in de theorie van psychologische tests en zijn toepassingen*. Houten: Bohn, Stafleu van Loghum.
10. Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszin, T.W & Reed, G.M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist*, 56, 2, 128-165.
11. Bloom, B.S., Engelhart, M.D., Walker, H.H., Hill, E.J., Furst, E.J. & Krathwohl, D.R. (1971). Vertaling: *Taxonomie van een aantal in onderwijs en vorming gestelde doelen*. Universitaire Pers Rotterdam/ Standaard Wetenschappelijke Uitgeverij Antwerpen.
12. Magnusson, D. (1967). *Test Theory*. Addison-Wesley Publishing Company. Reading Massachusetts.
13. Messick, S. (1989). Validity. In: R.L. Linn (Ed.). (3<sup>rd</sup> Edition). *Educational Measurement*. National Council on Measurement in Education. London: Collier Macmillan Publishers (pp. 13-105).
14. Strasser, S. (1961/1964). *Fenomenologie en empirische menskunde. Bijdrage tot een nieuw wetenschapsideaal*. Arnhem: Van Loghum Slaterus.
15. Barendregt, J.T. (1977). *Karakters van en naar Theophrastes*. Deventer: Van Loghum Slaterus.
16. Vestdijk, S. (1934). *Terug tot Ina Damman*. Amsterdam: Nijgh & van Ditmar.

17. Cook, T.D. & Campbell, D.T. (1976). The Design and Conduct of Quasi-Experiments and true Experiments in Field settings. In: M.D. Dunette (Ed.). *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally College Publishing Cy.
18. Cohen, J. (1994). The earth is round ( $p < .05$ ). *American Psychologist*, 49, 7, 997-1003.
19. Glass, G.V., Willson, V.L & Gottman, J.M. (1975). *Design and analysis of time-series experiments*. Boulder Colorado: Colorado Associated University Press.
20. Lamb, M.E., Suomi, S.J. & Stephenson, G.R. (Eds.) (1979). *Social interaction analysis: Methodological issues*. Madison: The University of Wisconsin.
21. Heatherton, T.F. & Weinberger, G.J.L. (Eds.). (1994). *Can Personality Change?* Washington DC American Psychological Association.
22. Denissen, J.J.A. (2014). Editorial: Roadmap for further progress in research on personality development. *European Journal of Personality*, 28, 213-215. Het nummer gaat over 'persoonsontwikkeling', maar de 'McCrosta' invloed is zichtbaar.
23. Asendorpf, J.B. (2003). *Psychologie der Persönlichkeit*, 3. Auflage. Berlin: Springer. (Nederlandse vertaling beschikbaar).
24. Goswami, U. (2008). *The learning brain*. Psychology Press. New York: Taylor and Francis Group.
25. Adolph, K.E., Young, J.W., Robinson, S.R. & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review*, 115, 3, 527-543.
26. Hermans, H.J.M., ter Laak, J.J.F. & Maes, P.C.J.M. (1972). Achievement motivation and fear of failure in family and school. *Developmental Psychology*, 6, 520-528.
27. Barelds, D.P.H. & Luteijn, F. (2002). Measuring Personality: A comparison of three personality questionnaires in the Netherlands. *Personality and Individual Differences*, 33, 499-510.
28. Van Kampen, D. (2015). Een kritische beschouwing: NEO-dimensies en de 5DPT. *De Psycholoog*, juni 2015, 49-58.
29. Hacking, I. (1975). *The emergency of probability*. Cambridge: Cambridge University Press.
30. Hacking, I. (1990). *The taming of chance*. Cambridge: Cambridge University Press.
31. James, W. (1983; 1890, first Edition). *The Principles of Psychology*. Vol. I and Vol. II. New York Cosimo Classics/ Harvard University Press. Linschoten heeft dit werk bekend gemaakt in Nederland.
32. James, W. (1902, 1982). *Varieties of religious experience*. Hammondswoth. In 1982 printed as Penguin Book. Nederlandse vertaling beschikbaar (2010).
33. Strasser, S. (1974). De dialogische dimensie der waarheid. *Tijdschrift voor Filosofie*, 36, 399-417.
34. Barendregt, J. (1965). 'Ingezonden' en P.C. Kuiper, 'Ingezonden'. *Nederlands Tijdschrift voor Geneeskunde*, 109, 1736.
35. Smaling, Adri (1987). *Methodologische objectiviteit en kwalitatief onderzoek*. Lisse: Swets & Zeitlinger.
36. Witteman, C., Van der Heijden, P & Claes, L. (2014). *Psychodiagnostiek: het onderzoeksproces in de praktijk*. Zoetermeer: De Tijdstroom.

36. Laak, J.J.F. ter (2011). *Elementair begrip van de psychologische diagnostiek*. Amsterdam: Pearson.
37. Bunge, M. (1996). *Finding Philosophy in Social Sciences*. New Haven: Yale University Press.
38. Bermúdez, J.L. (2005). *Philosophy of Psychology: A Contemporary Introduction*. New York: Routledge, Taylor and Francis Group.
39. Simon, H.A. (1991). *Bounded rationality and organizational learning*. Online beschikbaar: Permalink: <http://dx.doi.org/10.1287/orsc.2.1.125>.
40. Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: The Danish Institute for Educational Research.
41. Nijhoff, M. *Verzamelde Gedichten* (1963, 2e druk). Den Haag: Bert Bakker / Daamen N.V. (De regel is uit het gedicht 'Levensloop' uit de bundel Vormen (Bussum: Van Dishoeck, 1924, p. 105).
42. Wittgenstein, L. (1953; translated 1999). *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford, UK: Blackwell Publishing.
43. Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 3, 281-302.
44. Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multi-trait-multi-method matrix. *Psychological Bulletin*, 56, 1, 81-105.
45. Cronbach, L.J. (1990). *Essentials of Psychological Testing*. (6<sup>th</sup> edition). New York: Harper & Row.
46. Newell, B.R. & Shanks, D.R. (2014). Unconscious influences on decision making: a critical review. *Behavioral and Brain Sciences*, 37, 1-19.
47. Dijksterhuis, A., Van Knippenberg, A. & Veling, H. (2014). Newell and Shanks' approach to psychology is a dead end. *Behavioral and Brain Sciences*, 37, p. 25.
48. Dautzenberg, A.H.J. (2013). *Rafelranden van de moraal, Nouvelle*. Antwerpen Amsterdam: Atlas Contact.
49. *De Bijbel* (auteurscollectief). Nederlandse vertaling 2004: *De Psalmen* (pp. 835-1023). Amsterdam/ Antwerpen: Querido Jongbloed.
50. Bloem, J.C. (1981). *Verzamelde Gedichten*. Amsterdam: Athenaeum-Polak & Van Genneep.

## II Wat is psychologische diagnostiek?

*Handboeken beginnen met een werkdefinitie van diagnostiek en gaan over tot de orde van de dag. Vermijd de vraag 'wat is...' leerde ik bij filosofie. En, bij psychologie sloeg je die vraag over. Bij filosofie koketteerde men er echter ook graag mee. Het betekende dat je - hoe moedig en origineel - opnieuw de filosofie moest uitvinden, bijvoorbeeld Heidegger<sup>1</sup> met zijn 'Wat is dat-Filosofie' en de neo-Thomist Pieper<sup>2</sup> met zijn boek over filosofie als liefde voor wijsheid, geboren uit verwondering. Waarom zou filosofie niet uit kwaadheid, gekwetstheid, waarneming van een gevoel voor het gebrekkige, het onvolmaakte kunnen ontstaan? Heidegger zegt er bij dat het onderwerp onbepaald en mateloos is en ons allen raakt. Word daar maar eens wijs uit. Diagnostiek is ook onbepaald, daarover later meer. Pieper is 'sanft' en maakt filosofie tot het bezit van iedereen, net iets anders als dat het ons allen raakt en soms schuurt. Ze begint met verwondering over wat is en dankbaarheid omdat alles gegeven is. Ze eindigt in liefde voor wijsheid. En dit met WO I en II net achter de rug. Dat kreeg je voor je kiezen als jonge student in de jaren 60-70. Pragmatici als James<sup>3</sup> en Rorty<sup>4</sup> vermeden dit soort vragen. Ik neem het risico en ga op die vraag in. Schrijvers en lezers formuleren zo hun eigen omschrijvingen. Ieder heeft zijn unieke voor een deel ongerefecteerde referentiekader en kennisbestand. Dat valt nooit samen met dat van een ander. Niet voor niets dat er zoveel ondernomen moet worden om het zelfs over ogenschijnlijk eenvoudige zaken eens te worden.*

## 1. Het begin: Psychologie en de vraag, het probleem van de cliënt

Diagnostiek bevindt zich in een uithoek, het vooronder van de psychologieopleiding. Ze heeft een bijrol in de praktijk als opmaat voor de behandeling. In de universitaire vakgroepen is zij een cursus omdat de praktijk erom vraagt. Ze meestal is ondergebracht in een dienstverlenend Ambulatorium. Ze is geen onderzoeksthema. In de jaren 60 tot 80 bevatte een aanzienlijk aantal proefschriften de ontwikkeling en validering van tests. Dat is voorbij. De afdeling Methoden & Technieken vertelt over psychometrie, niet over testontwikkeling. Andere, op de praktijk gerichte afdelingen leren tests afnemen en een diagnostische procedure volgen. Om de activiteit onder de hoede van wetenschap te brengen is het proces gedisciplineerd en geprotocolleerd.

Diagnostiek wordt hier opgevat als onderzoek dat gaat over het verkrijgen van valide kennis over het gedrag van een cliënt. Diagnostiek lijkt echter meer op gezoek dan op onderzoek in de empirisch analytische betekenis van het woord. Zij toetst op  $n = 1$  niveau of het kennisbestand van de psychologie iets oplevert voor het analyseren van het gedrag of het probleem van een cliënt. Daarmee wordt een actie bedoeld die leidt tot een rijke (*rich*) en voldoende (*sufficient*) beschrijving (Simon<sup>5</sup>) van zijn gedrag. Dat gedrag is een vraag of een probleem voor de cliënt en/of voor zijn omgeving. Het is niet duidelijk wanneer je inhoudelijk diep genoeg gegaan bent als diagnosticus. Je weet ook niet wanneer je moet ophouden. Toch moet je naar diepte en volledigheid streven en de protocollen kritisch en onafhankelijk benutten.

Toen Duijker<sup>6</sup> in 1959 de psychologie verdeelde over vakgroepen werd diagnostiek niet genoemd. Hij had al moeite met Sociale Psychologie omdat 'sociaal' geen onderscheidend gedragskenmerk is. Persoons-, functie- en ontwikkelingsleer zijn disciplines want in het bezit van een materieel en/of formeel object. Diagnostiek heeft geen formeel object, dat wil zeggen geen eigen gezichtspunt om gedragingen te bestuderen. Ze heeft ook geen materieel object want ze richt zich niet op een specifiek gedragsdomein, maar op welk gedrag dan ook. Er zijn wel verlegenheidsoplossingen door diagnostiek aan een domein te verbinden. Sommige diagnostici richten zich bijvoorbeeld op disfuncties, of op een populatie, bijvoorbeeld kinderen of bejaarden, zoals in de pedagogiek, ontwikkelingspsychologie en klinische psychologie.

Pedagogen, ontwikkelingspsychologen en klinisch psychologen vragen naar de meerwaarde van diagnostiek voor de behandeling. Sommige klinisch psychologen slaan diagnostiek over. Het komt ook voor dat diagnostiek al als behandeling wordt opgevat. Verder verschuift het onderscheid tussen diagnosticus en cliënt van iemand die de vraag of het probleem analyseert naar twee deelnemers die overleggen in een gesprek of een dialoog voeren. Diagnostiek is dan geen toets meer voor wat de psychologie te bieden heeft om vragen over het gedrag van cliënten te beantwoorden. Ze gaat op in de behandeling. Het onderwijs wordt bij voorkeur verzorgd door docenten met ervaring in de praktijk.

Psychologische diagnostiek is in mijn ogen het analyseren van een vraag of een probleem van een cliënt over zijn gedrag. Het kennisbestand van de psychologie biedt daarvoor de theorieën en werkwijzen, het instrumentarium en empirische bevindingen.

*Iedereen diagnosticeert:* Diagnosticeren mag dan marginaal zijn in de opleiding, in het dagelijks leven is hij dat niet. Duidker merkte op dat iedereen aan psychologie doet en zijn *statement* is in 2015 tot een rubriek in de *Psycholoog* geworden. Waarom diagnosticeren we zo vanzelfsprekend? We willen weten wat voor vlees we in de kuip hebben. We selecteren omdat we ervan overtuigd zijn dat we het Toeval kunnen verslaan in de keuze van partners, vrienden, werknemers en ons beroep. We evalueren meteen of een persoon een voor- of nadeel voor ons is of kan worden. We verdelen mensen graag in: wij-zij. Deze tendensen horen bij onze Darwinistische overlevingsstijl.

*Op welk gedrag vallen we?* We zijn geïnteresseerd in gedragingen en sociale situaties die we waarderen: succes op school en in het beroep, kwaliteit van scholen, harmonieuze gezinnen en schone omgevingen; waar we bang voor zijn: agressie, terrorisme, milieuvervuiling en psychopathologie en die ons medelijden opwekken: gehandicapte, misbruikte kinderen, het lijden van de minder bevoorrechten, leerproblemen en pathologische gedragingen. Bij deze drie gaat het om de extremen: de meest (minst) succesvolle cliënten, de agressiefste personen en zij die extreem lijden of disfunctioneren.

*Wie diagnosticeren we?* De cliënt. Dat kan een individu zijn, een team, een schoolklas, een bedrijf of onderneming en zelfs een cultuur of samenleving. De nadruk ligt bij mij op de individuele persoon. De cliënt kan ook een wettelijke vertegenwoordiger van die persoon zijn, bijvoorbeeld van baby's, kinderen of van mensen met een beperking. Bovendien kan de overheid, bijvoorbeeld de rechtbank om een diagnostisch oordeel vragen. Ik ga er van uit dat er steeds iemand is die zelf of namens de onderzochte persoon 'terug kan praten'. Je diagnosticeert geen object.

*Waar halen we de kennis vandaan?* Waar valt diagnostiek op terug zonder formeel en materieel object? Zij haalt *ad hoc* overal iets vandaan en haar beoefenaars worden door schade en schande wijs. Na ongelukken komen de protocollen, denk aan de anatomische poppen voor het vaststellen van seksueel misbruik. Het ligt in de rede om terug te vallen op het kennisbestand van de psychologie. Dat is een *mer à boire*. 'Alles in Allen' zei amateurfilosoof Multatuli al. Zijn *Ideeën*<sup>7</sup> zijn een (aangename) grabbelton. Hoe dan een rijk geschakeerd en voldoende beeld te schetsen van de vraag/probleem van de cliënt? Elk voorstel bevat een bepaling van het kennisbestand van de psychologie. Die bepaling is in voor mij uitgekomen op drie elementen:

(1) Theorieën: Dat zijn formele modellen, zoals de Item Respons Theorie (IRT) en/of semantische netwerken of structuren van gedragsbegrippen. Een model is een wiskundige, logische constructie die theorie wordt als er een werkelijkheid mee beschreven is. Bij de IRT worden de functies theorie, zodra het antwoord op een item er mee beschreven wordt. IRT is een minitheorie over minigedrag. 'Mini' is niet denigrerend bedoeld; bij atomistische denkers is het zelfs een voordeel. Piaget als structuralist is een voorbeeld: om de concrete en formele operaties te beschrijven gebruikte hij formele structuren die hij inhoud probeerde te geven

in taken en experimentjes. In een groot deel van psychologie en filosofie is het construct, het begrip de eenheid.

(2) Operationalisaties en representaties van constructen die gedrag en sociale situaties concreet maken met behulp van voorbeelden. De items materialiseren of representeren gedragingen als functies in een (meet) model. Meten is het afbeelden van een empirisch waargenomen geheel van relaties in een formeel systeem van relaties. Dat systeem is meestal een categorie, schaal of aantal schalen. Operationalisatie is het werkendeweg concretiseren, materialiseren, empirisch realiseren van een (hypothetisch) begrip.

(3) Instrumenten, observaties, tests, vragenlijsten en elke manier om gedrag vast te stellen. De elementen van diagnostiek en daarbij horende vaklieden zijn:

Theorievorming door persoonlijkheids-, sociaal, ontwikkelings- en functieleer psychologen	Operationalisatie en meten door onderzoekers, itemschrijvers, psychometrici, statistici en modelbouwers	Instrumentatie in de vorm van procedures, tests en vragenlijsten ontwikkeld door testontwikkelaars en procedureontwerpers
---	---	---

Drie combinaties tussen elementen komen het meest voor: (1) Het *top-down* pad van theorie → modellering → instrumentatie lijkt op de hypothetisch-deductieve werkwijze uit de empirische cyclus van De Groot. Vooral onderzoekers zien dit als de koninklijke weg. (2) Het centrifugale pad sluit aan bij het maken van een instrument, bijvoorbeeld een vragenlijst om de mate van stress bij beveiligingsmedewerkers te meten. Gedrag dat bij het construct 'stress' in dit beroep hoort, wordt omschreven. Uit de literatuur wordt informatie gehaald over inhoud en effect van stress op werkprestaties. Ten slotte worden één of meer schalen gedefinieerd, die individuele verschillen in ervaren stress bij deze groep vaststellen. Instrumentontwikkelaars volgen dit pad. (3) Het *bottom-up* pad begint het afbeelden van een empirisch systeem in een formeel systeem, bijvoorbeeld een systeem van ziekte categorieën, een itemresponsfunctie. Psychometrici geven de voorkeur aan deze fundamentele weg.

Psychologische theorieën zijn *mers à boire*. Poppers<sup>8</sup> gedachte van het cruciale experiment werkt niet in de psychologie. Theorieën verdwijnen niet na een toets. Die overmaat roept weerstand op. Daar komt verlies van vertrouwen in grote systemen, theorieën, ideologieën en religies nog bij. Vergeet ook niet de nadruk op 'de praktijk. 'Veel te theoretisch wat u vertelt...', '...Ik kan er niets mee in de praktijk, 'Wat is uw meerwaarde'? Toch nemen we moeiteloos gehelen (*Gestalten*) en relaties waar. We beschrijven, benoemen, voorspellen en verklaren fysische en gedragsverschijnselen zonder veel nadenken, bijna achteloos. Methodologen letten daarbij op fouten: *biases* (vertekeningen) en *heuristics* (zelf uitgevonden, niet getoetste regels: de idolen van Bacon/Linschoten). Meehl<sup>9</sup> beweerde in 1978 dat theorievorming in de diagnostiek *soft* is, wetenschappelijk weinig voorstelt en nauwelijks aanknopingspunten voor handelen biedt. Deze sceptische houding wordt versterkt door het postmodernisme. Denk aan het echech van systemen als het strikte behaviorisme en de orthodoxe psychoanalyse. De houding is: theore-

tiseer niet, kijk naar de feiten: de mier aanpak van Bacon<sup>10</sup> in zijn *Novum Organon*. Toch kunnen we ons niet onttrekken aan theorieën al bevatten ze geen Cartesiaanse klare en distincte ideeën. Ze horen bij ons denken en handelen. Ze komen voort uit verwondering of onbehagen, behoefte aan kennis, begrip, inzicht en het streven het Lot te ontvluchten of te beheersen. Dit is wat Heidegger<sup>11</sup> *Seinsglaube* noemt. Het omvat uiteenlopende opvattingen, beliefs, theorieën, inzichten en ideologieën. Ze struikelen over elkaar en bestrijden elkaar.

In het dagelijks leven en als *eerste stap* in de wetenschapsbeoefening trekken we ons niets van voorschriften aan. We gaan moeiteloos verder dan wat we zien en ervaren. We zien de achterkant van een tafel niet maar weten dat die er is. We zien en ervaren onrechtvaardigheid, onbeschoftheid, maar weten dat er (zuivere) rechtvaardigheid en eerbied bestaat. Dit kenmerk van ons kennen staat al in Kants<sup>12</sup> *Kritiek der Zuivere Rede* (1878/1961). In vertaling: *Er kan geen twijfel over bestaan dat al onze kennis begint met ervaring. Want hoe zou het kenvermogen tot activiteit kunnen worden aangezet als dat niet gebeurde doordat objecten onze zintuigen in beweging brengen? Maar hoewel al onze kennis met de ervaring begint, betekent dat nog niet dat ze daarom uit de ervaring voortkomt.* Zo komt Kant op de apriori's van tijd, ruimte en causaliteit.

De intuïties van filosofen, van Plato's *Ideeën*, Schopenhauers *Wil* tot Merleau-Ponty's *Corps-sujet* zijn voorbeelden van dat getheoretiseer. Van Dale noemt het een woord met negatieve gevoelswaarde. Dat moet dan maar. Psychologen, van Wundt, Watson, Piaget tot Eysenck deden niet anders. Ons kenvermogen brengt iets uit zichzelf voort, enkel naar aanleiding van zintuigelijke indrukken. Die ervaring is de aanleiding, de gelegenheid om theorie te laten ontstaan, of is gelijktijdig met theorie, want ervaring is niet theorieeloos (Descartes' in- of aangeboren ideeën, Kants apriori's). Ze lijkt eerst direct en naïef gekoppeld aan wat waargenomen wordt maar er komt geleidelijk een louteringsproces op gang waar het zintuigelijk-stoffelijke afgestoten wordt en met zichzelf in tegenstelling wordt gebracht. En dat houdt nooit op. In De Groot's *Methodologie* staat het zo: *Een theorie is een systeem van logische samenhangende, niet-strijdige beweringen, opvattingen en begrippen betreffende een werkelijkheidsgebied die zo geformuleerd zijn, dat het mogelijk is er toetsbare hypothesen uit af te leiden. Op grond van houdbare voorspellingen komt men tot theorievorming.*

Aan psychologisch getheoretiseer is niet te ontkomen. Het bevat steeds interpretatie en waardering. De vorm - formeel of semantisch - heeft een aardse lezing nodig. Die kan er ook aan voorafgaan en vervolgens in een formeel systeem gegoten worden. De Griekse slaaf Meno die in het zand een meetkundige figuur tekende: de afbeelding als hulp en uitwerking van een Idee is daar een voorbeeld van, ofschoon Plato<sup>13</sup> het hem verbood. De tekening evenals een gedicht verduisterde immers het Idee. Het is ook nooit goed. Het mag dan vertroebelen maar we kunnen niet zonder getheoretiseer en geformaliseer.

Waardering blijkt uit Plato's intuïtie waartoe theorie dient: waarde creëren: het goede leven: *eudamonia*. Bij Heidegger moeten we waarde creëren door sluiers van verhullende betekenissen weg te halen en authentiek te kennen en leven. Ga er maar aan staan. Het moet nu nog en het heet valorisatie: is dat materiële opbrengst, rendement, of nog iets anders? Leg het mij uit, managers.

Interpretatie en waardering kleven aan iedere theorie. Ze vormen een liedje van verlangen. Het liedje is een formeel logische of semantische structuur, compositie, symfonie, tekst, een romantisch of tragisch verhaal. Het liedje heeft een relatie met een inhoud. Het verlangen drukt waarde uit: het tot stand brengen van inzicht, het te weeg brengen van een waardevolle ervaring. Methodologen kunnen het liedje van verlangen niet stoppen. Ze kunnen er op letten dat we niet ontsporen. In empirisch onderzoek komen we ons zelf overigens bijna altijd tegen: het is regelmatig een vals klinkend, niet zuiver gezongen liedje.

#### Drie oriëntaties

<p>I: Individuele verschillen: het vergelijken van gedrag van personen door categorieën, schalen te ontwerpen, bijvoorbeeld persoonlijkheidstypen, intelligentieschalen; schalen om sociale omgevingen en instituties in kaart te brengen</p>	<p>II: Ontwikkeling: het beschrijven en verklaren van huidig en toekomstig gedrag door te vragen hoe dat tot stand is gekomen; stadia van cognitieve en persoonsontwikkeling, ontwikkeling van sociale contexten, bijvoorbeeld het gezin, de buurt, een organisatie/ instituut</p>	<p>III: Sociale en fysische context: verklaren van gedrag als gevolg van de invloed van fysische en sociale omgeving, bijvoorbeeld reinforcement theorie, leren van modellen, sociaal behaviorisme</p>
---	--	--

Elk gedragstheorie bevat drie niet-willekeurige gezichtspunten: (1) Mensen verschillen in gedrag. Gelijkheid voor de wet is een democratische constructie, een regulatief principe, een contrafactisch streven. Als we gedrag categoriseren, voorspellen of verklaren, vergelijken we ons met anderen op fysische en psychologische kenmerken. We leren onszelf kennen door ons te vergelijken met anderen. We merken een verschijnsel, gebeurtenis, gedrag op als er verschillen zijn; denk aan de figuur-achtergrond waarneming. (2) Als we gedrag willen categoriseren, voorspellen, controleren, begrijpen, vragen we ons af hoe het zo gekomen is. Mensen construeren een verleden, hebben een heden en projecteren een toekomst. Dit berust op het feit dat we tijdelijk zijn en een geheugen hebben. Darwin heeft aannemelijk gemaakt dat het ontstaan van soorten - de natuurlijke historie - een tijdrovend toevalsproces is. Niettemin lezen we evolutie gemakkelijk als vooruitgang, bijvoorbeeld in het geval van cognitieve en persoonlijkheidsontwikkeling. Denk aan Piaget: van reflexen naar formeel operator denken en Freud of Erikson van symbiose naar de rijpe, onafhankelijke persoon. Deze vooruitgang wordt ons in de schoot geworpen. We nemen ook graag zelf het heft in de hand door middel van trainingen en interventies. We verwachten er veel van, soms tegen beter weten in, denk aan evaluatiestudies en stagnaties in kennen, denken en persoonsontwikkeling. Dawkins<sup>14</sup> benadrukt het onverschillige toeval: *Darwinistic nature is pitiless indifferent, neither good nor evil, neither cruel nor kind, but simply callous, indifferent to all suffering, lacking all purpose.* (3) Bij het beschrijven en verklaren van gedrag wijzen we de fysische omgeving en de sociale situatie aan als de daders, doeners, factoren. Zij hebben het gedaan! Elk gedrag is gesitueerd. Context verwees eerst naar specifieke fysische stimuli en is later uitgebreid tot sociale stimuli, inclusief

personen als modellen en sociale situaties, zoals gezin, buurt, school en cultuur. Dat zijn natuurlijke omgevingen, geen gefabriceerde zoals in experimenten, trainingen en interventies.

*Drie bronnen voor theorie*

<p>I: Impliciete, alledaagse theorie door leken over oorzaken en aard van individuele verschillen, ontwikkeling en sociale context; impliciete methodologie, bijvoorbeeld klinische predictie, Brunswiks lensmodel</p>	<p>II: Expliciete theorievorming: theorieën (behaviorisme, cognitieve, sociale ontwikkeling, persoonlijkheidstheorie) en methodologie: onderzoeksontwerpen, psychometrie, testleer, statistiek en instrumenten uit de tekstboeken</p>	<p>III: Alternatieve theorie: voorstellen van auteurs die witte plekken waarnemen in de expliciete theorievorming en methoden</p>
--	---	---

De kern van impliciete theorie is dat de definitie van termen, begrippen, constructen en methoden aan leken wordt overgelaten. De *Big Five* is een voorbeeld. Kennis van impliciete theorie is zinvol want de cliënt heeft zijn eigen theorie over zijn vraag of probleem. Impliciete theorieën en methoden zijn reactief en houden beperkt rekening met bestaande informatie over een probleem. Er is voorkeur voor concrete, direct relevante onderwerpen en voor het oplossen van praktische problemen. We benutten in het dagelijks leven geen onderzoeksmethoden, verrichten geen filosofische analyses en trekken geen lessen uit de geschiedenis. Het verzet van onze studenten tegen 'al die theorie' heeft voor een deel te maken met vasthouden aan hun impliciete theorieën. En dat kan ook in de psychologie. In andere studies is de *zelfbetrokkenheid* geringer. Bij de bètavakken moet je leren, leren en nog eens leren. Bij de alfavakken, bijvoorbeeld een taal, moet je die goed leren spreken en dat is oefenen en kost tijd en bij geschiedenis moet je oude teksten kunnen lezen en interpreteren. Als je aan de psychologiestudie begint, heb je al een theorie: je impliciete theorie en die voldoet goed. Onze studenten weten van wensen in het dagelijks leven. Ze hebben het ver geschopt. Ze zitten op de universiteit.

Expliciete theorieën en methoden zijn de paradigma's van Kuhn<sup>15</sup>. Ze houden het een tijdje vol en worden vervangen of opgeruimd. Dat is het verhaal van Kuhn. In de psychologie raken we oude theorie niet kwijt en er komt geen nieuwe, almachtige en alwetende monotheorie aan. Een enkele wetenschapper gaat ver: Skinner wilde bijvoorbeeld met zijn behaviorisme het menselijk gedrag en de hele wereld verbeteren. Watson en Crick (van het DNA) kondigden de mogelijkheid van volledige controle van menselijke eigenschappen aan. De doorsnee psycholoog kijkt daar besmuikt naar. Monotheorisme kan evenals monomethodologisme tot repressie leiden. Ze lijken op de drie monotheïstische leringen die waarheidsopvattingen opleggen. Dit sluit een monistische benadering zoals in de evolutietheorie niet uit. Die moet soepel zijn en eindeloos variaties toestaan.

Alternatieve theorieën en methoden zijn een reactie op de expliciete. Een met zoveel aplomb gebrachte en dominante theorie lokt tegenspraak uit. Er wordt iets afgewezen en toegevoegd, omdat interessante en relevante inhoud en/of methoden ontbreken. Als je iets zegt, zeg je iets anders *niet*. Elke bewering onthult en verhuult.

*Relatie tussen de bronnen:* De drie sluiten elkaar niet uit en zijn niet hiërarchisch te ordenen. De spanning tussen de drie is de motor voor verandering. Ze bewaken elkaar en gaan een kritische dialoog aan die niet tot overeenstemming hoeft te leiden. De drie oriëntaties in één zin: het is een oer-evidentie (Strasser) dat we ons hier (context) en nu (een even gestolde ontwikkeling) bevinden samen met anderen (van wie we verschillen en door wie we onszelf als een figuur tegen een achtergrond ontdekken).

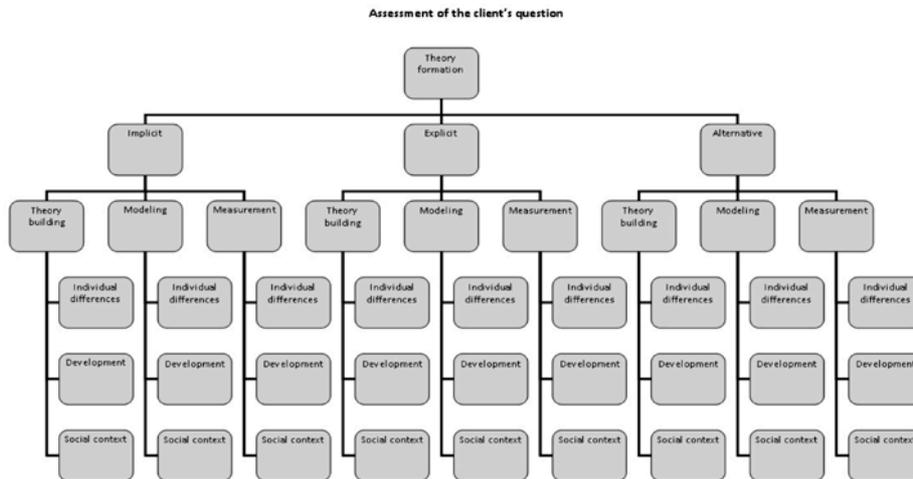
Elk gedrag kan geanalyseerd worden als individueel verschil, ontwikkelingsverschijnsel en als functie van de context. De bronnen kunnen niet tot strijdige feiten over hetzelfde gedrag leiden. Ze kunnen niet tot elkaar herleid worden. Reese en Overton<sup>16</sup> spraken in de jaren 70 over *incompatible world views*. Een *world view* is nodig om althans iets te zien en te begrijpen. De drie oriëntaties zijn autonoom en gelijkwaardig. Rawls<sup>17</sup> zou ze *equal in power, richness, and ability* noemen. Bij de drie bronnen kun je dat niet zonder meer en altijd zeggen: sommige alternatieven kunnen ver gaan en tot een schijnvertoning leiden.

Het tweede element van diagnostiek - operationalisatie en meten - verbindt theorie aan vragen, items en taken. Deze bevatten concreet gedrag, bijvoorbeeld het oplossen van een probleem als een index van een construct. Het kan weergegeven worden als een aantal categorieën, een structuur, bijvoorbeeld een ééndimensionele schaal of een latente trek. Testontwikkelaars en itemschrijvers komen hier in beeld. Psychometrici passen hun modellen op de bedachte taken en vragen. Ze passen deze niet zonder meer aan bij een geobserveerd systeem van relaties. De passing is bij begrippen als zwaartekracht en energie vanzelfsprekend maar leidt in de psychologie tot gedoe. De helft van de items wordt weggegooid omdat ze niet aan het gepostuleerde model voldoen. Om de itembakers (een benaming gebruikt op het Cito) te ondersteunen dit citaat uit James's<sup>18</sup> werk over de religieuze ervaring (Nederlandse vertaling, 2010, p. XIII '*...dat een brede kennis van voorbeelden meer wijsheid brengt dan het bezit van abstracte formules, hoe diep ook...*').

Operationalisatie komt op *meten* uit als het empirische systeem afgebeeld wordt in een formeel systeem. Doorgaans leidt dat tot (a) categorieën, bijvoorbeeld DSM-IV-TR, DSM-5, IDC-10, IDC-11; ontwikkelingsstadia; persoonstypen; onderwerpen in het diagnostisch interview en observatiecategorieën (b) schalen, bijvoorbeeld attitude-, intelligentie- of persoonlijkheidsschalen en (c) aantallen schalen, bijvoorbeeld meervoudige intelligentie- en persoonlijkheidsschalen. De afstand tussen (a) en (b) wordt kleiner. De DSM-IV TR is een categoriesysteem, maar Watson<sup>19</sup> (p. 522) beweerde al dat er genoeg bekend is om dit *rational system* van semantische, logische categorieën te vervangen door *...empirically based structure that reflects the actual similarities among disorders*. De nadruk op categorieën is een relict van de oriëntatie op ziektes. Die heb je of heb je niet.

Meetinstrumenten zijn het derde element van diagnostiek. Tests en vragenlijsten hebben de overhand ofschoon alle procedures die gedrag vastleggen tot het gereedschap van de diagnosticus horen. Instrumenten zijn losjes verbonden met theorie en meten. Dit leidt tot de kritiek dat een test een construct niet inhoudelijk en niet qua structuur dekt: U telt maar wat itemscores op en denkt dat ze over hetzelfde gaan en meer beter is, hoger is.

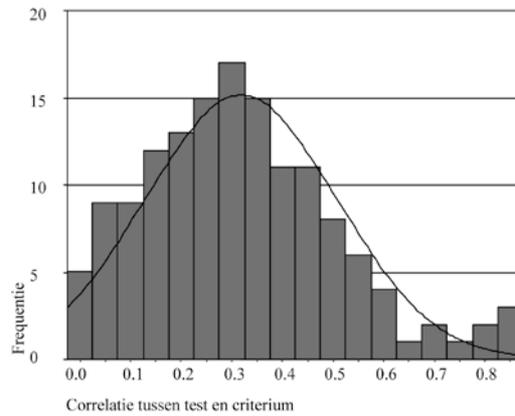
Hieronder een schema (Figuur 1) voor het kennisbestand van de psychologie, opgespannen door de drie elementen en drie oriëntaties. Modellen maken of modelleren verwijst naar afbeelden in een formeel systeem ook naar operationalisatie van een construct en meten.



Figuur 1: Schema om het kennisbestand van de psychologie ten behoeve van diagnostiek te beschrijven. Operationaliseren en meten zijn afzonderlijke activiteiten; het had er netter neergezet kunnen worden door het niet op een hoop te gooien onder 'modeling'.

Het kennisbestand van de psychologie is overwegend opgebouwd uit theoriegeleide, getoetste hypothesen over oorzaken en correlaties van gedragingen van steekproeven.

Diagnostiek is  $n = 1$  studie. De diagnosticus moet nagaan of de cliënt een element is van de steekproef waar hij iets over weet. De d- en r-waarden uit empirische (meta)studies zijn bescheiden. Een voorbeeld dat berust op 125 meta-studies over het *voorspellen* van gedragscriteria en ziekten is van Meyer et al.<sup>20</sup> (Figuur 2). De gemiddelde correlatie bedraagt .31; SD 0,19. De limiet van de verklaarde variantie lijkt ongeveer 50% in steekproeven.



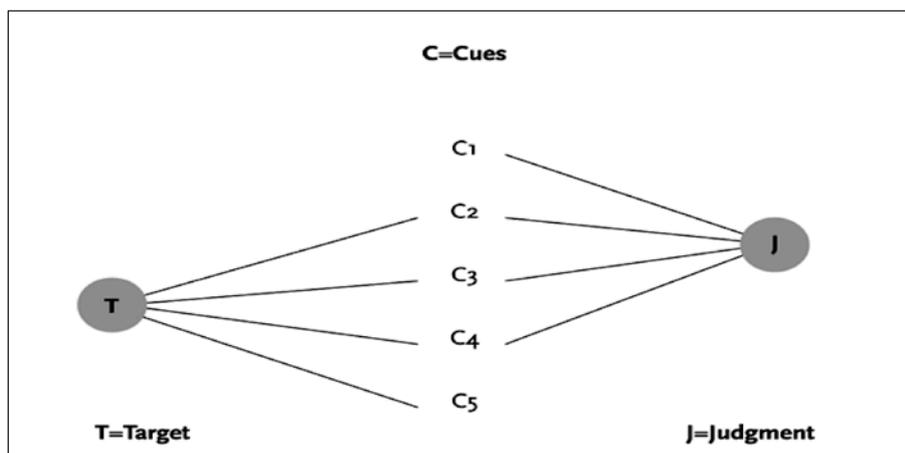
Figuur 2: De predictieve validiteitscoëfficiënten uit 125 meta-studies bijeengezocht door Meyer et al. (2001) en hier in een grafiek gezet.

De limiet lijkt bereikt. *Try harder* is een loze kreet. De twee *mers à boire* van r- en d-waarden helpen bij een rijke en voldoende analyse van de vraag van de cliënt. De diagnosticus moet datgene kiezen uit theorie, operationalisatie en instrumentatie wat past bij de vraag van de cliënt. Kiezen uit zo veel is overvragen en daarom zijn er protocollen. Hierbinnen opereert de diagnosticus. Het is een zee waar hij zijn hoofd boven water probeert te houden. De achter-eenvolgende acties die hij daartoe onderneemt, vormen het diagnostisch proces.

## 2. Het diagnostisch proces

Met informatie uit 'drie zeeën om te drinken' moet de diagnosticus aan de slag. Die activiteit krijgt vorm in het diagnostisch proces. Dat is een doelgerichte activiteit die uit een paar logische stappen bestaat. Het begint bij het probleem, de vraag en eindigt met een oplossing, een antwoord: de redenen of oorzaken van het (on)gewenste gedrag. We beperken diagnostiek tot beschrijving van het probleem. De verbinding tussen diagnose en behandeling is niet één op één. Na het verhaal van de cliënt verzamelt de diagnosticus informatie. Het antwoord is een beschrijving, voorspelling, verklaring, beslissing, advies. Bijvoorbeeld, kan ik binnen een redelijke tijd dokter worden, hoe kom ik van mijn dwanggedachten af, heeft het effect op mijn prestatie als ik deze training doe, wat moet ik doen: bij dit bedrijf blijven of van functie veranderen? Om vraag te beantwoorden, een probleem op te lossen is er een overmaat aan antwoorden en behandelingen.

*Het Lensmodel:* Om het verloop van het diagnostisch proces te beschrijven is het lensmodel van Brunswik<sup>21</sup> bruikbaar. Het is hieronder eenvoudig weergegeven (Figuur 3). Het vertelt hoe iemand informatie verzamelt om tot een geldige waarneming te komen:

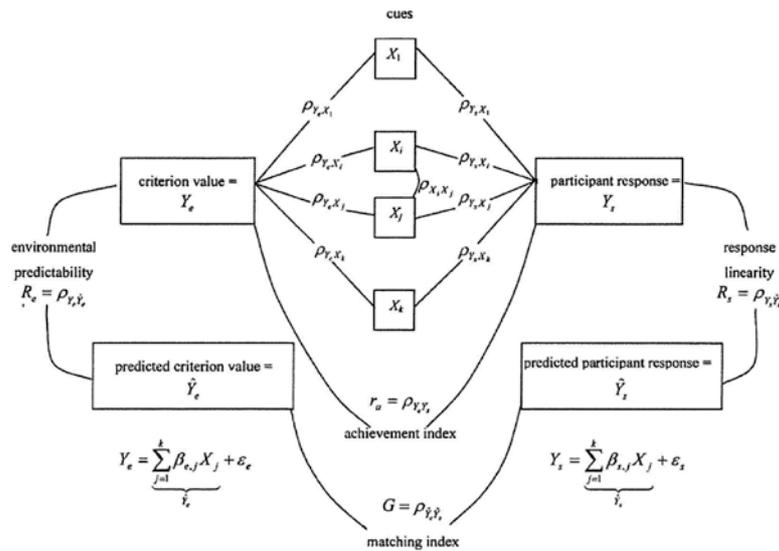


Figuur 3: Een eenvoudig schema om het lensmodel van Brunswik weer te geven

Er is uiteraard een ingewikkelder plaatje gemaakt dan dat van Figuur 3. Dat doet niets af aan het feit dat het gaat om enkele *cues* die we optellen om tot een *geldige waarneming*, diagnose te komen. Dat ingewikkelde bestaat niet uit het feit dat diagnostici gegevens, cues configuratief ordenen, al zeggen ze zelf van wel. Meehl was ook therapeut en vond dit aannemelijk. Je kunt interactieve en zelfs kwadratische combinaties bedenken. De lineaire regel zonder interactie kent alleen hoofdeffecten in termen van de Anova.  $Y$  moet voorspeld worden met  $X_1 \dots X_n$  als predictoren en  $b$ 's slaan op de gewichten:

De lineaire regel (zonder interactie):  $\hat{Y} = b_1X_1 + b_2X_2 + \dots + b_nX_n$   
 De lineaire + 'interactie' regel:  $\hat{Y} = b_1X_1 + b_2X_2 + b_3rX_1X_2$   
 De kwadratische regel (geen interactie):  $\hat{Y} = b_1X_1^2 + b_2X_2^2 + \dots + b_nX_n^2$

Er komt steeds een product uit: een klasstoewijzing, voorspelling van een waarde op een criterium. Het gaat over optelling van enkele cues, zo tussen de 3 en 7, schat ik. Het ingewikkelde model is ook lineair. Er zullen vast niet-lineaire modellen bestaan die met niet-Euclidische afstanden werken. Ik ben ze niet tegengekomen. Misschien heb ik niet genoeg of goed gezocht. Niettemin denk ik dat het product - de voorspelling - er niet substantieel beter door wordt.



Figuur 4: Een ingewikkeld lensmodel. Uit: Determinants of linear judgment door Karalaia, N. & Hogarth, R.M. (2008). *Psychological Bulletin*, 134, 3, 404-426 (p. 405). Het oordeel  $Y_s$  is gemodelleerd als een lineaire functie van een reeks cues  $X_j$ ,  $j = 1 \dots k$ :  $Y_s = \sum \beta_{s,j} X_j + \epsilon_s$ , waarbij  $\beta_{s,j}$  de gewichten weergeven, die de diagnosticus aan de cues geeft en  $\epsilon$  de foutenterm van  $Y$  is op de  $X_j$ . Het omgevingscriterium wordt gemodelleerd door  $s$  door  $e$  te vervangen:  $Y_e = \sum \beta_{e,j} X_j + \epsilon_e$ .

*Disciplineren met behulp van voorschriften:* Er is meer energie gestoken in het disciplineren dan het beschrijven van het diagnostisch proces. Diagnostici worden opgevat als een zootje ongeregeld. Ze doen maar wat als je ze hun gang laat gaan. Er zijn voorschriften ontleend aan de argumentatieleer, logica, statistiek en het experimentele ontwerp. Het laatste ontwerp domineert in Nederland en België als het Hypothese Toetsend Model (HTM). Het proces is niet alleen het afwickelen van een aantal lineair geordende stappen. Er is ook communicatie tussen diagnosticus en cliënt en dat kan alle kanten opgaan. Het proces is een sociaal systeem. Interactie en communicatie zijn dynamisch. De diagnosticus is geen boekhouder van de ziel

van zijn cliënt en die is op zijn beurt geen experimenteel subject, geen normproefpersoon. Zelfs aanhangers van- en geoeffenden in het HTM voeren de procedure niet als voorgeschreven uit (Groenier et al.<sup>22</sup>).

*Wat is diagnostiek dan nog?* De aard van de elementen, oriëntaties en bronnen maakt diagnostiek en het diagnostisch proces tot een geordende, open, onbepaalde, niet-deterministische en ondergedetermineerde activiteit om de vraag/het probleem van een cliënt te beschrijven/verklaren/duiden/ een plaats te geven, steunend op het kennisbestand van de psychologie. Open betekent niet: idiosyncratisch, willekeurig. Het systeem stelt de diagnosticus in staat om idiografische hypothesen over vragen/problemen van cliënten te duiden, verklaren en uit te leggen voor zover theorievorming en empirisch ondersteunende kennis uit de bronnen en oriëntaties, operationalisaties en instrumentarium dat mogelijk maken. De diagnosticus bepaalt daarbij zélf welke protocollen, constructen, metingen, instrumenten hij nodig heeft. Hij hoeft niet alles uit te vinden omdat er top-down (bijvoorbeeld het HTM) en ook door schade en schande totstandgekomen werkbare protocollen (Van Striens bottom-up praktijktheorieën<sup>23</sup>) zijn. De activiteit omvat een beperkt aantal aanbevolen stappen waar hij flexibel mee omgaat. De vraag/probleem van de cliënt is maatgevend voor zijn zoektocht. Hij is eerder *zoeker, detective* dan *onderzoeker*. Hij is vertrouwd met voorschriften die het proces stroomlijnen, disciplineren en overdraagbaar maken.

Het vertrouwen in de voorschriften berust op hun logica en plausibiliteit. Iets is plausibel als het ervaren wordt als waar. Het hoeft niet waar te zijn, maar een bewering over gedrag wordt meer waar als het daarnaast consistent is met andere ervaringen die ook als waar worden ervaren. Als je ook nog een empirische toetsing bij kunt leveren wordt de bewering stevig. Empirische vergelijkingen van protocollen met spontane werkwijzen zijn schaars. Het bewijs dat ze minder fouten opleveren is (nog) niet geleverd. Dit ontslaat de diagnosticus niet van de plicht tot verantwoording en met een protocol is dat gemakkelijker dan zonder. Het vermijden van een definitie van diagnostiek is een keuze. Het sluit aan bij de aarzeling van wetenschappers om hun theorie en onderzoek als hermetisch bouwwerk op te vatten. Deze behoedzaamheid vloeit ook nog voort uit de bescheiden prestaties van het kennisbestand voor beschrijven, voorspellen, verklaren/ controleren van- en beslissen over het gedrag van een cliënt en de effecten van behandelingen, interventies en therapieën. Dit getuigt niet van de weke aard van de diagnosticus, maar van realiteitszin.

### 3. Wat is gedrag

*Gedraag je!* Ieder van ons heeft de imperatieven: gedraag je, denk eens na, wees bewust, gehoord van opvoeders: ouders, leraren en politici. Ze gingen er van uit dat we heer en meester waren over ons gedrag. *Maîtres et possesseurs du monde*, zei Descartes. We wisten wel ongeveer wat ze ermee bedoelden. Zij hadden andere normen in hun hoofd. We moesten ons leven beteren. We zijn onvoldoende anarchist om ons niet te gedragen, niet te beheersen en niet na te denken. Volgens Proudhon<sup>26</sup>, de 19<sup>de</sup> eeuwse anarchist en voorloper van Marx en Piketty hield dat in dat we bewaakt, geïnspecteerd, gereguleerd, in een vakje gestopt, geïndoctrineerd, de les gelezen, gecontroleerd, beoordeeld, geteld, geformaliseerd, getarifeerd, geannoteerd, vermaand, berispt en verbeterd moeten worden.

*Een eerste omschrijving:* Als student namen we genoeg met de definitie uit het handboek en gingen over tot de orde van de dag. In de doctoraalfase kregen we een strikt behavioristische definitie van een glimlach te horen: alle er bij betrokken spiertjes vernamen we zonder een spier te vertrekken. We schreven die op in ons collegedictaat. Er werd geen *PowerPoint* getoond en de tekst stond niet op het internet. De docent had doorgaans geen uitgeschreven tekst. Maar wat is gedrag?

*Maak je niet druk:* Gedrag moet je niet thematiseren, laat staan problematiseren. We weten wel zo ongeveer waar het over gaat. Het is een *funktionierendes Begriff*, dat gaandeweg duidelijk wordt. Maar fenomenologie is niet meer *en vogue* en wordt tussen haakjes gezet, evenals vraagstukken over *mind-body*, theorie-praktijk en theorie-meten. De geest van het pragmatisme waart rond want waarom zou je, zoals James<sup>27</sup> in het begin van de 20<sup>ste</sup> eeuw opgeschreven heeft vragen stellen waar geen antwoorden op gegeven kunnen worden? Een pragmatische houding betekent dat je je niet laat verleiden om op onoplosbare, essentialistische, *foundational* (Rorty<sup>28</sup>) vragen in te gaan. Maar, we houden niet op onbeantwoordbare vragen te stellen: Wat is gedrag? Blijft dat niet boven de markt hangen? Leidt deze vraag, evenals 'Wat is filosofie?' er niet toe dat we de psychologie en onszelf telkens opnieuw uit moeten vinden? Er is immers nauwelijks cumulatieve kennis in deze tak van wetenschap. Is ze gedoemd om haar object voortdurend en nooit afdoende te definiëren: een variant op de mythe van Sisyphus?

Psychologen willen er al vanaf sinds het eind van de 19<sup>de</sup> eeuw. Maar nogmaals: Wat is gedrag? Vonden we geen sporen van alles en nog wat in de teksten die we moesten lezen zoals de objectieve behavioristische definitie, het onderscheid met dieren, de culturele praktijken en het bewustzijn? En, wat hebben mythen en literaire bronnen allemaal niet gesuggereerd over gedrag?

*Wat staat in woordenboeken?* Bergner<sup>29</sup> analyseerde 26 psychologiewoordenboeken en vond als dominante definitie: '*...any observable overt movement of the organism, generally taken to include verbal behavior as well as physical movement*'. Watson wegens succes verlengd! Het behaviorisme is wel wat opgerekt door veroorzakende stimuli complexer te maken.

Alles werd een stimulus van voedsel tot een model en zelfs het gezin en de cultuur. Het is de degelijke  $S \rightarrow R$  traditie. De S werd geprofileerd en genuanceerd maar de R, het gedrag werd kort gehouden. Geen discussies over hoe onze R verschilt van de R van dieren. Intentionaliteit en bewustzijn staan tussen haakjes. De S-R relatie is wel ingewikkelder gemaakt: ze is bidirectioneel en transactioneel, maar de R bleef wat Bergner vond in de woordenboeken. Een rookgordijn om de R niet te thematiseren lijken gesofisticeerde analyses van verbanden met complexe modellen, zoals *Structural Equation Modeling*, *multilevel*- en tijdserie analyse. Je vergeet te vragen wat de R eigenlijk is. Je hebt je handen vol aan het begrijpen wat de *ARIMA* modellen - *AutoRegressive Integrated Moving Average* - zijn en doen.

*Is de R van de leerpsychologie te mager?* Watson<sup>30</sup> heeft ons met zijn manifest uit 1913 bevrijd om door middel van introspectie psychisch functioneren en bewustzijn te onderzoeken. Er zit wel iets in zijn definitie van gedrag, ook al zeiden bezoekers van zijn laboratorium met ratten verontwaardigd: ‘...and they call this psychology’. Zijn betekenis wordt onderschat bij het diagnosticeren van vragen of problemen van de cliënt. Die onderschatting hangt wellicht samen met de atomistische aanpak en het manipulatieve karakter van gedragsmodificatie. Dat voelt niet goed. We zien ons gedrag als vrij en complex. Dat moet wel, denken we, want we zijn aangepast aan een technologische samenleving en aan een wereld vol complexe sociale rituelen. Redden we dat met S-R associaties, al zijn ze elegant en met mooie symbolen weergegeven?

Neem bijvoorbeeld Rotter<sup>31</sup>. Hij probeerde individueel gedrag te voorspellen met behulp van formules met controlerende en uitlokkende factoren. Ieder individu heeft een uniek *behavior potential* (BP), die een functie is van *expectancy* en *reinforcement value*:  $BP_x, s_1, ra = f(Ex, ra, \& Rva, S_1)$ . Het gedragspotentieel ( $BP_x$ ), dat verschijnt in situatie 1 [ $s_1$  in relatie tot reinforcement a ( $ra$ )], is een functie van de verwachtingen voor gedrag x in situatie 1 ( $s_1$ ) en de reinforcementwaarde van a in situatie 1 ( $Rva, S_1$ ). B = gedrag; R = reinforcement, E = de verwachting reinforcement te krijgen. E is opgebouwd uit de ervaring van steun of reinforcement en het ontbreken daarvan. Staat mooi, of is het een imponerende formulering van open deuren? Zo eenvoudig is het ook weer niet want de formule stelt de persoonlijke geschiedenis van steun, reinforcement aan de orde. Die leidt tot een verwachting van wat een bepaald gedrag de persoon gaat opleveren. Een persoon heeft immers een geheugen.

Volgens behavioristen is het doel van gedrag het vermijden van pijn en verkrijgen van genot en welbevinden. Toch zit er een adder onder het gras: het is moeilijk bij zulke formules concreet gedrag te vinden. Wat moet men immers allemaal te weten zien te komen? Wie is de betrouwbare informant? Is de informatie betrouwbaar en valide? Rotters idee over verwachtingen van reinforcement is omgewerkt tot een individuele verschillen dimensie, een trek: ‘interne versus externe controle’. Mensen die hun steun zelf kunnen regelen zijn intern gecontroleerd. Een trek werkt gemakkelijk in correlatieve onderzoek.

*Leren van modellen*: Je leert niet zo maar en niet van iedereen. Modellen moeten goed zichtbaar zijn en status en succes hebben. Net als Rotter heeft Bandura<sup>32</sup> een instrument

gemaakt dat individuele verschillen meet. De leergeschiedenis laat sporen na en personen ontlenen aan hun geschiedenis van succes en falen een inschatting van hun eigen mate van doeltreffendheid: *self-efficacy*.

*Gedrag volgens de psychometricus:* Hij weet evenals de (neo)behaviorist gedrag goed in de hand te houden: een goed of fout antwoord op een item. Dat is een klaar en distinct idee met lengte, breedte en hoogte. Verder is er expertise voor categoriseren, schalen en het vaststellen van het aantal schalen. Het intervalmeetniveau wordt daarbij het meest gebruikt. Dat wordt met een model geverifieerd (Item Respons Theorie) of voorondersteld (Klassieke Testtheorie).

*En, de diagnosticus?* Hij stemt *niet* in met deze omschrijvingen. Hij vat het gedrag van de cliënt op als een poging iets te bereiken, zoals groeten, stemmen of waarschuwen. Hij wil bijvoorbeeld succes op school of in het beroep vaststellen, condities identificeren die gedrag in stand houden, angst van een cliënt bepalen. Hij zoekt naar de oorzaken en redenen voor het gewenste of het problematische gedrag van zijn cliënt, of wil zijn vraag beantwoorden. Dat gedrag is een empirisch verschijnsel dat enigszins gedekt wordt door begrippen, constructen, die intelligentie, cognitie, persoonlijkheid en de sociale en fysische context beschrijven.

*Parameters van gedrag:* Bergner was niet tevreden met de gedragsdefinitie die hij distilleerde uit de 26 psychologiewoordenboeken. Hij stelt acht parameters voor die de kern van gedrag weergeven. Zijn idee van parameters komt uit de psychofysica. Kleur is daar volledig beschreven met de parameters helderheid, verzadiging en tint. In de fysica is het een golflengte met omschreven kenmerken en ook daar uiteraard volledig bepaald en beschreven. Waarom dit ook niet eens voor gedrag geprobeerd? De eerste parameter is identiteit: iemand zal zijn gedrag het zijne noemen; hij is de eigenaar van zijn gedrag. De tweede is wensen, willen, want een persoon wil iets, bijvoorbeeld een beroep kiezen, een opleiding volgen, een partner vinden. Volgens Schopenhauer<sup>33</sup> is dit de kernparameter van menselijk gedrag. Daardoor is leven lijden want een mens houdt nooit op met iets te willen dat hij niet heeft of is. De derde is een cognitie: gedrag betekent weten en onderscheid maken. Wat is bijvoorbeeld een beroep en hoe is het onderscheiden van andere. De vierde is weten hoe: bij weten hoort een competentie, een vaardigheid, een manier om iets totstand te brengen. De vijfde is *performance*. Dat duidt op proces- en procedure kenmerken van gedrag. De zesde is het bereiken van iets, bijvoorbeeld het slagen in een beroep, het vinden van een partner. De zevende is het profiel van persoonskenmerken. Deze parameter is in de persoonsleer uitgewerkt. De achtste parameter is de betekenis van gedrag met het oog op het doel van de cliënt en zijn sociale omgeving.

Acht kenmerken is veel. Het is niet gemakkelijk om elk gedrag er mee te beschrijven of elk construct op deze wijze uit te werken. Dus komt het er niet van. Dit neemt niet weg dat het een nuancering is van- en aanvulling op de definitie met een laag niveau subjectiviteit (de klassieke R), waarin gedrag nog net niet de uitkomst van een machine of automaat is.

Bergner zoekt de definitie bij de (psycho)fysica. Dat heeft iets ongemakkelijks. Snap ik het

eigenlijk wel? Gaat psycho(fysica) niet over iets anders? Eenzelfde ongemakkelijk gevoel kan een psycholoog/diagnosticus overvallen als er aansluiting wordt gezocht bij biologie, biochemie en neurologie. Wat moet hij met het DNA, de genetische uitrusting, EEGs, EVPs (EVoked Potentials)? Gedragingen zijn toch molaire eenheden die je met begrippen en constructen omschrijft en aanduidt. Dat aggregatieniveau kun je nog snappen. Hoe kan iemand nu aan het DNA of de hersenscan van iemand zien of hij een goede manager wordt? De diagnosticus gaat het om psychologische processen en constructen.

*Wat we allemaal weten van ons gedrag:* Smedslund<sup>34</sup> werkte, evenals Bergner de R uit. Hij wijst op gedragskenmerken die we *onmiddellijk* begrijpen: *Mentality*: mensen weten, denken, willen, voelen, nemen waar, horen en doen van alles om iets te bereiken, dat voor hen waarde heeft. *Reflectivity*: mensen zijn zich voor een deel bewust van wat ze willen en doen. *Verbality*: mensen kunnen voor een deel zeggen wat ze weten en denken en ze praten terug. *Hedonism*: mensen willen zich goed voelen als resultaat van wat ze weten, kennen en doen. *Learning*: wat mensen willen weten is verbonden met wat ze eerder geleerd hebben en nog willen leren. *Vulnerability*: mensen weten, voelen, denken, dat anderen goed of slecht kunnen handelen en zo behandeld kunnen worden. Iedereen is kwetsbaar. *Responsability*: mensen weten dat ze verantwoordelijk zijn voor wat ze willen, denken en doen. *Ethicality*: mensen willen doen wat goed is en vermijden wat fout is. Ze zoeken consensus: ze willen uiteindelijk dat iedereen met hen meedenkt, althans in eerste instantie, pas daarna mag er kritiek zijn. De eerste verwachting is: instemming van anderen en het idee dat ze gelijk hebben. Smedslund wijst op kenmerken die we aanvoelen en herkennen. Of je er snel en vruchtbaar mee aan de slag kunt in empirisch-analytisch onderzoek is een andere kwestie.

*En de wijsgerige antropologie?* Sommigen van ons kregen wat filosofie in het eerste jaar. De vakgroep Functioneleer 'doet het erbij' in de 'Inleiding in de Psychologie' op sommige universiteiten. Een filosoof doet 'het er niet bij'. Hij heeft iets te vertellen. Hij kan aannemelijk maken dat de persoon en zijn gedrag niet afdoende in behavioristische termen beschreven kunnen worden. Ook dat noteerden we en gingen over tot de orde van de dag. Verschilt ons gedrag niet van dat van dieren? Heeft De Waal<sup>35</sup> (*Bonobo en de tien Geboden*) ons er niet van overtuigd dat apen en olifanten altruïstisch zijn en rouwen? Rutten vertelde in 1966 vlak voor zijn pensioen in een college dat hij zich bevrijd voelde toen hij geen proefpersoon meer hoefde te zijn bij experimenten om via introspectie het *bewustzijn* te bestuderen. Eindelijk iets concreets: gedrag. Maar, daarna werd gedrag ongeveer alles wat je maar kunt bedenken: waarnemen, handelen, voelen en ... denken. Gedrag is van alles, bijna alles, ook bewustzijn, waar de psychologie mee begonnen is.

*Bewustzijn*: Dat is een uniek kenmerk van de mens. Hij is het *animal rationale* van Aristoteles, *res cogitans* bij Descartes en door Kant voorzien van synthetische oordelen apriori en aposteriori. Hegel<sup>36</sup> maakte het bont door ons bewustzijn als het verschijnen van de *Geist* op te voeren. De rationalisten en idealisten konden er iets van als het over de bijzondere

en verheven menselijke conditie en prestaties ging. Empiristen zijn bescheidener. Ze bouwen bewustzijn op uit ervaringen door middel van associaties volgens contiguiteitswetten: het bewustzijn als doorgeefluik van de ervaring uit de buitenwereld en een ook beetje van belevenissen uit de binnenwereld. Beide beschrijven en onderzoeken gedrag als informatieverwerking. Het idee van een gelaagd bewustzijn (Freud) is niet verdwenen en nu nog controversieel, denk aan de *priming* studies. We kunnen er als psycholoog niet om heen iets te zeggen over het bewustzijn. Je kunt omtrekkende bewegingen maken en uitwijken naar biologie en natuurkunde. Maar hoe van materiële processen te komen tot het besef dat we waarnemen, fantaseren, verlangen, voelen, herinneren, plannen, gedichten lezen, pooltochten dromen? We kunnen proberen kenmerken van bewustzijn te benoemen (De Mul<sup>37</sup>). Bewustzijn is mijn bewustzijn en weerspiegelt *mijn* perspectief. Ik kan me nooit volledig in een ander of iets anders verplaatsen. Mijn bewustzijn verschaft mij een *Gestalt*: een betekenisvolle samenhang, geen losse objecten en verschijnselen, inclusief de mens als een ondeelbaar geheel, een *individueum*. De bewuste ervaring biedt mij een *kwalitatieve* wereld. Een fysicus moet mij vertellen dat wat ik groen noem een golflengte is. We zijn ons daarom niet rechtstreeks bewust zijn van gevaren van bijvoorbeeld gammastraling. Het bewustzijn is niet in zichzelf *opgesloten*; het ligt niet verpletterd op zich zelf, zoals de dingen, *les choses* (Sartre<sup>38</sup>). Of het bewustzijn zo transparant is als hij beweert, kan betwist worden. Ik ben me bewust van de buitenwereld en ben er op gericht (*intentionaliteit*) en ik zie objecten en mensen, niet de representaties daarvan op mijn netvlies. Bewustzijn is ook *zelfbewustzijn*: ik ben me bewust dat ik iets zie, voel, fantaseer, lees, leer, haat, lach. Ten slotte is bewustzijn een *dynamisch* proces en geen pakketje losse producten. Het is een stroom van waarnemingen, gevoelens, gedachten, weetjes, feiten die alle kanten op kan gaan. Ik kan die stroom zelf een beetje verleggen maar het meeste overkomt mij; tweederde zegt de filosoof. Hij heeft het over de *Wandering Mind*. Daarom weten een ander en ik zelf nooit precies hoe het gedrag nu exact, stap voor stap in elkaar zit. Ik probeer het daarom maar rationeel, lineair na te bouwen en te modelleren. En, dat is psychologie! Een echte materialist zal zeggen dat hij de stroom radicaal kan verleggen en hij ons een nieuw zelf, een nieuwe identiteit kan bezorgen. Hij doet iets met onze hersenen door middel van medicijnen, drugs en operaties waardoor onze hersenen ons 'bedriegen'. Wie is hier de bedrieger, wie de bedrogen?

*Gedrag is raadselachtig*: Gedrag definiëren vooronderstelt zicht op wat de mens *is*. Naast gewoontedier is hij is dynamisch, grillig, zichzelf een raadsel, zonder klare en distincte gedachten, kwaliteiten, zonder doel en bestemming. Daarom moet hij van alles verzinnen om orde te scheppen. Die orde wordt door onderzoekers systematisch beschreven én gecreëerd. Dat omschrijven kan alle kanten op gaan. Ik interpreteer enkele zinsneden uit de literatuur. Ze zijn geheimzinnig (*Der Mensch is nur heim ins Ge-heim-nis*: Heidegger) en er valt van alles in te lezen: *intus legere*: volgens St Thomas is dat precies intelligentie. Gedrag is de enigsmatische ziel van de psychologie.

Als eerste een zinnetje van Heraclitus (6de - 5de eeuw BCE): *ethos anthropoi daimon*. Je kunt het vertalen als: je aard, je natuur, je gedrag is je demon, of je gedrag is je Lot. Gedrag is niet neutraal, het achtervolgt je als een demon, die goed of kwaad kan uitpakken. Of, het overkomt

je, zoals hoofdpersonen in tragedies die niets anders kunnen als het lot ondergaan, denk aan Oedipus.

Als tweede, iets uit Plato's dialoog met Protagoras. De laatste was sofist en vriend van Socrates. Er wordt een mythe vermeld die over gedragskenmerken van de mens gaat. Bij de Grieken zijn mensen stervelingen en dat onderscheidt hen van de goden. De eigen aard van de mens komt tot uitdrukking in een mythe: De oppergod Zeus geeft Prometheus en zijn tweelingbroer Epimetheus de opdracht om diegenen ter wereld te brengen die *niet* onsterfelijk zijn: planten, dieren en mensen. Ze mogen een aantal kenmerken, krachten, begaafdheden (*dynameis*) verdelen zodat uit klei planten, dieren en mensen kunnen ontstaan. Prometheus doet dat op verzoek van Epimetheus. Hij is slordig en verstrooid (lage score op *Conscientiousness*) maar hij weet achteraf dat hij fouten maakt en voelt zich schuldig (hoge score op *Neuroticism*) en denkt er herhaald en ernstig over na (het maalt in zijn hoofd: rumineren, is een depressie onderweg?). Hij komt erachter dat hij vergeten is mensen kwaliteiten te geven. En wat hij had, was uitgedeeld. Het gevolg is dat het met de planten en dieren wel goed zit: deze hebben hun eigenschappen en kunnen ermee uit de voeten/poten/vleugels/vinnen. Maar, de mens is ertoe veroordeeld om naar zijn kwaliteiten te zoeken: dat is zijn lot, zijn demon. Daardoor is hij het nooit met anderen en/of zichzelf eens over zijn doel en bestemming, denk aan Kouwers *Spel van de Persoonlijkheid*. Ter compensatie en om te voorkomen dat mensen in leegte leven (*horror vacui*) gaf Zeus ze via Hermes, van wie ons begrip hermeneutisch stamt, twee gevoelens: rechtvaardigheid en daarnaast schaamte/ terughoudendheid/ eergevoel. Prometheus steelt bovendien als goedmakertje het vuur van de goden en geeft het aan de mensen. Vuur kunnen we lezen als het vermogen technische middelen te ontwikkelen zonder welke ons huidig bestaan niet mogelijk is. Wat te doen zonder fiets, auto, Ipad, bril, plastische chirurgie, tablet? Mensen moeten met andere woorden hun bestaan *uitvinden*.

Bergner, Smedslund, wijsgerige antropologie, een enkele filosoof en oude geschriften laten zien dat er naast behavioristisch objectief waarneembare en bij tests objectief scorebare antwoorden - subjectieve, unieke en grillige gedragingen en uitingen zijn. Gedrag is voor de mens een raadsel. Strasser tijdens een mondeling tentamen '*...hoe kan de zich ontwikkelende, de ontwik- keling doorgronden?*

*Hoe 'staat de onderzoeker er in'?* Hij kan, mag, moet in zijn onderzoek de veelvormigheid, het bijna alles karakter van gedrag tussen haakjes zetten. De onderzoeker moet kiezen voor een segment én een opvatting van gedrag: '*Vous êtes libre, choisi!*' Hij/zij kan niet zonder reductie en dat betekent gelijktijdig onthullen en verhullen wat gedrag is. Als hij integratie opvoert om de reductie te overwinnen creëert hij meestal onwerkbare hybriden. In empirisch onderzoek doet reductie zich soms voor als doen alsof je neus bloedt. Men kijkt weg bij andere uitwerkingen en interpretaties van gedrag: verhullen zonder het er eerlijk bij te vertellen. Ook niet in de discussie waarin de onderzoeker doorgaans deemoedig de beperkingen van zijn onderzoek beschrijft. De steekproef was beperkt (wanneer niet?), de meetinstrumenten waren niet perfect betrouwbaar (zijn ze nooit). De statistische techniek was niet gevoelig genoeg (pak dan een andere). Geen longitudinale data (doe het dan). Hij toont zijn deemoed als een

ritueel. Gedrag is zo ongeveer alles. Onderzoek onthult een segment van wat mensen doen, denken en voelen en verhult - zo doende en denkende - andere. Je moet niet alles tegelijk willen, dat kan niet.

*Hoe 'staat de diagnosticus er in'?* Hij probeert - in tegenstelling tot de onderzoeker - contrafactisch meer ballen tegelijk in de lucht houden. Hij is niet de onderzoeker die een correlatie hoe complex ook geanalyseerd gemiddeld (bijvoorbeeld pad- en *multilevel* analyse), gemodelleerd (bijvoorbeeld *Structural Equation Modeling*, *Multilevel Analysis*) wil bepalen. Ook wil hij geen effect aantonen van een variabele hoe ingewikkeld gemodelleerd en complex theoretisch ingebed ook. Hij wil inzicht in het gedrag van zijn cliënt, dat wil zeggen in zijn vraag of probleem. En de cliënt is weliswaar een element van een steekproef maar mogelijk geen willekeurige en hij is niet het gemiddelde van die steekproef. Zijn vraag/probleem eist dat een diagnosticus oog heeft voor vele jaszjes en lagen waarin gedrag gepast kan worden: (neo) behavioristisch, een latente trek, een profiel op een aantal factoren, een transactie tussen persoon en context en als resultaat van een normatieve en idiosyncratische ontwikkeling. De diagnosticus wordt bedolven onder nieuwe gedragsconstructen, bijvoorbeeld *Mindfulness*: 'blijf bij jezelf broeder' (Koot en Bie, zonder jaartal) of *Positive Psychology*: 'één op de twee plofkippen loopt als een kievit' (jtl), *Post Traumatic Growth*. Kennis van die nieuwe constructen behoort tot *achtergrondkennis* van de diagnosticus, alleen en vooral omdat de cliënt zulke kenmerken vermeldt, als hij zijn vraag toelicht of probleem uiteenzet.

#### 4. Recht doen aan de cliënt

Het streven is de cliënt *objectief* te beschrijven. De Groot rekent tot 'objectief' ook het recht doen aan de vraag, het probleem van een cliënt. Dat wordt hier ruimer gelezen dan in zijn empirisch-analytische epistemologie. 'Recht doen aan' daagt uit om een relatie te leggen tussen theorie over het gedrag en het gedrag zelf. Je moet een epistemologisch standpunt articuleren. Methodologie gaat vooral over werkwijzen om vertekeningen en fouten bij observeren, registreren en toetsen van gedrag te voorkomen. Ze heeft ook betrekking op het vermijden van onjuiste conclusies over oorzaken van- en samenhangen tussen gedragingen. Ze is als het ware een commissie voor de Zuiverheid in de Leer. In de diagnostiek zijn dwalingen de valse positieven en valse negatieven. Dit is een negatieve formulering. Het gaat er om de diagnosticus op methodologische fouten te betrappen. Een andere formulering is dat de diagnosticus het gedrag van de cliënt naar boven haalt, het voor zich laat spreken. Dit klinkt fenomenologisch, maar heeft onvermoede aanhangers: Piagets proefjes lieten het kennen en denken spreken, aan het werk zien.

*Diagnostiek is geen experimenteel/correlatieel onderzoek:* Het object van studie, de cliënt met zijn vraag, probleem kunnen we niet neutraal of objectief, abstract, dat wil zeggen los van enig kader bestuderen. We kunnen epistemologie niet omzeilen. Het gaat er niet alleen om hóe je het zegt, maar ook wát je zegt over dat gedrag. In de diagnostiek is het startpunt het gedrag van de cliënt: zijn vraag of probleem in psychologische zin. Dat is geen onderzoeksvraag zoals in een tijdschriftartikel. Er wordt geen hypothese getoetst in een steekproef. Het startpunt is uiterlijk waarneembaar, onwillekeurig, *covered* en doelgericht gedrag, psychische toestanden en processen, informatieverwerking, handelwijzen en totstand gebrachte producten van de cliënt. Hij wordt opgevat als persoon/ kenner/ denker/ leerder met zijn individueel trekprofiel, idiosyncratische ontwikkeling en inbedding in zijn sociale context. Omdat we aan onze impliciete en expliciete theorieën, referenti kaders vastzitten, is verstandig ze te onderkennen in plaats van de neutrale objectieve buitenstaander te spelen. Wetenschappers willen die rol wel eens op zich nemen: 'Ik meld alleen de feiten; wat u daarvan vindt is uw zaak, dat zijn uw woorden, dat is *in the eye of the beholder*'. Dat kan niet want er zijn onontkoombaar filosofische, ideologische, wetenschapstheoretische impliciete theorieën in het spel. Beide groepen zijn belast met niet gereflecteerde achtergrondkennis en apriori opvattingen en over het gedrag van de cliënt.

*Diagnosticeren is aan het licht brengen:* Bij het diagnostisch proces komt meer kijken dan het methodologische verantwoord, dat wil onder meer zeggen objectief beschrijven, voorspellen en verklaren van gedrag. We geven De Groots 'recht doen aan' een draai. We gaan terug naar de wortels. Aan Plato wordt toegeschreven dat hij de taak van theorie (belangeloos schouwen) opvatte als sooisdzein (σώζειν τὰ φενομένα). Dat verwijst naar leven geven (soois, bijvoorbeeld in de naam So-crates), redden, ongeschonden bewaren, recht doen aan, behoeden voor verduistering. Positief gezegd, aan het licht brengen van de fysische en sociale verschijnselen. En dat is een hele toer, zoals Heidegger ons inwijft. Hij beschuldigt filosofen en ons allemaal

- wij zijn immers *Das Man* - (1962, pp. 126-130) woordvoerders van Zijnsvergetenheid. We praten na wat anderen zeggen en dat heeft geen betrekking op het Zijn. De filosofie en de natuurwetenschap redden het volgens hem ook niet om verschijnselen, gebeurtenissen en gedragingen waar te nemen en te begrijpen. Ze zijn bedolven onder theorieën, ideologieën, belangen, taalgebruik, religies en technische ontwikkelingen. De vraag is hoe komt hij dan dichterbij dat authentieke Zijn? Wat moet er allemaal weggehaald worden? Blijft er een brute, ruwe bron over? Is de ui metafoor van Kouwer van toepassing?

We komen, denk ik, niet verder dan het steeds weer verschuiven van het perspectief op het gedrag van de cliënt. Dat is soms en met wat geluk een verrijking. Heideggers opmerking bevordert het alert zijn op het verhullende karakter van theorie, geloof, techniek en politiek. Het zijn immers slechts perspectieven. Ieder is veroordeeld tot een tunnelvisie. Als we geen tunnel, kader, theorie hebben zien we niets. We moeten het - als het om gedrag gaat - doen met de schaduwen uit Plato's allegorie van de grot. Geen prettig vooruitzicht. Je krijgt je vinger er nooit helemaal achter. Toch kan de diagnosticus niet weglopen voor zijn verantwoordelijkheid om een uitspraak te doen over het gedrag/probleem van de cliënt. Hij is echter niet naïef, want wetten die perfect voorspellen en verklaren bestaan niet. Er is ook geen protocol dat feilloos de weg wijst.

*Een schrale troost:* Door denken en onderzoek van evolutiebiologen, de opkomst van de thermodynamica en de kwantumfysica heeft een aantal wetenschappers het dogma van universele causaliteit en rationaliteit losgelaten. Dat werkt beperkt door in psychologie en bijgevolg in de diagnostiek. Na Newton kwam Darwin. Evolutie is volgens hem een populatieverschijnsel, geregeerd door toeval. Daarna kwam het werk van Boltzmann dat de basis vormde voor de thermodynamica. Darwin was Boltzmanns held. Hij probeerde wat Darwin had beschreven voor levende wezens uit te werken voor *nonreproducing matter*. Het gedrag van de materie moet volgen uit populaties en wordt ook geregeerd door kans (Schneider & Sagan<sup>37</sup>, p. 48). Later kwam de kwantumfysica die afweek van Newtons natuurkunde. Nu hoef je niet te begrijpen waar de kwantumfysica over gaat. Einstein zal ook gezegd hebben: '...mijn vrouw begrijpt mij niet...' (Herman Finkers). De strekking er van helpt om een idee te vormen van beperkingen bij voorspellen en controleren van gedrag.

De flamboyante natuurkundige Schrödinger<sup>39</sup> schreef een boek voor leken (1944, Inleiding: p. 10), waarin hij zijn waarneming/opvatting beschrijft dat atomen altijd en totaal wanordelijk bewegen in hitte. Er is verzet tegen ordelijk gedrag en een klein aantal atomen engageren zich niet zo dat er een herkenbare systematiek ontstaat. Een wet ligt niet voor het oprapen: zij ontstaat in de tijd als atomen samengaan. Als grote aantallen atomen samenwerken dan pas worden statistische wetten toepasbaar.

Zou dat een metafoor kunnen zijn voor het recht doen aan het gedrag van de cliënt? De voorspellingen worden accurater als er meer van die verzamelingen bij elkaar komen. Uiteindelijk krijgen gebeurtenissen voor onze waarneming ordelijke kenmerken; ze worden voorspelbaar en controleerbaar. Schrödinger zegt dat alle fysische en chemische wetmatigheden die nodig zijn voor het leven van organismen altijd weer verstoord en onwerkzaam gemaakt kunnen

worden door de onophoudelijke beweging van atomen. Zou dat geen metafoor kunnen zijn voor het gedrag van onze cliënt? De Brownse beweging van stuifmeelkorrels in een vloeistof is gekenschetst als stochastisch, onvoorspelbaar en chaotisch. Kleine deeltjes gedragen zich grillig, *at random*. Een dergelijke metafoor voor gedrag van een persoon of groep bereidt de diagnosticus er op voor dat er altijd afwijkingen van een normatief patroon kunnen zijn.

Deze *zijstap* over het loslaten van universele causaliteit suggereert dat de psycholoog/diagnosticus de hoger in de pikorde staande wetenschappen er niet bij hoeft te halen om de cliënt het vergezicht van perfecte voorspelbaarheid en controle te bieden. Hij komt van een koude kermis thuis. Hoeveel assemblees van deeltjes, genen, zijn er niet nodig om tot een samenhang met concreet gedrag komen? Het laat ook zien dat in onderzoek het juiste niveau van aggregatie gevonden moet worden om statistische wetten toe te kunnen passen. Daarom werken we zo graag met steekproeven en met herhaalde metingen. Toch lijkt in psychologisch onderzoek er een voorkeur te zijn voor: 'hoe kleiner de gedragseenheid, hoe beter'. Zijn we op weg naar een deeltjespsychologie of hebben we die al in de vorm van gefragmenteerde studies in de tijdschriften? Willen we de echte - de kleinste - deeltjes vinden in de zoektocht naar het DNA, de genen en hun samenstellingen en de hormonen? Niemand kan dat allemaal begrijpen en de studies erover bijhouden. En is het voor psychologische diagnostiek nodig en nuttig? Een Nijmeegse, toentertijd nog aanstaande professor die me mondeling tentamen functioneel afnam, was het met me oneens - het kostte me één punt, nog wel geslaagd - toen ik een hoger niveau van aggregatie voorstond dan hij. Ik verdedigde de leerstappen uit het boek van de leerpsycholoog Gagné; hij zag meer heil in het geheugenonderzoek met behulp van de *memory drum* van Bower. Dat laatste zou je volgens hem uiteindelijk beter inzicht geven in hoe mensen leren. Is daar bijna 50 jaar later voldoende bewijs voor?

*Een zinvol aggregatieniveau:* Een effect of samenhang wordt opgespoord in een steekproef en geldt niet één-op-één voor een individu. Een alternatief is het herhaald meten van een individu. Dat is niet gebruikelijk in de psychologie: men gaat zonder argumenten van de steekproef naar het individu. Men gaat zelden van een individu naar de groep of steekproef, waar het individu een exemplaar van is. Waarom niet? Deze voorkeur maakt de interpretatie van bevindingen uit steekproeven voor individuele diagnostici lastig. Je hebt robuuste en stevige empirisch resultaten nodig. Bovendien nemen we met Schrödinger aan, dat je macro-verschijnselen als objecten, gebeurtenissen en gedrag niet kunt begrijpen, verklaren uit de kleinste samenstellende delen, zoals, atomen, genen, moleculen, EEGs, EVPs, RTs of behavioristisch geformuleerde responsen. Gedrag is een uitdrukking van levende wezens met een identiteit. Volgens Schrödinger gedragen individuele constituenten zich onvoorspelbaar. Slechts een groot aantal samenwerkende constituenten maakt het mogelijk statistische analyses te doen en probabilistische wetten te formuleren. Dat *aggregate* is in de diagnostiek vooralsnog het gedragsbegrip, het gedragsconstruct. Daar moeten we het mee doen, hoe veel kritiek er ook op mogelijk is. In de psychologie bereiken predictie en controle bescheiden (*modest*) tot gemiddelde (*medium*) waarden volgens de vuistregel van Cohen. De uitkomsten van meta-studies bevestigen dat keer op keer.

*Emergentie:* Turbulente, levende wezens die interacteren met drieste fysische en sociale contexten leveren onverwachte, nieuwe uitkomsten en betekenissen op. Emergentie doet zich voor bij complexe systemen en is een gevolg van het feit dat natuurlijke eenheden de tendens tot zelforganisatie vertonen. Emergente verschijnselen als gedrag en bewustzijn vertonen eigenschappen, zoals zelfregulering, groei en zelfreproductie. Deze kunnen niet afgeleid worden uit de elementen. Het erkennen van dit niet-reductionistische procedé (het geheel is meer dan de som van de delen) helpt afstand te nemen van een mechanistisch causaal begrijpen van gedrag. Het helpt om het idee van de tot op grote hoogte maakbaarheid van gedrag van individuen en groepen los te laten. Zou het streven 'recht te doen aan' geholpen zijn door rekening te houden met complexiteit, zelfregulering, groei en onvoorspelbaarheid van het gedrag van de cliënt?

*Conclusie:* De diagnosticus kan niet dan bescheiden zijn over zijn beschrijvend, voorspellend en controlerend vermogen. Dit komt voort uit respect voor- en kennis van de bijzondere aard van de cliënt en zijn sociale omgeving. Als hij uitdraagt dat hij het wel voor elkaar krijgt om perfect te voorspellen en verklaren - zodra ik alle factoren ken en onderzocht heb; het is een kwestie van tijd: *much research has to be done*, als de *funding* maar ruim genoeg is - dan hangt hij de Newtoniaanse idee van wetten aan. Die vlieger gaat in de psychologie niet op. Hij pretendeert een geldige analyse te maken van de vraag, het probleem van een cliënt. Er is geen andere basis voor deze analyse dan het kennisbestand van de psychologie. Hij is geen dokter, neuroloog, chemicus. Zijn verhaal is even sterk het geheel van theorie en empirische bevindingen. Zo doende, is het succes van diagnosticeren en van het diagnostisch proces een toets voor de psychologie: *Psychologische Diagnostiek is Diagnostiek van de Psychologie*.

## 5. Diagnosticeren is gedrag

Als gedrag ongeveer alles is, dan valt diagnosticeren eronder. De imperatief: 'gedraag je' geldt ook voor deze activiteit. En hoe, de diagnosticus is een brokkenmaker, hij moet gedisciplineerd worden door protocollen, statistiek, regels en supervisors. Voor elk gedrag geldt dat er individuele verschillen zijn, er zich ontwikkeling voordoet en het gevoelig is voor de sociale context. Voorschriften pretenderen dit terug te dringen maar dat lukt niet volledig. Ik gebruik de opbouw van betogen in de Heidelbergse Catechismus als hulp. Deze structuur werd in een ver verleden geëtaleerd in preken. Dat zijn toespraken waarop geen weerwoord geoorloofd is. Deze begint met het schetsen van de penibele situatie waarin de mens verkeert: de diagnosticus wordt bestraffend toegesproken en geconfronteerd met zijn falen. Hij verweert zich: 'Ik doe het toch af en toe ook goed, ik doe niet alles verkeerd'. Maar de wetenschap is streng en het publiek meedogenloos. Er is verlossing: protocollen, modellen, regels en statistiek. Daarna zou onze diagnosticus zich dankbaar moeten tonen. Dit impliceert dat hij zijn gebrekkigheid inziet en dankbaar alle hulp aanvaardt.

**Stadium 1 Falen:** De diagnosticus wordt bestraffend toegesproken. Hij is eenvoudig van geest: een paar diagnostische cues optellen en meteen voorspellen en beslissen. Hij neemt empirisch afgeleide formules niet serieus, terwijl die de gemiddelde fout minimaliseren: de kleinste kwadraten oplossing. Dat stelde Meehl<sup>39</sup> al in 1954 vast: ... *in all* (het waren er op dat tijdstip 20), *but one of which the predicti made actuarially were either approximately equal or superior to those made by a clinician* (p. 119). Dit is herhaald ondermeer door Grove et al.<sup>40</sup> in een meta-studie met 617 vergelijkingen en Aegisdotti et al.<sup>41</sup> met 92 vergelijkingen. De formules deden het beter bij recidive, schoolsucces, leiderschap, fraude, jeugddelinquentie, psychiatrische diagnoses, geweld, vrijlaten op borgtocht, huwelijksgeluk, productiviteit, succes als manager, diagnose van homoseksualiteit, faillissement van bedrijven, beroepstevredenheid, adoptie, succes als ondernemer en therapie succes. Van de voorspellingen was 47% ten gunste van de formule, 47% gelijk en in 6% kwam de spontane diagnosticus als winnaar uit de bus. Aegisdotti et al. vatten 56 jaar onderzoek naar succes van interventies in de geestelijke gezondheidszorg samen. Ze berekenden d-waarden van 92 vergelijkingen over hersenletsel, persoonlijkheidsstoornissen, duur van verblijf in het ziekenhuis, psychiatrische diagnose, aangepast zijn, prognoses van ziekteverloop, geweld, IQ, schoolprestaties, MMPI-profielen, zelfmoordpogingen en homoseksualiteit. Een geselecteerd deel van de steekproef bestaand uit 36 studies toonde absolute effectgrootten van  $d = 0,57$  en  $d = 0,73$ . De d-waarden waren in het voordeel van de statistische predictie.

*Waarom faalt de diagnosticus?* Hij grossiert in gemakzuchtige, onbewezen regels (*heuristics*) en vertekeningen (*biases*) en neemt statistische informatie niet serieus (Kahneman & Klein<sup>42</sup>). Zo'n fout is bijvoorbeeld de *availability heuristic*: we vinden gebeurtenissen waarschijnlijker, als we die gemakkelijk uit ons geheugen kunnen ophalen. Als een diagnosticus veel delinquenten in zijn praktijk ziet, zal hij gemakkelijk delinquentie waarnemen in gedrag van een cliënt ook als die geen delinquent is.

*Hij verwerkt informatie gebrekkig:* Leken en diagnostici zijn selectief, nemen waar in de richting van hun eigen standpunt, negeren zinvolle beschikbare informatie en houden geen rekening met het ontbreken van belangrijke informatie. Ze verdisconteren in het oog springende informatie meer dan empirische gegevens die over langere termijn verzameld zijn. Eén recent geval van kindermishandeling in de krant krijgt meer gewicht dan statistieken over jaren.

*Hij zondigt tegen kansrekening:* Hij schat de frequentie van verschijnselen onjuist. Als men vraagt wat gevaarlijker is, het weerlicht, de bliksem of het ontbreken van een trapleuning in huis, wordt het weerlicht regelmatig gevaarlijker geacht. Hij en leken onder- of overschatten verandering en groei. Verbetering op basis van interventies wordt over- en exponentiële groei onderschat. Ze baseren oordelen op infrequente, extreme waarden. Ze schrijven oorzaken onjuist toe. Stel, je goed - slecht voelen fluctueert. Als je je niet goed voelt, ga je naar de dokter. Hij (m/v) geeft je een medicijn of placebo. Omdat je toestand fluctueert is er een kans dat de toestand van slecht voelen overgaat in goed of neutraal voelen. De fluctuatie is de 'oorzaak' van de verandering.

*Hij ziet verbanden die er niet zijn:* Een voorbeeld is het beoordelen van homoseksualiteit op basis van menstekeningen. Men veronderstelde dat er een verband is tussen het tekenen van gespierde mannen en homoseksualiteit. Empirisch onderzoek heeft dit verband niet bevestigd.

*Hij denkt controle te hebben:* Dat komt behalve bij diagnostici voor bij begrotingen van ministeries, vijfjarenplannen van de Sovjetecconomie, kosten van de Noord/Zuidlijn, tijdstip van in gebruik nemen van de hogesnelheids- en de Betuwelijn. Ooit met enige regelmaat een hogesnelheidstrein zien rijden? Als effect van begeleiding van criminele jongeren in inrichtingen wordt vermeld dat er winst is geboekt, terwijl bij gelijkblijvende criminaliteit het geweld afneemt (ex directeur Rentray: H. Lodewijks op Radio 1, 2009). Het effect van een HALT programma voor jongeren die te veel alcohol gebruikten hadden en betrappt waren, werkte omgekeerd, averechts. Het was iatrogeen.

De diagnosticus en de leek krijgen er van langs. De psychiater Holt<sup>43</sup> (p. 386) nam '*...in non clinical psychologists a patronizing attitude to the clinical assessment enterprise* waar.' Stevenson van *Treasure Island* en *The strange case of Dr. Jekyll and Mr. Hyde*: '*Waartoe we ook bestemd zijn, we zijn niet bestemd tot slagen. Falen is het ons toebedeelde lot*'. En dat moeten we volgens hem met vrolijk gemoed dragen. Zal hem als van oorsprong strenge Schot soms moeilijk gevallen zijn. Misschien hielp het dat hij met een Amerikaanse gehuwd was.

**De diagnosticus verweert zich:** Zoveel kritiek lokt 'ja maar...' reacties uit. Van Dam<sup>44</sup> spreekt in haar monografie *Fixatie op fouten* over het plezier dat de onderzoekers ontleen aan het feit dat ze anderen op fouten betrappen. Zij legt de bewijslast voor het falen van de leek en diagnosticus bij de onderzoeker. De laatste gebruikt laboratoriumonderzoek om de proefpersonen er in te laten lopen, te bedotten. Hij moet ecologisch valide onderzoek doen. De diagnosticus is onschuldig tot het tegendeel is bewezen, zegt ze Funder na.

Brunswik schreef al vóór Meehls (1954) *disturbing little book* verscheen over het totstandkomen van een waarneming of een diagnose. Deze bevat enkele cues die opgeteld worden, elk gewogen naar zijn validiteit. De waarnemer/diagnosticus leert in de loop van de tijd want de werkelijkheid en andere waarnemers corrigeren hem als hij fouten maakt. Hij is adaptief. Brunswik gaat er niet van uit dat waarneming een slap aftreksel is van de Platoonse Idee en hij maakt van de diagnosticus geen structuralist die alles naar zijn had zet. Maar zonder zonden is ook weer niemand. Hij gaat hij er dan ook niet van uit dat we objecten, verschijnselen, gebeurtenissen en gedrag onmiddellijk en foutloos waar te nemen. Hij is ook geen naïef realist.

*Hij beroept zich op intuïtie:* Dat valt slecht want de wetenschapper heeft het Streng Methodologisch Geweten hoog in het vaandel. Intuïtief denken en handelen verwijzen naar moeiteloos begrip van bijvoorbeeld het gedrag van iemand. Het sluit de overtuiging in dat er een vermogen is om kennis te verwerven zonder stapsgewijs rationeel te denken en zonder langs logische weg conclusies te trekken. Mensen vellen oordelen en trekken conclusies zonder bewuste, gearticuleerde en verdedigbare stapsgewijze processen. Het proces verloopt snel, zonder overleg en wordt niet achteraf ge(re)construeerd.

Intuïtie komt overigens niet uit de lucht vallen. Er is een voorraad aan kennis over een onderwerp, ook al kan de persoon die niet articuleren. Een goede voetballer kan niet verklaren hoe hij de schitterende steekpass geeft, ook al heeft hij uitentreuren van de trainer gehoord dat hij vaste patronen, het tikkie-takkie van FC Barcelona, op het veld moet leggen. Intuïtie bestaat omdat een persoon beschikt over verschillende en complexe informatieverwerkende systemen. Ze kunnen los van elkaar opereren. Intuïtieve oordelen berusten op goed om je heen kijken, vrijuit speculeren en generaliseren, maar ook toetsen aan feiten en geïnteresseerde anderen serieus nemen.

Het lijkt er op dat de tijdgeest toestaat om de rol van intuïtie in de diagnostiek te onderzoeken. De DSM-5<sup>45</sup> vraagt bijvoorbeeld een klinisch oordeel: het schatten van de lijdensdruk. Dat gaat niet zonder gevormde, getrainde intuïtie.

*Natural Decision Making:* Experts moeten soms onmiddellijk beslissen wat te doen, bijvoorbeeld in het verkeer, bij calamiteiten, het schaakspel: Het is al beschreven in het werk van De Groot over *Het denken van de schaker* en in zijn *Methodologie* en in dat van Simon. Experts beslissen onder tijdsdruk, maken gebruik van hoog-niveau gespecialiseerde kennis en van goed getrainde vaardigheden. Denk aan brandweerofficieren, schakers, personeel op boortorens en soldaten/officieren in gevecht. Ze gebruiken niet-gearticuleerde kennis, bijvoorbeeld hoe de vlammen zich verspreiden, tekenen van het ineenstorten van een gebouw, inschatten van de gewoonten van de tegenstander. Ze beginnen niet met een aantal opties maar met één. Dit is de *Recognition-Primed-Decision* (RPD) strategie. Ze passen hun strategie meteen aan als die onjuist blijkt, zegt Klein<sup>46</sup>. Als ze *achteraf* een protocol moeten maken, construeren ze een lineair, rationeel verhaal.

*Met gestrekt been er in:* Gigerenzer<sup>47</sup> wilde laten zien dat wat in de *Heuristics-Biases* traditie zonde is, soms het beste is om te doen. Onder bepaalde condities verslaan snelle en spaarzame, zuinige (*fast* en *frugal*) heuristieken van deskundigen en beginnelingen de lineaire modellen. Als men Amerikaanse studenten vraagt welke stad meer inwoners heeft San Diego of San Antonio beantwoordt tweederde dat correct. Duitse studenten gaven vrijwel alleen goede antwoorden. De vuistregel lijkt: *...if one of two objects is recognized more or easier than the other, then infer that the recognized object has higher values with respect to the criterion* (Goldstein & Gigerenzer<sup>47</sup>). Zo'n regel zal niet altijd gelden, maar is soms effectief vanwege zijn ecologische rationaliteit. Bij de *Take The Best* heuristiek komen opeenvolgende *cues* beschikbaar. De eerste onderscheid makende aanwijzing wordt gekozen. De rest laat je links liggen. Als iemand moet beoordelen welk van twee steden de meeste inwoners heeft, dan kijkt hij naar een voordehandliggende *cue*, bijvoorbeeld welke stad heeft een topuniversiteit of een voetbalkampioen. Utrecht of Eindhoven dus.

**Stadium 2 Verlossing:** Gehoorzaamheid, volgzzaamheid en 'prevelen tot de God der methodologie' (Barendregt<sup>49</sup>) gaan aan verlossing vooraf. Het volgen van de voorschriften leidt tot afname van het aantal foute diagnoses. Het bestaat uit gedisciplineerd en geüniformeerd worden. Dit laatste wil niet zeggen dat ze dezelfde kleding dragen zoals militairen, politieagenten, parkeerwachters, toezichthouders & handhavers en medisch specialisten. Het gaat erom dat ze gelijkvormig denken en handelen bij dezelfde vragen en problemen.

*Het schenden van logische regels tegengaan:* Ze zijn al opgeschreven door Aristoteles en benut door Van Hoesel<sup>50</sup> in *Zindelijk Denken* (1955). De wetenschapsfilosoof Derksen<sup>51</sup> wijst op dergelijke fouten in het proces Lucia de B. Hij speelde een rol in de vrijlating van de van acht moorden beschuldigde verpleegkundige. Hij spreekt in zijn boek *De ware toedracht* over cognitieve handicaps, zoals het zien van oorzakelijke verbanden in willekeurige reeksen; het gedrag van iemand verklaren uit zijn karakter en drijfveren en de sociale context veronachtzamen; geen oog hebben voor toeval; denken in simpele tweedelingen terwijl er steeds nuances en glijdende schalen zijn.

*Correct argumenteren:* Toulmins<sup>52</sup> procedure stelt in staat om aan het eind van het stappenplan vast te stellen of er sprake is van een stand van zaken of niet, bijvoorbeeld de aan- of afwezigheid van een gedragskenmerk, probleem of ziekte. Er moet een grond zijn die verwijst naar feitelijke steun voor het argument. Als ik beweer dat ik morgen het tentamen Onderzoeksmethoden zal halen, mag je mij vragen om dat te verantwoorden. Ik moet gronden kunnen noemen, bijvoorbeeld: ik ben tot nu toe steeds geslaagd; ik heb me goed voorbereid; ik heb de colleges en practica gevolgd. Een grond is een controleerbare gebeurtenis. Er is (a) een claim: aanspraak op de geldigheid van een bewering over een stand van zaken, bijvoorbeeld docent Jan faalt, of Jan is een casus van een incompetente onderzoeker (b) gegevens nodig om de claim met de werkelijkheid te verbinden. Het is gerechtvaardigd om deze bewering te doen als Jan de kenmerken A, B, C, D, E en F vertoont: hij kijkt tentamens slordig en te laat na; hij houdt

zijn vakliteratuur niet bij; hij is nooit op zijn werkkamer; hij schrijft droog, hij krijgt slechte evaluaties van decaan en studenten, hij publiceert niet. Er moet (c) aangetoond worden dat de kenmerken voldoende zijn om van een incompetent docent/onderzoeker te spreken. Als dat niet lukt, is de claim van tafel. Als het wel lukt moet nieuwe, onafhankelijke informatie worden ingewonnen die opnieuw aantoont dat de kenmerken A tot en met F op Jan van toepassing zijn. Dit is *backing*: een ruggensteun creëren. Als dat weer lukt moet de claim genuanceerd (*qualified*) worden: bijvoorbeeld: 'hij is competent voor onderwijs aan Masterstudenten, maar kan eerstejaars niet boeien (*rebuttal*: weerlegging). Het is mogelijk dat de kenmerken A tot en met F na heroverweging niet volstaan om de conclusie te trekken dat Jan een casus van een incompetent docent/onderzoeker is.

In Groningen werd een tijd een vertaling van Perelman en Olbrechts-Tyteca's<sup>53</sup> retorica gebruikt. Ik vermoed dat het niet meer het geval is. Het boek is van 1958 en er is een heruitgave van 2009. Ik herinner me het bruine kaft. De tekst is als pdf te downloaden van het internet. Dit boek sloot aan bij de klassieke retorica van de Grieken. Het vormde een breuk met de Cartesiaanse opvatting van rede en redeneren dat zijn stempel op de westerse filosofie heeft gedrukt. Ik dacht dat Toulmin het niet zou halen in de psychologie. Het werk steunt op de klassieke retorica en dat is filosofie. Niettemin vond ik nog een referentie waaruit bleek dat de procedure aanbevolen werd om studenten te leren onderzoeksresultaten kritisch te evalueren. In de verpleegkundige diagnostiek wordt de werkwijze gebruikt. Brown et al.<sup>54</sup> bevelen de werkwijze aan voor studenten om hun redeneren over wetenschappelijk verklaringen aan te scherpen. Ze wordt verbonden met *evidence-based* redeneren en beslissen. Ik zit er overigens naast want bij het Cito wordt onderzoek verricht en worden ontwerpen gemaakt die gebaseerd zijn op Toulmins argumentatieleer. Deze draagt bij aan het evalueren van de validiteit van onderwijs en trainingen. Dit is winst omdat het validiteitsbegrip zowel verruimd, als bepaald, omschreven wordt. Er wordt een ontwerp gemaakt voor elke validiteitssoort. De dissertatie van Wools<sup>55</sup> vat deze aanpak samen en past hem toe op bijvoorbeeld het rijvaardigheidsonderwijs.

*De Multi-Attribuut-Utiliteits-Theorie omarmen*: Het gaat om het toepassen van regels om de winst zo groot mogelijk te maken. Er is een aantal keuzemogelijkheden, bijvoorbeeld in het geval van een ouder echtpaar: in een groot, mooi, oud huis blijven wonen, een nieuw appartement, een woning in een verzorgingscomplex. Er is een aantal attributen, relevant voor de keuze, nabijheid van de kinderen, bijvoorbeeld kosten, verhuizen, toegankelijkheid, dagelijkse verzorging op de drie plekken, in staat zijn tot zelfstandig reizen en persoonlijke vrijheid. De attributen krijgen een gewicht, bijvoorbeeld cijfers van 1 tot en met 10. Het gewicht drukt het belang van elk attribuut uit voor dit echtpaar. Als het stel rijk is, zijn kosten niet zo belangrijk. Het gewicht moet het belang van dat attribuut uitdrukken in vergelijking tot de - in dit geval - zes andere. Dat wordt geregeld door de som van de attributen een constante waarde te geven, bijvoorbeeld 1.0. Als één erg belangrijk is, bijvoorbeeld persoonlijke vrijheid, geeft men het een waarde van 0.40 en dan blijft er voor de overige zes 0.60 over. De gewichten worden vermenigvuldigd met de relatieve waarden en daarna opgeteld. De hoogste waarde

wordt gekozen als men rationeel is. Als de gewogen sommen op ongeveer dezelfde waarde uitkomen, helpt de onderneming niet. We gaan er ondertussen vanuit dat het kiezen van de attributen en de weging zo gepiept zijn en dat beiden het er over eens zijn.

*Bayes' hulp aanvaarden:* Zijn benadering kan toegepast worden op één persoon of een verschijnsel. De apriori kans dat een verschijnsel, gedrag voorkomt wordt een aposteriori kans nadat informatie is toegevoegd. De Bayesiaanse aanpak zegt dat de kans op het voorkomen van een verschijnsel (*base rate*, incidentie) en de kans dat het verschijnsel goed gediagnosticeerd wordt, gecorrigeerd moet worden voor de kans dat het verschijnsel voorkomt in de populatie en het niet dát, maar een ander verschijnsel is en het niet perfect gediagnosticeerd wordt. De uitkomst is soms tegen-intuïtief en bevat, als de regel geldt in de empirie - en daar is steun voor - een waarschuwing om niet te snel tot een categorie te besluiten. We zien bijvoorbeeld de incidentie van een verschijnsel over het hoofd; we denken met zekerheid de oorzaak van een verschijnsel te kunnen vaststellen en houden geen rekening met andere oorzaken. Om de regel te kunnen toepassen, moeten we over kennis beschikken: hoe vaak komt een verschijnsel voor en in welke mate en over hoeveel kennis van empirische verbanden hebben we. Verder moeten we in staat zijn het probleem zo te structureren dat de formules toegepast kunnen worden. Doorgaans is de kritiek op deze modellen en regels juist daarop gericht. De aanpak past bij diagnostiek maar wordt zelden toegepast.

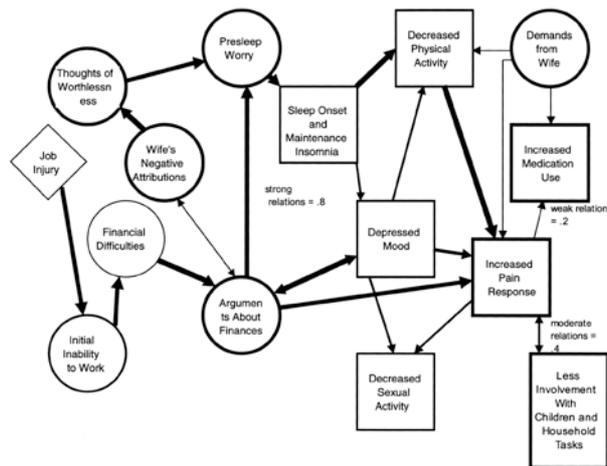
*Nog meer hulp:* Stapsgewijze procedures (i.c. stroomdiagrammen) voorkomen dat onderdelen worden overgeslagen. Elke keten is zo sterk als de zwakste schakel. Weinstein en Fineberg<sup>56</sup> stellen een aantal stappen voor. Na elke stap volgt een beslissing. De stappen structureren het probleem. Bij elke stap hoort een antwoord op de vraag of, en in welke mate deze bijdraagt tot een oplossing. Dit voorstel gaat in tegen *fast* en *frugal* heuristieken en tegen de tendens van experts om het eerste het beste idee te beginnen.

De *Guidelines for the Assessment Process*<sup>57</sup> bevatten zowel de diagnose als de behandeling. Westhoff et al.<sup>58</sup> maakten een uitvoerig stappenplan. De twee laatste procedures gebruiken beslissingsanalyse, statistische voorschriften voor optimale predictie en het HTM als bronnen. De lijst van Westhoff et al. bevat 36 stappen waarin 275 vragen beantwoord moeten worden. Het is een *gründliche* procedure die ontworpen is voor het in Duitsland benodigde *Gutachten* voor instellingen en de rechterlijke macht.

*Het Hypothese Toetsend Model (HTM)* is gebaseerd op de empirische onderzoekscyclus. Het is een combinatie van de denkpsychologie van Otto Selz (Württembergse school, begin 20<sup>ste</sup> eeuw; en Amsterdam) en Poppers idee van falsificatie (1928/2006). Beiden zorgden er voor dat probleemoplossen een epistemologisch thema werd. De Groot liet zich inspireren door Selz voor zijn analyse van het denken van de schaker (1946) en benut de denkpsychologie ook in zijn *Methodologie*. In zijn onderzoekscyclus speelt onder meer Poppers falsificatie argument een rol. De cyclus is gebruikt als protocol om de diagnostische activiteiten te sturen: de diagnosticus als *scientist-practitioner*. Het HTM neemt het hypothetico-deductieve model als uitgangspunt en maakt diagnostiek tot het beantwoorden van een onderzoeksvraag.

Ik heb eerder beweerd dat de diagnosticus niet als onderzoeker te werk gaat. Er is een vraag en er is informatie die geordend en getoetst moet worden om de oorzaak/reden van het probleem op te sporen. In het HTM worden theoretische analyse en *evidence-based* kennis geïntegreerd en hypothesen over de oorzaken getoetst. Hypothesen toetsen is het schatten van een populatieparameter (Kirk<sup>59</sup>): de nulhypothese dat er geen verschil is tussen twee gemiddelden en ze daarom uit één populatie komen. De alternatieve hypothese zegt dat de gemiddelden *niet* uit één populatie komen.  $H_0: \mu_1 - \mu_2 = 0$  en  $H_1: \mu_1 - \mu_2 \neq 0$ , tweezijdig getoetst. Steekproeven worden getrokken om de parameters te schatten. Het is niet duidelijk hoe je een nulhypothese toetst bij één cliënt. Er is geen steekproef van scores van elementen uit een populatie. Uiteraard kan er beschrijvende statistiek gedaan worden met herhaalde metingen bij één persoon. Dat doet zich in de diagnostiek echter vrijwel nooit voor. In het HTM moet het toetsen van hypothesen iets anders betekenen. Geobserveerde scores worden bijvoorbeeld vergeleken met een criteriumwaarde of er wordt vastgesteld of een gedrag aan- of afwezig is.

*Casus*: Analyses van casus waren gewoon in het begin van de 20ste eeuw en werden ook gepubliceerd. Dit verdween maar lijken in de functionele gedragsanalyse terug te keren. De vraag wordt uitgewerkt in een *case formulation*: ...it can be defined as the process of operationalizing target behaviors (determining the form) and evaluating relationships among target behaviors and potential controlling factors (determining function) for an individual client (O'Brien et al.<sup>60</sup>). Veel interacterende factoren zijn werkzaam. Dit leidt tot een functioneel analytisch causaal model voor een cliënt. Een voorbeeld is een gevalformulering voor meneer X door Haynes en Williams (2003) in Figuur 4.



Figuur 4: Casus van het probleem van meneer X. 'Case formulation and the design of behavioral treatment programs'. Ontleend aan: S.N. Haynes & A.E. Williams. *European Journal of Psychological Assessment* 2003; 19, 3, 164-175 (met toestemming van de uitgever Hogrefe).

De causale structuur is opgezet na interviews met meneer X en zijn vrouw. Er zijn vragenlijsten afgenomen, observaties verricht, ook door de cliënt zelf (*self-monitoring*) en er is een werkbespreking met een team geweest. De rangorde van de causale variabelen is: 1) ruzie over financiën 2) financiële problemen 3) eisen van echtgenote 4) moeite met in slaap komen 5) negatieve attributies door echtgenote; 6) gedachten over waardeloosheid; 7) gevoel van het werk niet meer aan kunnen. De werkwijze is tijdrovend en ondergedetermineerd. Dat is het probleem van iedere *rich and sufficient* beschrijving van een cliënt. Hoewel de casus schatplichtig is aan de functionele gedragsanalyse, bevat hij hermeneutische stappen.

De Verlossing bestaat uit een parade van protocollen, modellen en richtlijnen. Bewijs voor de kwaliteit van de protocollen is verbetering van het product: meer juiste en minder onjuiste diagnoses. Er is nauwelijks empirisch onderzoek waarin de voorgeschreven werkwijzen onderling en met spontane diagnostiek vergeleken worden. Zou de uitslag niet dezelfde kunnen zijn als die bij de vergelijking van klinische versus statistische predictie? Zou het antwoord niet luiden: het is aan te bevelen één van de disciplineringsmethoden te volgen, maar in iets minder dan de helft is een eigen, spontaan gefabriceerde strategie gelijkwaardig? Is er niet veel vertrouwen in een stelsel met regels? Houden practici zich er strikt aan?

**Stadium 3 Dankbaarheid:** Een beetje dankbaarheid is er wel want wie maakt er niet graag gebruik van steun en sturing bij de lastige activiteit die diagnosticeren is? En, je hoeft niet steeds je werkwijze uit te leggen en te verantwoorden aan de supervisor. Diagnostici volgen de regels waarschijnlijk niet strikt maar we weten het niet echt, omdat er weinig onderzoek is dat leert hoe de diagnosticus met de modellen omgaat. Zou het niet zo gaan: een beetje van het model en een beetje van mij, de diagnosticus? Er is wel nagegaan of het HTM gevolgd wordt. Dat blijkt slechts voor een deel het geval. Het diagnostisch proces is niet te vergelijken met dat in chemische procestechnologie. Een proces is een rationele (re)constructie met een vaste architectuur. Werkelijke processen zijn chaotisch. Ze verlopen niet lineair. Een protocol maakt kans om volgens de regels uitgevoerd te worden als het aansluit bij de natuurlijke werkwijze van de diagnosticus en zijn informatiesysteem niet overbelast. Is dat ooit nagegaan voor al deze modellen? Ik ken geen voorbeelden.

**Uniformeren door protocollen** Als diagnosticeren gedrag is gelden de drie oriëntaties van thema I, sectie 1. De mechanische, klerk opvatting ziet de diagnosticus als inwisselbaar. Lukt dat met behulp van een protocol? Gaat de diagnosticus met de cliënt om zoals een operator achter de computer; zoals een procestechnoloog in het laboratorium? Is hij een klerk (Cronbach), kan hij door een robot vervangen worden (Linschoten)? Een protocol heeft als doel uitvoerders zo te sturen dat ze bij gelijke vragen van cliënten tot dezelfde diagnose komen. Diagnosticeren is zich gedragen en gedrag is gekenmerkt door individuele verschillen, ontwikkeling en invloed van de sociale context. Het lijkt zinvol om uit te zoeken of deze kenmerken zich feitelijk voordoen in plaats van erop te vertrouwen, dat bijvoorbeeld het HTM zo sturend en doorzichtig is dat elke diagnosticus het als bedoeld ten uitvoer legt.

*Diagnosticci verschillen individueel:* Accuraatheidsonderzoek bij persoonswaarneming kan toegepast worden op individuele verschillen tussen diagnostici. Dit thema was een tijd niet populair. Het oordeel en de waarneming van een persoon werden gezien als een product van kenmerken van de oordelaar en niet van de gediagnosticeerde (Cronbach et al.<sup>61</sup>; Funder & West<sup>62</sup>). Een substantieel deel van de variantie in oordelen was gebonden aan variantie in kenmerken en belangen van de oordelaars. Als deze verschillen echter aanvaard worden, mag je vragen om een goede diagnosticus. *Person Perception* studies laten zien dat er verschillen zijn. John en Robbins<sup>63</sup> toonden bijvoorbeeld aan dat naarmate personen hoger scoorden op een narcismeschaal minder accurate oordelen over anderen en henzelf gaven. De authentieke behoefte om anderen te leren kennen (*NC: need for cognition*) verwijst naar het plezier dat mensen ervaren in het - niet evaluatief - denken over en analyseren van gedrag van anderen. Personen met een hoge NC-score: ... *occasionally give more normative responses than their lower NC-counterparts (especially in within-subjects contexts)*, maar ook: ...*increased thought as indexed by NC-scores do not appear to rid respondents of observed inconsistencies and biases*. Normatieve antwoorden zijn gepaste, sociaal correcte en wenselijke antwoorden, maar de hogere NC score verlost ze niet van de bekende vertekeningen van het oordeel. Letzring et al.<sup>64</sup> correleerden persoonskenmerken van oordelaars met accuraatheid. Naast vragenlijsten gebruikten ze video-opnamen van interacties tussen oordelaar en beoordeelde. De goede oordelaars toonden hogere scores op sociale vaardigheden en op Aangenaamheid van de Big Five. Ze vertoonden zelf ook gedrag dat overeenstemde met hun score op de schalen. Ze waren geïnteresseerd in hun interactiepartners en deze waren op hun beurt open ten opzichte van deze *judges*.

*Diagnosticci ontwikkelen zich* tijdens het onderzoek (microgenese) en hun professionele loopbaan (ontogenese). Microgenese was een tijd uit de gratie van onderzoekers. Siegler<sup>65</sup> introduceerde het opnieuw bij het onderzoek naar hoe kinderen het balansschaalprobleem oplossen. Om het micro-proces te bestuderen kun je uitzoeken of en hoe diagnostici hun hypothesen veranderen gedurende de rit, hoe ze beslissen over de instrumenten en hoe hun contact met de cliënt wijzigt. Dit vraagt observeren van het proces en het ondervragen van practici over hun werkwijze. In rapporten staat regelmatig dat er contactgroei tussen diagnosticus en cliënt is.

Longitudinaal ontogenetisch onderzoek kan aan het licht brengen hoe en of diagnostici theorievast zijn en of ze zich aanpassen aan de vernieuwing van procedures, methoden en instrumenten. Barnes en Moon<sup>66</sup> beschrijven de ontwikkeling van therapeuten gedurende supervisiesessies. Mogelijk geldt dit ook voor diagnostici. Door de tijd heen worden ze efficiënter, leren ze van supervisie en zijn ze bereid en in staat om supervisie te ontvangen en te geven. Een professionele identiteit ontwikkelt zich. De auteurs stellen vier stadia voor: *Role shock*: worstelen om een professionele identiteit te omschrijven, wat is het vak? *Role recovery and transition*: een realistische kijk op de ontwikkeling en het vormen van een identiteit als diagnosticus. *Role consolidation*: zich bekwaam achten en voelen. *Role mastery*: de overtuiging krijgen dat men stabiel, consistent en vakbekwaam is en verantwoordelijk taken uitvoert en

die kan overbrengen aan anderen. Diagnostici ontwikkelen zich en dat kan blijken uit een toename van accuraatheid, zelfvertrouwen en bereidheid resultaten met collega's en supervisors te checken. Het gaat niet altijd goed. Ontwikkeling kan immers ook stagneren.

*Diagnostici verkeren in een sociale context:* De institutionele en sociale omgeving beïnvloedt persoon en werk van de diagnosticus. We veronderstellen dat supervisie nuttig is maar er is geen kwantitatief onderzoek dat het effect daarvan in kaart gebracht heeft. Supervisie kan de identiteit van de diagnosticus versterken, maar verhoogt het ook de kwaliteit? Wordt er geprofiteerd van kennis van collega's of is de hiërarchie bepalend voor wat beweerd wordt over cliënten? Zijn expertmodellen winst voor de diagnosticus? Wat is de rol van de institutie waar de diagnosticus werkt? Wat levert bijscholing op naast de punten voor de BIG (her)registratie (Beroepen in de Individuele Gezondheidszorg)? Er zijn weinig studies/essays waarin diagnostici hun *Werdegang* beschrijven. Geen essays met als titel: 'Wat heb ik geleerd van zoveel jaren diagnosticeren en behandelen' en 'Wat heb ik goddank niet geleerd', 'Wat heb ik afgeleerd'. Welke valkuilen ben ik tegengekomen'? Studenten vragen dit zelden aan docenten, stagebegeleiders en practici. Doen!

*Tot slot:* Diagnosticeren is gedrag en dat opereert tussen *logos*, wetenschap, verantwoordende, theorie, modellen, procedures van het HTM, MAUT, Bayes-regels en logisch redeneren en *mythos*: mythes, alledaagse opvattingen, intuïtie en ervaringskennis van communicatie en conversatie met de cliënt, het maken van een verhaal, het subjectief aanvoelen en het construeren van betekenis. Logos en mythos zijn in een kritische dialoog verweerd. Ze zorgen samen voor dynamiek in de studie van het diagnostisch proces, als de logos: de Wetenschap, het Model, het Streng Methodologisch Geweten de waarheid niet claimen.

## 6. De weg van een gedragsconstruct, bijvoorbeeld van de Big Five

Eerder is een zinvol aggregatieniveau voor diagnostiek gesuggereerd. Het werd niet atomistisch gezocht bij het DNA, EVP, RT, genen, maar bij molaire eenheden: gedrags- en situatiebegrippen. We zijn misschien niet verder dan Aristoteles' definitie van een begrip. Via de Scholastiek luidt dat: *Ens Rationis cum fundamento in Re*: een Zijnde van het Verstand met een basis, fundament in de Werkelijkheid. In de uitwerking zijn we een stuk opgeschoten door validiteitsonderzoek met onze intelligentie- en persoons- en situatieconstructen. Met deze eenheden moet de diagnosticus het doen. Constructen zijn meestal op intervalniveau gemeten individuele verschillen op een aantal dimensies van typerend en cognitief gedrag en van situaties. Er zijn daarnaast categorieën, vooral persoonstypologieën. Voorbeelden zijn hechtingscategorieën en de DSM-IV en 5.

De 'uitstraling' van de Big Five (BF) wordt nagegaan in construct- en predictief validiteitsonderzoek. Experimenteel onderzoek biedt inzicht in de mogelijkheid tot controle/verandering van dat gedrag door trainingen en interventies. Beide leveren in meta-studies bescheiden *r*- en *d*-waarden met aanzienlijke variantie op. De verklaring, interpretatie van verschillen biedt de parade van biochemie tot sociologie en culturele antropologie. Is dat te zien bij de BF? Wat is de weg die een populair construct, in dit geval vijf trekken (*traits*), moet gaan? U bent gewaarschuwd, de BF bevindingen zijn omvangrijk en de resultaten bescheiden. De referenties voor BF studies zijn in iedere overzichtsstudie te vinden. Ze zijn op een enkele uitspringende na hier niet in de referenties opgenomen.

*De Individuele Verschillen benadering domineert*: In persoonsleer en intelligentieonderzoek overheerst de individuele verschillen benadering die van ontwikkeling en sociale context. Daar is iets voor te zeggen. We nemen objecten waar als ze verschillen van de achtergrond. De wiskundige Brouwer (Heyting<sup>67</sup>) verdedigde het verschilprincipe: 'iets is verschillend, of het is niet', dat wil zeggen we merken het niet op. Constructen beschrijven die verschillen in typerend en cognitief gedrag en in situaties met behulp van categorieën, bijvoorbeeld de typologie van de vier lichaamsvochten (Hypocrates) en de 39 karakters (Theophrastes). Van recenter datum zijn de hechtingstypologie en de DSM-IV en 5.

Een wrang voorbeeld van een typologie is het beoordelingssysteem van de Nederlandse Zorgautoriteit (NZa). Het is ontleend aan Gotlieb<sup>68</sup> een ambtenaar die zelfmoord pleegde na tegenwerking en onvoldoende beoordelingen. Deze zijn postuum ongeldig verklaard. Het is een kwadrant met twee kenmerken die beoordeeld worden door een personeelschef. Er vindt geen betrouwbaarheidsonderzoek plaats en er valt de *facto* niet - de *jure* of *soms*, na een zelfmoord wel - tegen in te gaan. De dichotome kenmerken van de werknemer zijn: Toegevoegde Waarde (TW 0,1) en GroeiPotentieel (GP 0,1). Dit resulteert in *learners* (TW 0, GP 1), *werkbijen* of *bakens* (TW 1, GP 0), *high potentials*: *hipo's* (TW 1, GP 1) en *achterblijvers* (TW 0, GP 0). De laatsten werden het NZa jargon *payroll bevuilers* genoemd.

Psychologen beschrijven - in tegenstelling tot medici - de verschillen overwegend met behulp van schalen, dimensies, latente trekken. Persoonsconstructen zijn meestal samengesteld,

denk aan de *Vectors of Mind* van Thurstone<sup>69</sup>: de *Primary Mental Abilities*. Je kunt analoog spreken over de *Vectors of Personality*.

*Geschiedenis van 'traits'*: Hartshorne en May<sup>70</sup> onderzochten in 1928 bedrog en eerlijkheid: *Studies in Deceit*. Ze zagen er geen transsituatieve trek in maar een situatief bepaald kenmerk. Het was vlak voor de crisis, de armoede was groot en de situatie maakte de dief. In de jaren 50 werden lekenoordelen over personen (*person perception*) als vertekend, *gebiased* beschouwd. Belangen spelen een rol in de waarneming wie en wat een ander is. In 1968 schreef Mischel<sup>71</sup> een boek waarin hij liet zien dat trekcores weinig variantie in criteriumgedrag voorspelden. De slagzin *traits don't exist* laat niets aan duidelijkheid te wensen over. In de tussentijd deden twee Amerikaanse luchtmachtpsychologen, Tupes & Cristal<sup>72</sup>, onderzoek met 35 *representative traits* bij acht steekproeven. Hun rapport verscheen in 1961, maar kreeg geen aandacht. Mischels boek van 1968 trok die naar zich toe. In 1992 werd het nog eens als artikel gepubliceerd. Hun benaming van de vijf factoren luidde: *Surgency, Agreeableness, Dependability, Emotional Stability, Culture*. Evenals bij intelligentietests plaveide het Amerikaanse leger de weg voor een nieuwe start van het diagnosticeren van persoonlijkheid. John<sup>73</sup>, een kiene samenvatter van wat er in de persoonsleer te koop is, zag aankomen dat BF centraal zouden komen te staan. Hij bracht in een hoofdstuk in een persoonsleerhandboek alle trektheorieën onder in het BF frame (1990). Transsituatieve trekken werden weer aanvaard. En intussen hadden Costa en McCrae<sup>74</sup> de BF geconfisqueerd. Hun vragenlijst de NEO-PI van 1985 wordt bijna over de hele wereld gebruikt. Het instrument is gemakkelijk af te nemen en de uitslag eenvoudig te interpreteren. Het is eenvoudig om er construct en predictief validiteitsonderzoek mee te doen.

Onderzoekers bedenken verklaringen en geven interpretaties voor *aangetroffen* verschillen. Experimenteel onderzoek zou ingezet kunnen worden om de persoonlijkheid te *veranderen*. Dat zie je weinig. Het instrument is er niet op gemaakt, maar dat deert veranderingscoaches uiteraard niet. We nemen de *Big Five* (BF) als voorbeeld. Ze gaan al een tijdje mee maar waren er niet altijd. Een construct kan een tijd populair zijn en verdwijnen, bijvoorbeeld Witkins<sup>75</sup> *Field Dependency*. Het startte als onderzoek naar welke factoren een rol spelen bij het vaststellen of een lijn loodrecht of scheef wordt waargenomen in een frame. De lijn was fysisch 90 graden, loodrecht dus. Manipulatie van het frame had effect, maar er bleken individuele verschillen. Dat leverde een massa correlaties op van de waarneming van loodrecht met persoonlijkheidskenmerken en intelligentie. De strekking was hoe onafhankelijker je van de context (het frame) bent, des te efficiënter pak je problemen aan en lost ze op.

Dit zal met de BF niet gebeuren want er zit geen specifiek meetapparaat aan vast met één getal als uitkomst. De BF biedt een overzichtelijk profiel met scores op vijf trekken. Je ziet meteen hoe de scores op de Vijf onderling verschillen en of ze afwijken/samenvallen met de score van de gemiddelde persoon.

Ik stel *vragen* over construct- en predictieve validiteit van de BF en de gangbare interpretaties van de factoren worden vermeld. Het is veel van hetzelfde. U bent gewaarschuwd. U kunt de

tekst tussen de cursief gedrukte 'vragen' en 'antwoorden' overslaan, als U de berg correlaties, die onderzoeksliteratuur over ons uitgestort heeft niet interessant vindt. In dat geval moet U mijn antwoord vertrouwen. Het luidt kort door de bocht: het aantal factoren verschilt tussen de 1 en 16. De Vijf zijn met van alles en nog wat gecorreleerd. De construct- en predictieve validiteitscoëfficiënten zijn bescheiden tot een enkele keer gemiddeld. De interpretatie van de oorzaken varieert van statistisch artefact, semantisch label tot genetische bepaaldheid. Als we met Kouwer serieus nemen dat de persoon terugpraat, kun je dit met de BF uitproberen. Stemt hij in met het resultaat op de BF? En zo niet, wat zegt u dan?

Eenzelfde verhaal kun je vertellen over intelligentie en cognitie. Het aantal factoren varieert tussen 1 en 150. De samenhangen met criteria zijn hoger dan bij de BF en de verklaringen zijn ruwweg dezelfde. Theorievorming en onderzoek zijn daar verder gevorderd door de verbinding met waarnemings- en geheugenonderzoek en kennis over informatieverwerking.

*Vraag: Wat zijn de BF?* Een factoranalytisch geconstrueerd profiel van persoonstermen. De constructen kenden voorlopers in het werk van Galton en Thurstone. Ze kregen een impuls door Allport en Odbert<sup>76</sup> in een studie uit 1936. Zij verzamelden ongeveer 18.000 woorden uit de *Oxford English Dictionary* die naar een persoon zouden kunnen verwijzen. Cattell<sup>77</sup> koos daaruit 4000 termen en reduceerde die tot 171 clusters. Na het berekenen van de correlaties in een steekproef besloot hij tot 35 clusters: de basis van zijn *16 Personality Factors* (16PF). Cattell leverde er een interpretatie bij. Zijn lexicale hypothese houdt in dat alle kenmerken van de persoon die de tand des tijds doorstaan hebben in ons lexicon zijn blijven hangen als bijvoeglijke en zelfstandige naamwoorden die personen beschrijven.

Een woordenboek verschilt van cultuur tot cultuur. Het bevat ongeveer vijf miljoen woorden met 50.000 lemma's die naar personen verwijzen, 5000 persoonsadjectieven waarvan er 500 regelmatig gebruikt worden en 50 daarvan kunnen in facetten (5 x 10) bij de Grote Vijf geplaatst worden. Hofstee et al.<sup>78</sup> gebruikten 5960 termen uit het Nederlands woordenboek. Zij vroegen zich bij elke term af of het adjectief, de term paste in de zin: 'Dit is een ... persoon', of als bijvoeglijk naamwoord in: 'een ... persoon'. Dat leverde 1203 termen op die op basis van waargenomen overeenkomst geordend werden in 42 categorieën. Principale Componenten Analyse (PCA) op elke categorie leidde tot 143 factoren met eigenwaarden > 1.00. Vervolgens werden items geselecteerd met een lading > 0.40. Dat resulteerde in 96 groepen met drie bijvoeglijke naamwoorden. PCA op deze groepen leidde in 1981 tot zeven factoren. Dat zijn de bekende Grote Vijf aangevuld met VI: Progressiviteit en VII: Agressie.

*Antwoord: De Vijf zijn het kleinste gemene veelvoud van persoonsadjectieven uit het woordenboek.* De Big Five hebben de pretentie alles te omvatten wat over een persoon gezegd wordt. De vijf letters vormen het anagram OCEAN: wat een diepte en breedte! Hieronder de BF. In de eerste kolom zijn de nederlandse en engelse factornamen vermeld. In de tweede staan de hoogst ladende adjectieven op die factor.

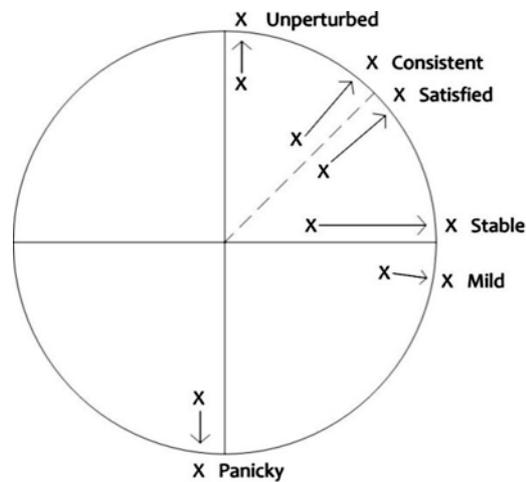
I	E: Extraversie (extraversion/ surgency)	energiek, enthousiast, rustig, gereserveerd, verlegen, spraakzaam, assertief, enzovoort
II	A: Aangenaamheid (agreeableness)	altruïstisch, affectief, koud, onvriendelijk, sympathiek, empathisch, warm, enzovoort
III	C: Gewetensvolheid (conscientiousness, dependability, orderliness)	verantwoordelijk, methodisch, precies, chaotisch, effectief, planmatig, enzovoort
IV	N/ES: Neuroticisme/ Emotionele Stabiliteit (neuroticism, emotional stability)	Zenuwachtig/ stabiel, kalm, onemotioneel, enzovoort
V	O: Openheid/ Autonomie (openness, autonomy, culture)	origineel, geïnteresseerd, open, eenvoudig van geest, intelligent, nieuwsgierig, enzovoort

*Vraag: Wat is het aantal factoren en welke labels wordt gebruikt? Vijf natuurlijk, op de vingers van een hand te tellen. Toch varieert het aantal van één: de General Personality Factor (GPF) tot de zestien van de 16PF. De eerste, sterke factor wordt gelezen als Stabiliteit/Conformisme of Sociale Wenselijkheid (SW). Meestal varieert het aantal tussen vijf en acht. De Raad en Barelds (2008) voegden drie aan de Vijf toe: VI Virtue, VII: Competence en VIII: Hedonism. Er is een zeven-factor oplossing: van Ashton et al. (2004). Ze voegen VI: Honesty-Humility en VII: Religiosity toe: aardige kenmerken voor de mensen. De GPF kwam in beeld naar analogie van Spearman's G-factor. Een twee-factor oplossing komt regelmatig voor: we houden van tweedelingen: als alfa óf bèta ben je duidelijk; niet vreemd dat gamma's het lastig hebben. Van der Linden et al. (2010) komen op één meta-factor uit maar achten een twee meta-factoren oplossing acceptabel: I Stabiliteit: de alfa factor en II Plasticiteit: de bèta factor.*

*Antwoord: Tussen de 1 en 16. Vijf is wel handig voor het gebruik en onze voorkeur voor tweedelingen is bij de BF merkbaar. Verklaringen: van biochemie tot culturele factoren.*

*Vraag: Wat is de relatie tussen de factoren? Orthogonaal natuurlijk: het profiel van vijf onafhankelijke factoren. Dat kan getoetst worden. Je zou verwachten dat Costa en McCrae<sup>79</sup> er als de kippen bij waren om dat te doen. Het kwam er eerst van in 1996. Het resultaat kwam hen niet goed uit (p. 522): *Despite the empirical robustness of the 5-factor model of personality recent confirmatory factor analyses (CFA's) of the NEO-PI data suggests that they do not fit the hypothesized model. In a replication study of 229 adults, a series of CFA'S showed that the Revised NEO-PI scales are not simple-structured but do approximate the normative 5-factor structure. CFA goodness- of- fit indices, however, were not high.* Ze benadrukken de juistheid van de *simple structure* van de Vijf. Het gebrek aan fit (p. 563): *...points to serious problems with CFA itself when used to examine personality structure.* Zo doe je dat als Big Five monopolist. Hofstee et al. <sup>80</sup> *aanvaardden* het empirische feit dat items bij varimax rotatie op twee en zelfs meer factoren laden. Als judoka's gingen ze met de beweging van de feiten mee en*

konden de persoon er gedifferentieerder mee beschrijven. Ze namen steeds combinaties van twee factoren uit de Vijf en bekeken welke items op de twee redelijk hoog laadden. Elk paar dimensies (bijvoorbeeld I en II, I en III, I en IV) leveren een circumplex op die is opgebouwd uit de twee polen van elke dimensie. In de cirkel kan men segmenten maken met de items. Zie Figuur 5.



Figuur 5: Een voorbeeld van een circumplex van twee van de Vijf factoren. De items kunnen in de cirkel geplaatst worden volgens de ladingen op de twee dimensies (ontleend aan De Raad, met zijn goedvinden, 1995, p. 96).

Een voorbeeld van Hofstee et al.: Stel, u zoekt een manager die stressbestendig is. We gaan uit van de circumplex met II (Aangenaamheid: + is hoog en – is laag) en IV (Emotionele Stabieleit: + hoog en – laag). Wilt u een geduldige manager (II+ en IV+), of een ongeduldige die opschiet (II– en IV–), of een stabiele (IV+ en II+) of prikkelbare manager (IV– en III–), of een nuchtere (IV+), enz. De auteurs hebben de cirkel verdeeld in twaalf segmenten van 30 graden en dat leidt tot de Abridged 5 Factor Circumplex (AB5C). Zo zie je, wat de een afwijkt ('McCrosta') kan bij een ander tot een genuanceerder interpretatie leiden.

*Antwoord: de Vijf Correleren tussen de .10 en .40, maar worden als orthogonaal behandeld worden in het profiel.* Als de correlaties tussen factoren serieus genomen worden, stelt dat in staat tot een genuanceerdere beschrijving van een persoon dan met het profiel van vijf onafhankelijke factoren.

*Vraag: Vind je de BF terug in andere culturen?* Costa en McCrae beschouwen de BF als even duidelijk en onderscheiden als de zeven continenten. Dus moet je ze overal vinden. Onderzoekers zijn eropuit om hetzelfde aantal en dezelfde structuur van de Vijf terug te vinden. Katigbak et al. (1996) rapporteerden de Filippijnse *indigenous* persoonsdimensies: '*responsibility, development of social abilities, emotional control*'. Hoewel dit verschilt van de Vijf, blijven de auteurs in het spoor van de BF, want er zijn ... *slight cultural differences in the constructs or*

*in the relevance of some items for the Philippines and the U.S. sample...* (p. 102). De Procrustes techniek van factorrotatie werd gebruikt om de Filipijnse dimensies in de Amerikaanse Vijf te dwingen. Katigbak et al. (2002) bevestigden dat nog eens in een steekproef van 508 studenten. Ze vermelden nu een klein verschil in de inhoud van de factoren en voegen toe: *social inquisitive behavior, including gossip, risk taking*: een facet van Extraversie, *Confidence to pursue new challenges* en *Religiosity* (p. 98).

En de Russen in 1993, net na de kortdurende perestrojka? Smelyov en Pokhil'ko vonden in 2029 Russische persoonswoorden zes factoren. Aan de Grote Vijf werd een zesde factor toegevoegd met een grillig ladingenpatroon: A-, E+ en -, O- en N+. Ze lazen de zesde als een Russische trek: uitdrukking van hun *sadness about facts of life*. Er moest wel iets Russisch te herkennen zijn en dat is niet het Amerikaanse optimisme.

En de Chinezen, nog maar net bezig een kapitalistische samenleving te maken? Cheung et al. (2001) legden ook de nadruk op de gelijkheid van de factoren in een Chinese en Amerikaanse steekproef. Ze vonden zeven factoren en interpreteerden ze als de Vijf. Een extra factor werd geïnterpreteerd als typerend voor de Chinese cultuur: *Interpersonal relatedness*. Cheung et al. (2003) vonden in Azië trekken die op verbondenheid met anderen en het gebrek aan koersvastheid van het zelf wezen. Zhou et al. (2009) vermelden als voor China kenmerkende factoren: *unselfishness, emotional volatility, dependency/fragility en negative valence*.

En de Europeanen? Hendriks et al. (2003) staan wisselende identiteiten toe en benadrukken minder de stabiliteit en koersvastheid van een BF profiel. Benet-Martinez en John (2000) vonden *Los Cinco Grandes* in een Noord-Spaanse steekproef terug. Ze merkten op dat het Spaans-Castiliaans dialect meer bijvoeglijke naamwoorden kende voor humor, socio-emotionele betrokkenheid, onconventioneel handelen en spiritualiteit. De Fruyt et al. (2009) verzamelden in 24 culturen met hulp van studenten *ratings* van gedrag van 5109 adolescenten van 12 tot 17 jaar. Ze bevestigden de structuur van de Vijf en voegden geen Belgische toe, terwijl de Belg toch 'zijn eigen plan trekt'. Misschien doet hij dat op deze manier: doe als Amerikanen succes gegarandeerd. Een studie van McCrae en Terraciano (2005) in vijftig culturen vermeldt eenzelfde conclusie: de factoranalyses in de culturen toonden *the normative American self-report structure*.

*Antwoord: Studies uit landen van België tot de Filippijnen suggereren dat we op Amerikaanse studenten lijken. We voegen iets van onszelf toe: 'iets van Maggi en iets van onszelf'.*

*Een nuance:* Eerlijkheid gebiedt te melden dat er ook terughoudendheid is over de cultuur-overstijgende kwaliteiten van de BF. Heine en Buchtel<sup>81</sup> wijzen erop dat de steekproeven vooral uit Amerikaanse studenten bestaan. Zij nemen verschillen tussen de VS en andere culturen waar in: *approach-avoidance, internal-external frames of reference, seeing one-self as independent or dependent, the possibility of personal growth and relational mobility*. Ook binnen de VS zijn er verschillen gevonden die verbonden zijn met sociale indicatoren. Een voorbeeld: de steun voor presidentskandidaten verschilt per staat. *State-level* persoonskenmerken dat wil zeggen trekken die situatiegevoelig en bijgevolg instabieler zijn dan traits, waren gecorreleerd met stemgedrag: het percentage stemmen voor Democratische kandidaten was positief gecorreleerd met *state-level* Openheid en negatief met *state-level* Gewetensvolheid ( $r = .64$  en  $r = -$ .

37;  $p$  uiteraard  $< .05$ ). Deze regionale verschillen maken duidelijk dat je iemand geen typische New Yorker, Amsterdammer, Utrechter, Groninger, Tilburger, Maastrichtenaar, Nijmegenaar, of een typische katholiek, protestant, hindoe, soenniet, wahaabiet kunt noemen. De verschillen binnen deze groepen zijn immers even groot als die tussen deze groepen. Dé Nederlander bestaat niet volgens Maxima, Koningin der Nederlanden. Ze heeft gelijk.

McCrae et al.<sup>82</sup> doen mee in dit koor, maar net iets anders. Samen met 37 auteurs uit 26 landen (alle continenten, behalve Afrika) vroegen ze aan 3323 deelnemers om kenmerken te noemen van een typische adolescent, volwassene en oudere (m/v). Ze gebruikten de NEO-PI, metingen verkregen in een *Adolescent Personality Profiles of Cultures Project*, gegevens van een *National Character Survey*, van data van De Fruyt et al. (2009) en McCrae en Terraciano (2005). Naast de BF (NEO-PI) werden 30 persoonskenmerken beoordeeld van een typische adolescent, volwassene en oudere. De deelnemers toonden *shared beliefs* over de kenmerken van de drie groepen. De auteurs interpreteren ze als stereotypische sekse- en leeftijdsverschillen. De accuraatheid van de oordelen werd bepaald door de stereotypen te vergelijken met de 'waarheid': de NEO-PI scores. De correlaties waren significant, maar verwaarloosbaar tot laag ( $r$ -waarden tussen de .01 en .20). De accuraatheid van de *shared beliefs appeared to be extremely limited* (p. 839). Terzijde: de terughoudendheid om landen en culturen te vergelijken treft men minder aan bij vergelijking van IQ scores in landen, denk aan het werk van Jensen en Rushton<sup>83</sup> en het boek van Lynn<sup>84</sup> over *The chosen people*.

*Vraag: Hoe doen BF het bij de concurrentie; hoe is de relatie met andere constructen die ongeveer hetzelfde of uitdrukkelijk iets verschillends pretenderen te meten: constructvaliditeit?* De Vijf zijn gecorreleerd met de 16PF van Cattell: Rossier et al. (2004) voerden een hiërarchische factoranalyse uit en concluderen dat de twee instrumenten ongeveer dezelfde factoren bevatten. De tweede laag van de 16PF moet wel samenhangen met de BF want Costa en McCrae gebruikten items van Cattell. Aangenaamheid ontbrak in de 16PF. Eysenck ontwierp een hiërarchische persoonlijkheidstheorie met Psychoticisme, Extraversie en Neuroticisme als *supertraits*. De *Eysenck Personality Questionnaire* (EPQ) is opgebouwd uit 21 primaire trekken. Ng et al. (1998) vatte de EPQ resultaten tot 1997 samen en deden een nieuwe studie met adolescenten uit Hong Kong. De Big Five kregen de voorkeur als beschrijving van de data boven Eysencks *Giant Three*. Er was enig verschil: Aangenaamheid ontbrak in de EPQ en Extraversie ging gepaard met de facetten Sociabiliteit en Opwinding zoeken. Impulsiviteit werd niet gedekt door Psychoticisme. Woods en Hardy (2012) zochten, vonden en interpreterden een hogere-orde structuur voor vijf vragenlijsten: *Hogan Personality Inventory*, *Occupational Personality Questionnaire*, *16PF*, *Personality Preference Inventory* en de *Profile Match*. Veel werk maar dan heb je ook wat. Alle lijsten leverden een interpreteerbare structuur op en de BF werden steeds teruggevonden. Studies met de *California Personality Inventory* (CPI) van Gough (Fleener & Eastman, 1997; Soto & John, 2009), de *Jackson Personality Inventory* (JPI) (Ashton et al., 1998), Tellegens *Multidimensional Personality Questionnaire* (MPQ) (Church, 1994) en met *Attachment Styles* (Nofle & Shaver, 2006) tonen hoge correlaties met de BF. Moosbrugger en Fishbach (2002) gebruikten de sneeuwbalwerkwijze om hun Duitse proef-

personen bijeen te krijgen. Ze bereikten 2006 deelnemers met een gemiddelde leeftijd van 31 jaar, meest studenten. De BF factoren Extraversie en Emotionele Stabiliteit waren vergelijkbaar met de EPQ-typen. Avontuur-Voorzichtigheid (het P type van Eysenck) was een mengsel van drie BF factoren: Gewetensvolheid, Aangenaamheid en Openheid.

Barelds en Luteijn (2002) vergeleken drie Nederlandse vragenlijsten: de BF, EPQ (*Edwards Personality Questionnaire*) en de NPV (Nederlandse Persoonlijkheids Vragenlijst) die gebaseerd is op de *Edwards Preference Schedule* (EPS). De factorstructuren waren gelijk in een steekproef van 254 volwassenen. De auteurs kiezen voor vier persoonlijkheidsfactoren: I: Neuroticisme, II: Extraversie, III: Conformiteit en IV: Egoïsme. Ze achten deze stabiel en ze hebben er één minder nodig. Ze durven.

*Antwoord: Er is substantiële overeenkomst tussen de BF en een aantal bekende persoonlijkheidsvragenlijsten.* Enkele instrumenten meten specifieke kenmerken: de 16PF mist de factor Aangenaamheid. De Vijf kunnen ondergebracht worden in de drie van Eysenck: Psychoticisme, nu gelabeld als *Adventure/Caution*, Extraversie en Neuroticisme. Enkele studies vermelden een eigen lezing van de factoren. Er is geen belangstelling voor het vinden van persoonskenmerken die *niet* met de BF samenhangen.

*Vraag: Tonen de BF dezelfde gemiddelden, SDs, factorstructuur en correlaties in verschillende populaties?* Sekse, leeftijd en SES zijn de bekendste covariaten. Ofschoon sekseverschillen bijna entertainment zijn: De Jongens en de Meisjes van de Tros, noemen we enkele studies over m/v-verschillen. Feingold (1994) en Costa et al. (2001) vermelden vrijwel dezelfde resultaten: vrouwen zijn Neurotischer, Aangenamer en in sommige studies Gewetensvoller dan mannen. Mannen zijn assertiever - een facet van Extraversie - en Opener. Er is geen verschil in factorstructuur. Een evolutionaire verklaring is dat mannen agressiever en vrouwen zorgzamer zouden zijn (Budaev, 1999). Niettemin lijkt een culturele verklaring evengoed mogelijk. Verklaarende hypothesen die berusten op verschillen in mannelijke en vrouwelijke biochemie en hersenontwikkeling zijn er ook. We houden van m/v verschillen. Je komt er mee in de bladen en op de TV. *Differential Item Functioning* (DIF) op de items zou een aanvulling zijn. Zo kun je kijken wat er over blijft van de verschillen. Betekenen de items hetzelfde voor mannen en vrouwen?

Is de structuur te generaliseren over leeftijd? Kohnstamm et al. (1994) vonden in de vrije beschrijvingen van kinderen door ouders vijftien clusters; 70%-80% van de beschrijvingen paste in de Vijf. Daarnaast vonden ze negen specifieke factoren voor kinderen. Er is overeenstemming én leeftijdgebonden specificiteit. Een ander voorbeeld is een studie van John et al. (1994). Zij vergeleken een vragenlijst bij jonge kinderen: de California Q set van Block: de kenmerken *ego resilient* (veerkrachtig) en over- en ondergecontroleerd met de Grote Vijf. De moeders leverden de Q set gegevens van 350 jongens van 12-13 jaar. De gegevens pasten goed in de BF. De auteurs voegden *irritability* en *positive activity* toe om de variantie in de antwoorden recht te doen. In Nederland deden Van Lieshout en Haselager (1994) eveneens een studie met de Q set en de Vijf. Zij handhaafden de Vijf en voegden *Motor Activity* en *Dependency* toe.

*Antwoord: Nauwelijks verschillen*, al vinden we graag gemiddelde sekseverschillen. Ze zijn er, maar bescheiden, ofschoon significant door de grote aantallen. De structuur verschilt niet tussen de seksen. Leeftijd zorgt voor meer dan Vijf trekken en voegt met leeftijd verbonden verschillen toe. De structuur verschilt niet in de meest onderzochte populaties.

*Hoe doen de BF het als voorspellers?* Als ik iets voorspel en het komt uit, dan weet ik iets, zei De Groot. Uitstraling van een construct is vooral het kunnen voorspellen van criteriumgedrag. Dat is gedrag, zoals eerder gezegd, waar we bang voor zijn (agressie, ziekte), medelijden mee hebben (armoede, kindermishandeling) en die tot succes, gezondheid en geluk leiden (hoge opleiding, beroep met aanzien, prestaties). Er is een overmaat aan studies.

*Vraag: Voorspellen BF scores lichamelijke en psychische gezondheid?* Marshall et al. (1993) correleerden waargenomen lichamelijke en geestelijke gezondheid met de Grote Vijf. E, N en A voorspelden dimensies van gezondheid, zoals optimistische controle: hoop, zelfwaardering en interne controle en expressie van angst en inhibitie en introspectie. Costa en McCrae deden in 1992 een poging om As II van de DSM-IV TR (2000): persoonlijkheidsstoornissen (PDs) te herschrijven in termen van de Vijf. Het *Handboek DSM-5*<sup>85</sup> bevat twee modellen voor PDs. Het eerste (pp. 645-685) herhaalt de IV versies. In de DSM-5 worden vrijwel dezelfde stoornissen vermeld maar ze worden beschouwd als *impairments in personality functioning and pathological personality traits* (pp. 761-781; onderlijning door DSM-5 auteurs). Er is kennelijk naar Costa en McCrae geluisterd maar het blijft niettemin een categoriesysteem met dimensies, gevormd op basis van kenmerken en oorzaken en expert-oordelen over de ernst van PDs. De mogelijkheden van multivariate analyses en IRT kunnen niet benut worden omdat er geen kenmerken en symptomen van PDs in klinische steekproeven zijn vastgesteld. Eysencks P-factor is al herschreven als combinatie van E+, ES+ (of N-) en C-. Miller (1991) vond dat N gerelateerd was aan de duur en de intensiteit van de klacht. E voorspelt het enthousiasme om in therapie te gaan en C correleert met de bereidheid om suggesties van de therapeut te volgen en opdrachten uit te voeren. McCrae (1991) correleerde de BF met MMPI scores en de Millon schalen. De helft van de correlaties was niet significant; de overige varieerden tussen  $r = .12$  en  $r = .63$ . Angst, depressie en psychotische klachten correleerden gemiddeld tot hoog met N en drugsgebruik met E.

Quirk et al. (2003) beweerden dat de BF *incremental validity* vertonen ten opzichte van de MMPI schalen voor sommige stoornissen en stellen voor om de BF te gebruiken voor klinische casus. De Eysenck typen Extraversie en Neuroticisme correleren substantieel met PDs. Lee et al. (2003) voegden de BF dimensie *Honesty-Humility* toe die ook in Aziatische steekproeven werd gevonden. Deze factor was bijvoorbeeld negatief gerelateerd aan de kans op seksuele intimidatie. Haslam (2001) vond dat de E facetten impulsiviteit en sensatiezoeken gecorreleerd waren met *risicogedrag*: onbeschermd seks met meer homo- en heteroseksuele partners. Sensatiezoeken was bij heteroseksuele mannen verbonden met drugsgebruik, afgezien van alcohol, de meest gebruikte drug. De Vijf voorspelden bij 459 adolescenten tussen 11 en 18 jaar de bereidheid tot het nemen van risico's. Opwinding, het opzoeken van een kick corre-

leerde  $r = .23$  met E; antisociale uitingen  $r = -.40$  met N, weerspanning en opstandig gedrag  $r = -.19$  met C en roekeloosheid  $r = -.14$  (Gullone & Moore, 2000).

Penley en Tomaka (2002) onderzochten de relatie tussen BF en *coping* processen, zoals cognitieve inschattingen, subjectieve reacties en het gebruikmaken van *copingstrategieën* met prestaties bij 97 volwassenen. Ze vermeldden ook de correlaties *tussen* de vijf factoren: N en C waren gecorreleerd evenals E en O:  $r = -.38$  en  $r = .43$ . Enkele onderdelen van elke coping strategie met de Vijf lieten  $r$  waarden tussen  $.21$  en  $.38$  zien.

DeNeve en Cooper (1998) voorspelden *welbevinden* op basis van persoonskenmerken. Ze ordenden alle scores op verschillende vragenlijsten in de Vijf factoren. N (ES) blijkt de beste predictor van tevredenheid, geluk en positieve gevoelens. E en A predicieren eveneens de hoeveelheid positieve gevoelens. De  $r$ -waarden zijn significant maar laag volgens de Cohen vuistregel ( $r =$  rond de  $.20$ ). Hayes en Joseph (2003) rapporteerden  $r$ -waarden die varieerden van  $r = .38$  tot  $r = .72$  bij 111 volwassenen tussen de BF en geluk, depressie en tevredenheid met het leven. Emotionele stabiliteit kreeg in de regressieanalyse voor het voorspellen van *happiness*, *depression* en *satisfaction with life* een  $\beta$ -gewicht van  $.59$ . Deze waarden zijn hoger dan in andere studies. De zinnen en woorden uit de ene vragenlijst stonden voor een deel ook in de andere lijst. Smedslund waarschuwde daarvoor. Het gaat in dat geval voor een deel niet over empirische verbanden maar over semantische gelijkenissen. Deze studies zijn het topje van de ijsberg. De grote omvang is te begrijpen want gezondheid is immers de eerste wens van ieder met als gevolg veel onderzoek.

*Antwoord: De Big Five voorspellen lichamelijke en psychische gezondheid significant maar bescheiden.*

*Vraag: Kunnen de BF iets zeggen over agressie?* Dat is het gedrag waar we bang voor zijn en dus iets over willen weten. Yoliffe (2013) onderzocht agressie bij 720 normale Engelse adolescenten (gemiddelde leeftijd 14.8, SD 0,41; 344 meisjes en 90% was Caucasisch). Onaangenaamheid en Gewetensvolheid waren - onafhankelijk van elkaar - gecorreleerd met zelfrapportage van geweld bij jongens. Lage scores op Gewetensvolheid waren gecorreleerd met geweld bij meisjes. Bij ingesloten meisjes van 14 tot 18 jaar correleerde het aantal gerapporteerde criminele gedragingen positief met Openheid en negatief met Gewetensvolheid (ter Laak et al., 2003).

*Antwoord: Er is een relatie tussen de Vijf factoren en agressie. De waarden zijn bescheiden (correlaties tussen  $.15$  en  $.40$ ).*

*Vraag: Voorspellen de BF werkprestaties?* Dat willen personeelsfunctionarissen en werkgevers graag weten. Schmidt en Hunter (1998) ontdeedden in hun meta-analyse de gegevens eerst van administratieve fouten en statistische artefacten en berekenden daarna correlatie tussen integriteit, gewetensvolheid en werkprestatie. Het resultaat is resp.  $r = .30$  en  $r = .40$ . Volgens Cohens norm is dit (bijna) gemiddeld. Mikulay en Goffi (1998) namen als criterium *counter productivity* : fraude, overtreden regels, stelen, fouten niet toegeven, de Grote Vijf en een aantal integriteitstests als predictoren voor *job performance*. De correlaties varieerden van  $r = .01$  tot

$r = .62$ . Een hiërarchische regressieanalyse toonde de Vijf als eerste en tests voor integriteit als tweede in rang bij het voorspellen van fraude, regels overtreden, stelen en fouten niet toegeven. Judge et al. (1997) voorspelden *absenteïsme* van personeel met de factoren E en C:  $r = .26$  en  $r = .23$ . Deze waarden zijn bescheiden. Verantwoordelijkheid en Risico nemen voorspelden de werkprestaties beter dan de BF (Ashton, 1998). Een zesde factor *Honesty-Humility* werd door Ashton en Lee (2005) toegevoegd. Johnson et al. (2011) lieten zien dat deze factor positief correleerde met oordelen van chefs en feitelijke werkprestatie. Eerlijkheid en gehoorzaamheid van personeel worden gewaardeerd, meer nog dan creativiteit en intelligentie. Dat viel Proudhon in de 19<sup>de</sup> eeuw ook al op.

Sackett en Walmsley<sup>86</sup> verrichtten een meta-analyse over persoonskenmerken die van belang zijn op de werkvloer. Ze dachten daarbij aan productiviteit, zelfstandig je werk kunnen doen, omgaan met stress en soepele verhouding met collega's. Als steeds blijken Gewetensvolheid en Aangenaamheid de belangrijkste voorspellers. Dat geldt voor functies die variëren in complexiteit, benodigde training en ervaring.

*Antwoord: De BF voorspellen werkprestaties bescheiden tot gemiddeld. Ze zijn een goede tweede na het IQ.*

Ik benut Schmitts<sup>79</sup> *Annual Review* artikel van 2014 met meta-analyses om de relatie tussen persoonskenmerken en werkprestatie samen te vatten. Er is in dat artikel een *Aptitude Treatment Interaction* (ATI) effect gesuggereerd. Dat verwijst naar de hypothese dat het effect van een behandeling samenhangt met het niveau en/of de soort van een individueel verschil variabele: een persoons- of intelligentiekenmerk. Hij stelt eerst vast dat intelligentie, geschiktheid en eerdere prestaties *performance outcomes* voorspellen in de meeste, zo niet *alle* functies en situaties. De persoonskenmerken voorspellen minder goed maar regelmatig substantieel en in ieder geval significant. De auteur merkt op dat de kenmerken moeten passen bij de werksituatie. Er worden hogere validiteitscoëfficiënten behaald met *tailor-made, context-specific personality measures, subfacets of the BF and combinations of relevant personality constructs*. Dit klinkt als een onvervalst ATI succes en dat is zeldzaam. Als de *tailor-made* metingen vooraf theoretisch uiteengezet en beschreven zijn en de voorspellingen vooraf gedaan zijn (geen *HARKING*: hypothesen formuleren als de data binnen en de eerste analyses verricht zijn) dan is er sprake van een ATI effect. Als er achteraf interacties vastgesteld worden is het een vorm van vissen naar de beste voorspellende combinatie uit een grote hoeveelheid. In dat geval is kruisvalidatie nodig. De correlaties variëren van verwaarloosbaar tot gemiddeld.

*Vraag: Hangen de BF samen met beroepsinteresse?* Larson et al. (2002) combineerden data van 24 steekproeven met drie beroepsinteresses. De laatste werden geordend volgens de indeling van Holland<sup>87</sup>. De dertig correlaties toonden gemiddelde samenhangen. Eerst wordt de beroepsinteresse-indeling van Holland genoemd en daarachter de persoonlijkheidsstrek: Artistiek en Openheid  $r = .38$ , Ondernemen en Extraversie  $r = .41$ , Sociaal beroep-Extraversie en Aangenaamheid, de correlatiewaarden zijn respectievelijk  $.31$  en  $.19$  en Onderzoeker-Openheid  $r = .28$ . De Realistische beroepsvoorkeur is niet gerepresenteerd in de BF (De Fruyt & Mervielde, 1997).

*Antwoord: De samenhang tussen de BF en Hollands beroepstypen varieert van bescheiden tot gemiddeld. Welke factor met welk beroep samenhangt is intuïtief plausibel. Het zou aardig zijn dit nog eens uit te proberen. Zou de correlatie tussen Interesse in Onderzoek en Openheid dan weer bescheiden zijn?*

*Vraag: Voorspelt de BF studiegedrag en ijver: Okun en Finch (1998) gebruikten structurele vergelijkingstechnieken en vonden dat Gewetensvolheid van 240 eerstejaars met  $r = .29$  samenhang met het voortzetten van de studie. Er was een indirect effect van Gewetensvolheid. Het verband ging via verbondenheid, toegewijd zijn (*commitment*) aan de organisatie en via intelligentie. Uitstellen (procrastinatie) is een bekend (studenten)probleem. Schouwenberg en Clay (2001) hebben er een meetinstrument voor gemaakt. Het bleek bij studenten gecorreleerd met de zes facetten van Gewetensvolheid en met het facet *tentativeness*: aarzeling, voorzichtigheid, niet *to the point* zijn van Neuroticisme. Zhang (2003) vond een bescheiden relatie tussen een diepe manier van verwerken van de leerstof: verbindingen leggen, zinvolle vragen stellen - met andere woorden niet alleen lezen en uit het hoofd leren - en Gewetensvolheid en Openheid. Dit werd in Nederland eerder gevonden door Busato et al. (1999). Zij lieten bovendien zien dat Extraversie verbonden is met een op betekenis gerichte leerstijl en Gewetensvolheid met een op reproductie en toepassing gerichte stijl.*

*Antwoord: De BF en studiegedrag en ijver hangen in bescheiden mate samen.*

*Vraag: Hangende BF samen met schoolprestaties? Gewetensvolheid en Extraversie voorspelden in een studie van John et al. (1994) schoolprestaties matig ( $r = .17$  en  $.24$ ). Mervielde et al. (1995) kwamen tot correlaties rond de  $.60$  tussen schoolcijfers en Gewetensvolheid, Openheid en Extraversie. Dat zijn uitzonderlijk hoge correlaties. Doorgaans worden waarden tussen de  $.15$  en  $.35$  gevonden. Een schatting bij ongeveer 70.000 deelnemers biedt de meta-studie van Poropat<sup>88</sup>. Schoolprestaties en Aangenaamheid, Gewetensvolheid en Openheid hangen steeds significant, maar laag ( $r = .15$  tot  $.25$ ) samen. Dit type studies gaat maar door, nu met complexere data analyses: Beaujean et al. (2011) gebruikten latente structurele modellen en vonden significante relaties tussen leesprestaties en Openheid, Gewetensvolheid en Aangenaamheid. Rekenen en wiskunde was gecorreleerd met Openheid, Gewetensvolheid en Extraversie.*

*Antwoord: Persoonlijkheid en schoolprestaties hangen samen. Vooral Gewetensvolheid doet het redelijk. Let wel, er is grote variatie in de resultaten van de studies.*

*Vraag: Kunnen politieke voorkeur en burgerschap voorspeld worden met de BF? Vecchione et al. (2011) analyseerden de relatie in vijf Europese landen. De resultaten verschillen per land. Een acceptabele samenvatting is dat er lage tot gemiddelde correlaties zijn tussen de links-rechts dimensie en de BF in de te verwachten richting. Links is Gewetensvoller en Neurotischer. Oud PvdA minister Wouter Bos zei eens dat Links tobberig is want er moet iets verbeterd worden. Rechts vindt de samenleving zoals die is goed. Rechts is zelfverzekerd en zetelt aan de rechterhand van God zei Canadese filosoof Charles Taylor ooit. Chiaburu et al.*

(2011) correleerden de BF scores met kenmerken van burgerschap: individu-, organisatie- en veranderingsgericht. Hun meta-analyse is gebaseerd op 87 steekproeven. Alle Vijf voorspelden de gecombineerde burgerschap score:  $r = .18$  met C,  $r = .14$  met A,  $r = -.12$  met N,  $r = .09$  met E en  $r = .14$  met O. De correlaties zijn door de grote aantallen significant maar verwaarloosbaar tot laag. Een beproefd recept is in een dergelijk geval het combineren van scores met hoop op hogere correlaties, zo ook hier. De prosociale alfa combinatie (C, A, N) leverde echter ook een verwaarloosbare correlatie op met een op verandering gerichte burgeroriëntatie ( $r = .06$ ). Bij de Individu- en Organisatiegerichte oriëntatie werden waarden van  $r = .19$  en  $r = .22$  gevonden. De drie burgerschapstijlen zijn nog eens geanalyseerd in een meta-studie door Berry et al. (2011). Zij vonden dat N, E en O *incremental valid* waren bij het voorspellen van de burgerschapstijlen boven C en A. De correlaties waren verwaarloosbaar, laag en bescheiden: van .09 tot .32.

*Antwoord: Er is een samenhang tussen de vijf factoren en politieke voorkeur en burgerschap.* De waarden zijn naar verhouding laag en variëren tussen verwaarloosbaar en bescheiden. *Ich hab' genug*, zult U J.S. Bach (Cantate 82) nazeggen. En dit is nog maar een fractie van predictieve studies met de BF.

*Conclusie:* De Vijf correleren significant maar bescheiden en een enkele keer gemiddeld met uiteenlopend gedrag, zoals lichamelijke en psychische gezondheid, prestaties en beroepsinteresse. De BF zijn de Haarlemmerolie voor predictie van gedrag waar we ons druk over maken. Ze zijn een goede predictieve tweede na het IQ. Er is grote variatie in correlaties tussen de BF en criteria (hier al tussen .01 en .72). Studies waar correlaties ontbreken, zijn niet gevonden. De BF stralen, dat wel, maar variabel en alle kanten op. Ze zijn niet 'het witte licht dat schitterend breekt in kleuren', 'kleuren die de daden zijn van het licht dat breekt'. Eigenlijk weinig daden bij al die persoonswoorden.

*Vraag: Welke verklaringen en interpretaties zijn er voor de relaties tussen de BF en typerende gedragingen?* Vele, *this is psychology, man!* Als je niet weet of er één of meer factoren zijn, dan wordt verklaren lastig. Je moet voor ieder ander aantal een nieuwe verklaring verzinnen. Daar stappen onderzoekers overheen als ze aan verklaren beginnen. Ze zoeken en vinden oorzaken voor verschillen op Vijf en interpreteren de factoren. Het laatste is gemakkelijker dan het eerste want het berust op de bereidheid een betekenis te geven aan de Vijf en betekenissen die anderen geven serieus te nemen (Gadamer<sup>89</sup>). In de onderzoeksliteratuur lopen de twee soms door elkaar.

We hadden de *lexicale* functionele verklaring al. We houden termen over die functioneel voor ons zijn. Je hebt er wat aan als je bijvoorbeeld weet dat iemand hoog scoort op Gewetensvolheid want die voert zijn plannen trouw en volgens de regels uit. McCrae en Costa gaan verder. De basis is *genetisch* (van wat niet eigenlijk?). De factoren/trekken hebben een betrekkelijk hoge  $h^2$ : de overerfbaarheidscoëfficiënt. Dat is het percentage variantie van een trek of kenmerk in de populatie dat verklaard kan worden door de hoeveelheid gedeelde genen: de mate van genetische verwantschap: monozygotische tweelingen 100%, broers en zussen 50%, enzovoort. De sommen zijn vaak gemaakt en de correlaties bij MZ zijn rond de .50. en bij voor DZ rond de .27. Gedragsgenetisch onderzoek is grof. Er komt geen meting van omgeving of genen aan te pas. Alleen de variatie in verwantschap wordt gecorreleerd met

variatie van scores op de BF. McCrae et al. zijn zich er van bewust dat veel genencombinaties betrokken zijn bij elk van de Vijf. Omdat genen apart kleine effecten hebben, proberen ze clusters van mogelijk verbonden aparte nucleotide polymorfismen (SNP's) te identificeren. Ze omschrijven die clusters als moleculaire polymorfismen (MPSs), waarvoor tussen de 4 en 2.497 SNP's nodig zijn. Ze deden een studie op het eiland Sardinië, omdat door het isolement van de bevolking het DNA grotendeels bekend is. De MPSs voor Neuroticisme, Openheid, Aangenaamheid en Gewetensvolheid (niet Extraversie) bevatten genetische informatie. Ze trekken deze conclusie op basis van bescheiden correlaties tussen MPSs en de scores op de NEO-PI. Het bij elkaar brengen van de SNPs laat zien dat het moeilijk is om rechtstreeks verbanden te leggen van genen (MPSs) met de antwoorden op vragen uit de NEO-PI.

De BF worden regelmatig tot twee factoren samengevoegd: de alfa-factor: Aangenaamheid, Gewetensvolheid en Emotionele Stabiliteit en de bèta-factor: Extraversie en Openheid. De Young (2006) stelt een *biochemische* verklaring voor en labelt de alfa als Stabiliteit en de bèta als Plasticiteit. Het verschil is gebaseerd op variatie in de functie van serotonineproductie, een stof die in de pijnappelklier omgezet wordt in melatonine en het ritme van lichaamsfuncties regelt, en in de functie van dopamineproductie, een stof die de productie van adrenaline verzorgt en als neurotransmitter fungeert.

*Antwoord: Het interpreteren en verklaren gaat van semantische label, statistisch artefact tot een neurologisch, biochemisch, genetisch gefundeerde basis voor verschillen in trekken.*

Als er al zoveel interpretaties zijn mag ik er wel een aan *toevoegen* en misschien bestaat die zelfs al: Extraversie en Openheid wijzen op individuele verschillen in relatie tot de wereld van objecten en sociale verschijnselen, Neuroticisme drukt de relatie van de persoon tot zichzelf uit en Gewetensvolheid en Aangenaamheid die tot anderen. Deze interpretatie vat persoonskenmerken als *relaties* op van het Ik/Mij met de fysische, sociale buiten- en binnenwereld. Stelt u zich iemand voor die bij een aanzwellend onweder alleen op het strand staand kijkt naar de ziedende zee *...en hij die uitzag huivert, naar binnen keert hij zich, waar zijn stille dode neerligt, en wacht er de uitspraak en het openende weerlicht, en smacht dat voltrekking moog beginnen over zijn schuld: hij, die aan zijner dromen duistere vervoering haar aanhanke-lijk leven ontredderd in vereenzaming gedreven, prijs gaf, hij roept, dat vergelding kome...* (Roland Holst<sup>90</sup>). Je komt niet zonder kleerscheuren van welke relatie met wat en wie dan ook af. Een mens is een *'moi, insatiable du non-moi'* (Baudelaire).

Ik heb me afgevraagd of deze bespreking niet veel eer is voor de vijf dimensies. Ik heb het gedaan, omdat het mij liet zien hoe er met constructen omgegaan wordt, zowel theoretisch als methodologisch. Dit is exemplarisch voor persoonlijkheidsconstructen, zowel voor nieuwkomers als Emotionele Intelligentie en *Mindfulness* en oudgedienden zoals Depressie, Neuroticisme (Freud) en Extraversie (Jung). Het zijn misschien termen die een

gesprek over het karakter, de gedragskenmerken en eigenaardigheden van een cliënt kunnen stroomlijnen. Hoef je niet meer naar hobby's, sportclubs, relaties en *good breeding* te vragen, al mag dat laatste officieel niet.

Een tekst met eenzelfde strekking kun je schrijven over intelligentie- en situatieconstructen. De predicties van intelligentie komen hoger uit. Voorspellingen op basis van verschillen in situaties en binnen een situatie bereiken ongeveer dezelfde waarden als die van persoonsconstructen. Het ziet er eender uit voor populaire *effecten*. Verderop blijkt dat uit onderzoek van het Flynn effect. Ook daar raak je overvoerd met studies die verschillen in de omvang van het effect, de generaliseerbaarheid (geldt het voor enkele of vele facetten van intelligentie, voor beide seksen, over leeftijden, voor hoogbegaafden) en de interpretatie (wat veroorzaakt het effect?). Er zijn ook studies waar het effect omgekeerd is. Het verschaft je geen helder beeld dat je kunt toepassen op de individuele cliënt.

*Vraag: Wat gebeurt er als ik het BF profiel voorleg aan de terugpratende cliënt?* De diagnosticus kan de Vijf en de verklaringen, oorzaken, interpretaties aan de cliënt voorleggen. Hij kan ook twijfel zaaien met antwoorden zoals, het is genetisch, een statistisch artefact, een semantische gelijkenis. Hij kan opmerken dat van al deze kennis de BF het begin en het einde is. Geef het profiel maar, daar heb je iets aan: tussen de 2 en 7. Dat kan iedereen onthouden. Je kunt er bovendien 'Kouweriaans' met de cliënt over in gesprek gaan. Hij mag terugpraten, bijvoorbeeld dit: 'Wat de BF - de gemiddelde burger - Gewetensvol noemt, vind ik saai, initiatiefloos, niet *out of the box* durven denken'. Zoiets kan een cliënt nog verzinnen. En, wat hebt u als diagnosticus dan te zeggen? Wat moeilijker is de tegenwerping dat de correlaties bescheiden zijn en de samenhang of het effect voor een beperkt deel van de populatie geldt: U zult een cliënt niet gauw horen zeggen: 'Het is niet zeker of het effect of verband voor mij geldt; ik val misschien substantieel buiten de regressielijn'. Over mondigheid gesproken! De diagnosticus kan de genetische interpretatie nuanceren. Ze wordt vaak gelezen als onveranderlijk. *The genetic myth* project in de VS gaat daar tegen in.

*Antwoord: We weten het niet, nog niet uitgezocht.*

*Een voorstel:* Probeer eens een gesprek over het BF resultaat te voeren met de cliënt. Het kan verschillend uitpakken. Uw cliënt kan het als een bondige samenvatting beschouwen: 'Kort, maar krachtig zegt u aardig hoe het zit met mijn karakter, hoe ik in elkaar zit'. Het kan zich ook voordoen dat hij zich *tekort gedaan* voelt door het neerslaan van veel en genuanceerde informatie in het platte vlak van de Vijf. Dit overkwam mij toen ik een vraag van Berkowitz, een onderzoeker van morele ontwikkeling, beantwoordde. Hij vroeg het auditorium een moreel hoogstaande persoon in gedachte te nemen en die te beschrijven. Ik nam Leo Vroman<sup>91</sup> in gedachte, de joodse biochemicus, gevlucht naar Indië, wachtend op zijn Tineke (door hem beschreven in brieven), een beroemd bloedonderzoeker, verblijvend in de VS ('liever heimwee dan Holland') respect tonend voor SS-ers want zij hadden immers bloed: het wonder, dat ieder mens gelijk en kwetsbaar maakt, enzovoort. Na mijn verhaal met allerlei details, zelfs met een ter plekke gemaakte parafraze van

enkele regels uit 'Voor wie dit leest': *Gedrukte letters laat ik U hier kijken/ maar met mijn warme mond kan ik niet spreken/mijn hete hand uit dit papier niet steken/ wat kan ik doen? Ik kan U niet bereiken.* Mijn palaver werd neergeslagen in drie trek-steekwoorden. Eén van de deelnemers vermoedde dat ik Leo Vroman beschreef en vond dat de dichter (en ik) tekort gedaan werden. Geen nieuwsgierigheid naar wie het was die ik beschreef; het was toch een bijzondere man, vond ik. *Scientia non est individuorum.* Ik sluit me aan bij wat Piketty zegt over Amerikaanse economen: wetenschappelijk en apologetisch. Voor Amerikaanse psychologen geldt dat ook een beetje.

*Uw persoonlijkheid veranderen* Een cliënt wil vaak weten over hoe veranderlijk zijn persoonlijkheid is. Hij wil met de kennis over zijn persoonlijkheid iets bereiken: inzicht, een doel (*target*) behalen, een probleem oplossen, zijn leven veranderen. Studies geven zelden concrete aanwijzingen. Er is vooral onderzoek verricht naar stabiliteit en continuïteit door de dominantie van de individuele verschillen benadering en de populariteit van (latente) trekken. Hudson en Fraley Hudson en Fraley beweren voor het eerst te hebben laten zien dat je je trek-scores als je dat wil, kunt veranderen. Er moet nog nagegaan worden of dat zo blijft in uiteenlopende situaties en gedurende een grotere tijdsperiode dan de 16 weken van hun onderzoek. Ontwikkeling en verandering liggen niet empirisch voor het oprapen. Je hebt een theorie nodig die ontwikkeling of verandering van een gedrag of een reeks gedragingen beschrijft. Omdat ontwikkeling niet uit data blijkt, gebruikte Werner<sup>92</sup> een apriori definitie van ontwikkeling: toenemende differentiatie en integratie volgens het orthogenetisch principe. Ontwikkeling koerst volgens hem naar een 'goede Gestalt'. Werners theorie is vergeten maar duikt af en toe op. Bowlby's volgorde van hechtingspatronen wordt omschreven als differentiatie en integratie door Raeff<sup>93</sup>: Er is eerst een fase van *pre-attachment*: kinderen glimlachen tegen iedereen, vervolgens is er een overgang naar 'hechting in de maak is' en daarna ontstaat 'echte' attachment. Dat is de vorming van een wederkerige verhouding die behouden blijft door de tijd heen. Eerst gaan kinderen objecten van mensen onderscheiden en daarna maken ze onderscheid tussen mensen en zijn nog weer later in staat het perspectief van een ander in te nemen: sociaal-perspectief-nemen. Raeff ziet de organismische ontwikkelingstheorie als een inspirerend kader.

De ontwikkelingsconcepten differentiatie, integratie en hiërarchisering zijn abstract en kunnen op veel verschijnselen toegepast worden: van het gedrag van een molecuul, een cel tot de aanpassing aan een nieuwe baan. Bijgevolg zijn er volop illustraties van de principes te bedenken maar abstracte specificiteit is meestal geen aansporing tot toetsend empirisch onderzoek.

Individuele verschillen vragenlijsten zijn ongeschikt om ontwikkeling af te beelden (Heymans<sup>94</sup>). *Can personality change?* Er zijn zelfhulpboeken volop om intelligentie en persoonlijkheid te verbeteren. Meestal treft men open deuren aan, zoals, verhoog uw zelfdiscipline, wees geduldig en emotioneel stabiel. De zelfverbeteringsmarkt had in 2008 in de VS een budget van 10 miljard dollar en hield 30.000 *personal coaches* aan het werk (Marketdata Enterprise Inc., 2010). Interindividuele verschillen instrumenten als de Big Five zijn niet gemaakt om verandering/ontwikkeling vast te stellen. Je leert er niets over ontwikkeling/verandering.

*En, de diagnosticus* Hij moet zich zien te redden door op de individuele cliënt te letten en op de vele draadjes waarmee hij verbonden is aan zijn eigen kenmerken in vergelijking tot anderen, zijn context en zijn ontwikkeling. Hij vermijdt het om te komen onder de complexe bedrading die studies over de BF hem bieden. Het zijn voor hem handige termen die een gesprek over het karakter, de gedragskenmerken en eigenaardigheden van een cliënt kunnen stroomlijnen. De diagnosticus neemt in tegenstelling tot de onderzoeker het terugpraten van de cliënt serieus. Als het over veranderingen gaat, kun je de BF overslaan en moet je een ontwikkelingsconstruct verzinnen.

## 7. Het einde: Ware beweringen, valide diagnoses

Dit einde gaat over wat is waarheid, wat is een valide diagnose? Ik beweerde eerder dat je de beginvraag van dit thema 'Wat is diagnostiek' beter kon vermijden. Dit geldt nog sterker voor de eindvraag: 'Wat is een ware bewering, wat een valide diagnose?' Onttrek je aan deze vragen, laat het over aan gepensioneerde leunstoelfilosofen met veel vrije tijd! Gun ieder zijn eigen waarheid, scheelt gekrakeel. Het is maar een spelletje van de elite, denk aan Pontius Pilatus die graag wilde babbelen: *Quid est veritas?* En, het antwoord van de in het Westen beroemdste goeroe - en van alle voor en na Hem - luidde en zal luiden: 'Ik ben de Waarheid'. En, ze leveren de methode er meteen bij: 'Ik ben de Weg'. Toch waag ik me aan dit onderwerp. Ik heb het concept van waarheid bij kinderen onderzocht. Piaget heeft de ontwikkeling van abstracte natuurkundige en morele begrippen bij kinderen geanalyseerd maar ik herinner me geen onderzoek naar het waarheidsconcept.

Eerst scheer ik oppervlakkig - kan niet anders - over de geschiedenis van de westerse filosofie. Daar hebben psychologen hun waarheidsopvatting uit samengesteld. Waarheidsvinding verschilt tussen disciplines. Als je naar de rechtspraak kijkt, zie je meteen andere accenten. Al willen we waarheid, dé waarheid bestaat niet. Toch kunnen we niet zonder. Als wetenschapper en professional moet je weten wat je denkt en doet. Je kunt dat niet ongearticuleerd laten. Vervolgens is er de Sartrianse keuze. Je bent vrij, maar je voelt je er ongemakkelijk bij. Waarom juist dit en niet dat? Waarom op deze en niet op die manier? Je bent vrij, kies! Vervolgens ga ik in op het belang dat we als leken, professionals en onderzoekers hechten aan waarheid.

Impliciete theorie is so wie so interessant, bovendien kun je er als diagnosticus niet omheen. Dichters en schrijvers maken inspirerende opmerkingen over waarheid. Verder is er naast het emotionele belang - je voelt je naar, ontdaan als de waarheid niet boven tafel komt, denk aan het neergestorte vliegtuig in de Oekraïne - het kennisbelang. Wat wil je met die kennis bereiken? Beheersing/controlle van gedrag? Inzicht en begrip van het gedrag? De cliënt emanciperen?

In de diagnostiek en psychologie komt validiteit het dichtst bij het begrip waarheid. Zij is vooral methodologisch uitgewerkt als test- en researchvaliditeit. Methodologisch betekent in de psychologie het vermijden van fouten en vertekeningen, de *Idolen* waar Linschoten indringend over schreef. Dat mondt ondermeer uit in het vermijden van subjectiviteit. Die methodologie schuurt voor een deel in de diagnostiek als  $n = 1$  onderzoek. Objectiviteit is meer dan en anders als objectief scoren van antwoorden op items. Ik pleit voor een objectieve diagnose, dat wil zeggen een die recht doet aan de cliënt als een idiosyncratisch zich ontwikkelend subject.

*Het achtergelaten, wijsgerige verleden van de psychologie:* Hoe komen we aan ware kennis, hoe weten we dat onze kennis geldig is? Dit zijn vragen die al in de oudste ons bekende geschriften gesteld zijn. We haalden ze uit Griekenland, de bakermat van het westers denken. De Grieken moeten misschien Europa uit, maar hun denken raken we nooit kwijt. Als je een ruwe schets moet maken aan de hand van regels, principes van die geschiedenis, is dit er

misschien een: a. zoeken naar zekere kennis op welke manier dan ook (van Plato, via Descartes tot Locke en Hume) b. het inzicht dat kennis nooit volledig is, maar je kunt ermee leven (Kant) en c. denken en kennis moeten iets opleveren (van Marx tot de moderne toegepaste filosofie: combinaties met thema's als leven, gezondheid, techniek, voeding, enzovoort). Hieronder ga ik enkele 'waarheidszoekers' langs. Het is een selectie op basis van wat er is blijven hangen uit het onderwijs in de geschiedenis van de filosofie<sup>96</sup>. Mijn selectie verantwoord ik niet. Bij geen interesse, kunt u het overslaan en onder het kopje: 'Wat heeft de psycholoog eraan over gehouden?' verder lezen.

Bij Plato is de bron van kennis de *herinnering* aan absolute Ideeën over Eenheid, Waarheid, Goedheid en Schoonheid. We kunnen daar alleen de schaduwen van waarnemen. We zijn gebrekkige kenners. Aristoteles vatte menselijk kennen en intelligentie op als *spontane actieve krachten*. Het verstand vormt begrippen door het algemene uit het individuele en bijzondere te halen. Materie zonder vorm is chaos, maar met de potentie om in de greep te komen van de vorm. Ziel, geest is vorm en maakt materie tot iets intelligibels, dat wil zeggen iets dat toegankelijk is gemaakt door ons verstand. De constructen die we van de wereld vormen zijn een compromis tussen elementen uit de objectieve wereld en wat bedacht, uitgevonden of gegeven is door of met ons verstand. De ziel is het levensprincipe van wat beweegt: planten, dieren en mensen. De laatste tonen naast kennen emoties als verlangen, wraak, overheersing, vreugde. De basis van zijn ethiek is het bewaren van het midden: 'niets té', 'de gulden middenweg': zelfcontrole. Bij Thomas van Aquino bijna 16 eeuwen later spelen de zintuigen een rol in het kennen van de wereld en God. Eén van zijn vijf 'godsbewijzen' berust op de oneindige eenheid, waarheid, goedheid en schoonheid van de wereld (Plato). De overige ontleende hij aan Aristoteles: dat van de Onbewogen Beweging, de Onveroorzaakte Oorzaak, het uit zich Volstrekt Noodzakelijke. De vierde is de Platoonse. De vijfde weg is nu nog populair: de Intellectuele Ordenaar der naturen (Peters<sup>95</sup>, pp. 451-461). Sommigen erkennen het bestaan van een Ontwerper - de vijfde weg. Daar bedoelen ze niet iemand mee die in Delft bouwkunde gestudeerd heeft of de Eindhovense *Design Academy* bezocht heeft.

Descartes (1569-1650) is de filosoof van de nieuwe tijd. Doordat in de 16<sup>de</sup> eeuw schepen uitvoeren en er nieuwe continenten ontdekt werden, kwam men tot het inzicht dat kennen cultureel bepaald, relatief is. Descartes probeerde zich zowel te onttrekken aan het stellen van zinloze vragen als aan een hulpeloos relativisme. Hij zocht *zekerheid*. We kunnen aan alles twijfelen, zelfs aan het bestaan van de wereld, want dat kan een illusie zijn, maar we kunnen de twijfel zélf niet betwijfelen. Het 'Ik denk dus ik ben': 'Cogito ergo sum' is een onontkoombare intuïtie. Kennen en waarheidsvinding worden daardoor een aangelegenheid van het bewuste denken, van de *ratio*. Filosofen moeten heldere en duidelijke ideeën formuleren. Hij vond die helderheid van de wereld van objecten en verschijnselen in de kenmerken lengte, breedte en hoogte (en beweging). Het bewustzijn beschikt over met de geboorte gegeven, ingeboren ideeën. De belangrijkste zijn: (1) een oneindige substantie van het meest volkomen zijn, dat uit zichzelf bestaat en God heet - daarmee redde hij zich uit een materialistisch atheïsme dat hem de kop gekost zou hebben (2) uitgebreidheid, dat wil zeggen van lichamen en voorwerpen en (3) niet-uitgebreidheid van de denkende substantie, de ziel of het bewustzijn. Hij noemt daarnaast nog andere ingeboren ideeën, zoals getal, tijd, plaats, beweging en vorm of gestalte. Enkele daarvan

zullen in de 19<sup>de</sup> eeuw terugkeren in het werk van Kant als synthetische oordelen apriori. Zijn tweedeling achtervolgt de psychologie in het *mind-body* vraagstuk, het onderscheid objectief-subjectief en de doorzichtigheid van ons bewustzijn en nog veel meer. De westerse rationalistische traditie kreeg stem door Descartes en is nu nog herkenbaar.

Locke (1623-1704) gaf de nieuwe filosofie een radicale wending. Hij stelde vast dat we bij kinderen niets merken van ingeboren ideeën. Kennis komt volgens hem tot stand door *zintuiglijke ervaring: sensations* en door waarneming van onszelf: *reflection*. Die twee samen leiden tot voorstellingen: *ideas* die met elkaar verbonden raken door *associations*. Hiermee is afgerekend met de inbreng en voorrang van het werkzame verstand en ingeboren ideeën. Hoe die verbindingen tot stand komen en hoe we tot gemeenschappelijke voorstellingen, *ideas* komen, is een probleem. We hebben immers unieke ervaringen en we kunnen ook nooit de ervaring en het bewustzijn van een ander doorgronden. Hume (1711-1776) beweerde dat de associaties niet willekeurig zijn, maar wetmatig, ze raken verbonden door gelijkenis, eenzelfde oorzaak en dezelfde plaats en tijd. We kunnen tot een gemeenschappelijk beeld van de fysische werkelijkheid komen, omdat we in een wereld leven met universele verschijnselen en natuurwetten. Dit past ook op de sociale werkelijkheid en doet denken aan de behavioristische leerwet van 'contiguiteit' dat wil zeggen dat een stimulus en een respons verbonden raken, omdat ze in tijd en plaats nabij zijn, maar ook aan de Gestaltwetten die op natuurlijke wijze orde in de waarneming van particulariteiten aanbrengen.

De Britten formuleerden de *empiristische traditie* waarin psychologie en diagnostiek voor een deel in geworteld zijn. We lopen achter de feiten aan en proberen weinig of niets aan de feiten op te leggen. We verkrijgen op deze wijze objectieve kennis en deze kennis heeft autoriteit, dat wil zeggen, je kunt er niet om heen en moet ze accepteren.

De continentale en Britse tradities van rationalisme/idealisme en het empirisme zijn door Kant verbonden. Hij vroeg zich af wat we kunnen weten en wat niet. Wat en hoeveel kan het verstand (*Vernunft*) los van de ervaring kennen? Is er geen kennis die al in een construct of in ons zit opgesloten, los van enige ervaring? Dat de cirkel rond en lichamen en objecten uitgebreid zijn, is een gegeven vooraf: een apriori. Het is niet een resultaat van onze waarneming van veel cirkels of het zien van veel objecten. Dat oordeel over de rondheid en uitgebreidheid is geen empirisch oordeel achteraf, maar een analytisch oordeel vooraf. Andere oordelen over kenmerken vellen we achteraf na waarneming en studie. Bijvoorbeeld: deze objecten zijn zwaar, hebben een kleur. Om van de toevallige associaties van Hume en de ingeboren ideeën van Descartes af te komen, ging Kant op zoek naar samengestelde (synthetische) oordelen vooraf (apriori). Voorbeelden van deze oordelen zijn ruimte, tijd en causaliteit. Ons verstand kan niet onder het synthetisch oordeel apriori uit en dus 'moeten' we objecten, gebeurtenissen en gedragingen als een oorzaak-gevolg ketens opvatten. Daarnaast kunnen we informatie over de werkelijkheid verkrijgen door empirisch onderzoek, bijvoorbeeld door kenmerken van objecten en mensen vast te stellen.

Volgens Kant is dit een Copernicaanse revolutie, want het kennen richt zich niet naar de

voorwerpen; ze richten zich naar ons: de kenners. We kennen de dingen op zichzelf (*an sich*) niet, we kennen ze in het licht van onze synthetische oordelen apriori en kunnen daarnaast onderzoek doen en geldige synthetische oordelen aposteriori over verschijnselen en mensen vormen en uitspreken. Kennen en weten zijn bij Kant relationeel, dat wil zeggen met een eigen inbreng, waaraan je je niet kunt onttrekken en met een inbreng van de objecten met hun kenmerken.

In de vroege 20<sup>ste</sup> eeuw zijn er nog invloedrijke filosofen geweest, bijvoorbeeld Heidegger. Hij spreekt in zijn hoofdwerk *Sein und Zeit* (1927, 1986) over *Existentialen*: onontkoombare zijnswijzen, zoals het 'in-de-wereld-zijn'; 'het met-anderen zijn'; 'het zelf-zijn' en 'het ten-dode-zijn'. Deze drukken de tijdelijke, ruimtelijke en sociale aard van het menselijk bestaan uit en doen herinneren aan Kants synthetische oordelen apriori. Heidegger zegt immers dat gebeurtenissen, voorwerpen, de wereld, de medemens geen bestaan op zichzelf hebben, maar slechts in relatie tot elkaar. Bovendien komt in deze *Existentialen* tot uitdrukking dat de mens een individu, een zelf is. Hij is historisch, dat wil zeggen dat hij zich ontwikkelt. De gebeurtenissen, verschijnselen, de wereld en medemensen vormen het brute, ruwe materiaal, de ruwe stof die door de kennende geest gevormd wordt. Merleau-Ponty heeft zich vooral ingezet om het Cartesiaanse dualisme te overwinnen: het lichaam is geen puur object en geen puur subject maar een *Corps-sujet*.

De filosofie van het eind van de 20<sup>ste</sup> en begin 21<sup>ste</sup> eeuw is niet meer in te delen volgens het schema Rationalisme, Empirisme, Idealisme, Verlichting, enzovoort. Ze bevat thema's die aansluiten bij maatschappelijke vragen. Ook filosofen moeten valoriseren. We zien titels als toegepaste filosofie, filosofische therapie, levensfilosofie en filosofie en techniek, filosofie en gezondheid. Sommige filosofen geven cursussen in Franse landhuizen of begeleiden reizen naar de bakermatten van onze beschaving. Er zal een kleine groep vakfilosofen zijn die zich met atomaire onderdelen van de filosofie bezighoudt en daarover publiceert. Dat ligt buiten het veld van de psycholoog.

*Wat heeft de psycholoog eraan over gehouden?* Zijn waarheidsbegrip is een combinatie van Engels empirisme en Frans/Duits rationeel formalisme. De vraag wat we aan de werkelijkheid buiten ons ontleen en wat we zelf inbrengen in het kennen van gedrag is in de psychologie in de vorm van een compromis opgelost. Je neemt zowel de eigen inbreng in de vorm van een theorie serieus als direct observeerbaar of waarneembaar gemaakt gedrag. Het zijn twee polen die spanning moeten houden, dat wil zeggen, dat niet de een en ook niet de ander zijn stempel volledig kennisverwerving moet drukken. Zo blijf je nadenken & construeren en kijken & registreren. Er is eerder in de tekst afstand genomen van zowel het pure verzamelen van feiten (Bacons mier) als van het bouwen van logisch consistente systemen en onwrikbare structuren (Bacons spin).

Psychologen maken zich niet druk over hun wijsgerig verleden. Ze hebben de logisch positivistische stijl aanvaard met inbegrip van de empirisch-analytische wijze van verwerven van kennis over gedrag. En dat heeft de psychologie tot een serieus vak gemaakt.

*Dé waarheid bestaat niet:* De waarheid en niets dan de waarheid! Ieder wil dat, maar net even anders. In de wetenschap kan dat om een logisch consistente structuur gaan die in eerste instantie de werkelijkheid niet hoeft te beschrijven: je mag van je eigen ogen het ontzaglijk glanzende beschouwen en als de spin uit je eigen lijf de prachtigste webben (theorieën) weven. Empirische wetenschappers moeten daarentegen een verband verzinnen en bewijzen tussen de constructen, formules, structuren en de werkelijkheid van verschijnselen, objecten en gedrag.

Je hoeft maar een stap buiten je vak te zetten of er is een andere opvatting. Waarheidsvinding betekent bijvoorbeeld in de nederlandse rechtspraak het leveren van het wettig en overtuigend bewijs, dat X het delict gepleegd heeft. X kan ontkennen, maar als dit bewijs geacht wordt geleverd te zijn, is hij schuldig. De rechter is vooral beducht voor de fout dat hij iemand veroordeelt die onschuldig is. Hij laat X liever lopen ook al vermoedt hij dat X het delict gepleegd heeft. Hij krijgt het dan niet voor elkaar dat wettig en overtuigend bewijs te leveren. In de diagnostiek wegen valse positieven en valse negatieven echter even zwaar. In China behoort de bekentenis van de verdachte tot de bewijsvoering. De politie zegt dat 99,5% van de verdachten bekent. Hij vertrouwt op zijn verhoortechniek, inclusief de leugendetector. In de VS gaat het weer anders. Je kunt daar bekennen om strafvermindering te krijgen, ook als je onschuldig bent. De verschillende procedures en de uiteenlopende strafmaten berusten op gedachten en verwachtingen of en hoe je de waarheid over een misdaad boven water krijgt.

Omdat er niet één waarheid is, houdt elke tak van wetenschap en zelfs iedere persoon er een op na. De bewering van communisten dat er één waarheid (*Pravda*) of géén waarheid is, begrijp ik wel als wens en als dictaat van de macht, maar ze is precies daarom onjuist en onethisch. Ook met Duijkers bewering: 'Er is één psychologie of er is geen psychologie', heb ik moeite. Hij bedoelt het waarschijnlijk 'scholastisch', niet zo vreemd gegeven zijn wijsgerige achtergrond. Zijn uitspraak spoort met het eerste scholastische logische axioma, dat van non-contradictie: 1.1: '*It is impossible that the same thing be and not be at the same time and in the same respect*'. De axioma's zijn van het internet te downloaden: *A scholastic list of philosophical axioms*. Feitelijk houden psychologen er heel verschillende opvattingen op na, denk aan de verschillende wijsgerige stromingen. Ze bestrijden elkaar wel degelijk en ze hoeven het eigen ongelijk niet te erkennen. Wetenschappers, denk aan alfa's en bèta's, betwisten elkaars waarheidsclaims. Ieder heeft zijn waarheid, geldigheid. Dat mag je zeggen zonder een radicale relativist te zijn.

Kun je niet beter als de pragmatist zeggen dat er geen antwoord op die vragen is. Ga over tot de orde van de dag! Zoek dingen uit die iets betekenen en waar de mensen iets aan hebben. Of, kies na een beetje nadenken en overleg een werkdefinitie en zeur niet. Of, doe empirisch onderzoek, dan kom je jezelf tegen en zuiver je de begrippen uit. Of, is het gesteld als met het *kritische* en *praktische Vernunft*: de dingen *an sich* kunnen we niet kennen, maar dank zij ons *praktische Vernunft* kunnen we er aardig mee uit de voeten.

*We kunnen niet zonder waarheid, al bestaat ze niet:* We, dat zijn leken, professionals en onderzoekers. Waarheid is eerst een emotioneel belang. Leken, diagnostici en professionals

willen objectief handelen en de waarheid dienen. Niemand, of bijna niemand, wil constant bedriegen en in ieder geval niet doorlopend bedrogen worden. Dat houdt voor leken in: de waarheid spreken, hoe vaag dat ook klinkt. Wetenschappers willen onderzoek doen dat iets zinvols aan het licht brengt. Ze houden daarbij methodologisch hun emoties op afstand maar ze zijn wel degelijk betrokken bij het onderwerp en bij de geldigheid van hun onderzoeksresultaten. Diagnostici willen rapporten schrijven met diagnoses die de cliënt recht doen. We willen niet alleen reclame horen, maar zinvolle en juiste informatie.

Dat emotionele belang weten dichters tot uitdrukking te brengen, bijvoorbeeld Emily Dickinson<sup>97</sup>: *Tell all the truth/ But tell it slant-/ Success in CIRCUIT lies/ Too bright for our infirm delight/ The Truth's superb surprise/ As lightening to the Children eased/ With explanation kind/ The Truth must dazzle gradually/ Or every Man be blind.* In ouderwetse vertaling: *Zeg de Waarheid zijdelings/ Een omweg voert naar het doel/ Te fel is Waarheids grootste schok/ voor ons krank lustgevoel/Als bliksem rustig uitgelegd/Aan het beangstigd kind/ Moet waarheid lichten gaandeweg/ of anders maakt zij blind.*

Ze richtte zich feitelijk tot niemand. Haar gedichten werden na haar dood in een la gevonden, maar ze zijn er voor iedereen. Haar brieven bereikten een klein publiek. Bij wetenschappers is dat anders: Er zijn er ongeveer 15 miljoen. Velen onder hen zijn tevens lezer.

Romans over wetenschappers komen van schrijvers die deze wereld van binnenuit kennen. Er zijn er vele. Een enkel voorbeeld: Kingsley Amis van *Lucky Jim* (1953) was een universitair literatuurdocent. Hij beschrijft de slangenkuil/apenrots, die de werkring van zijn hoofdpersoon Dixon is. Deze haat zijn professor, is zich bewust van de trivialiteit van diens onderzoek, maar voegt zich naar de Britse universitaire mores. Ieder vecht voor zijn carrière. Hij is bereid daarvoor een liefdeloze verhouding aan te gaan. De Amerikaanse literatuurdocent John Williams schreef de roman *Stoner* over het universitaire leven. De roman (1<sup>e</sup> druk, 1965; Nederlandse vertaling: 2013) gaat over de loopbaan van een universiteitsmedewerker. Zijn taak wordt opgeheven. De hoofdpersoon neemt inhoudelijk werk serieus: hij is zich bewust steeds onvoldoende te weten. De romans en gedichten die hij besprak waren er niet om er tentamen over te af te leggen. Ze moesten de studenten iets leren en door hen ervaren worden.

W.F. Hermans' *Onder professoren* is een Nederlands voorbeeld. Nu weten we tenminste hoe het op de Rijksuniversiteit Groningen (RUG) toeging en misschien nog toegaat. Goed dus. De boeken van Voskuil over het Meertens instituut zijn een voorbeeld van de - in de ogen van de hoofdpersoon Maarten Koning - venijnige sociale structuur van een wetenschappelijk instituut en van het triviale onderzoek dat daar verricht wordt. Het gaat om waarheid, dat wel, maar ook om posities in een hiërarchie.

Grossman, de Russische oorlogscorrespondent (WO II), schreef in februari 1962 Chroesjtsjov een brief om zijn boek *Leven en Lot* (2014, pp. 953-957) gepubliceerd te krijgen: *... mijn fysieke vrijheid is zinloos als het boek waaraan ik mijn leven heb gegeven zich in gevangenschap bevindt. Ik heb het ten slotte geschreven; ik heb er geen afstand van genomen, en dat zal ik ook nooit doen. Het is nu twaalf jaar geleden dat ik aan dit boek begon te werken. Ik geloof nog steeds dat ik de waarheid heb geschreven, uit liefde en medelijden, omdat ik in mensen geloof. Ik verzoek u mijn boek de vrijheid te geven (onderlijning: jtl).* Op zijn brief kreeg hij geen antwoord.

*Ontwikkeling van het waarheidsbegrip:* Onze waarheidsopvatting is ondermeer een vrucht van de cognitieve ontwikkeling. Als men kinderen tussen de acht en twaalf jaar ondervraagt over wat liegen en de waarheid spreken inhoudt en hoe ze weten of een bewering over een persoon of gebeurtenis waar is, dan geven ze antwoorden die we in uitgewerkte vorm in onze tekstboeken aantreffen. Het is waar 'omdat het logisch is', omdat 'het echt gebeurd is', 'omdat je het kunt zien'. Dit is de correspondentie tussen een bewering en een stand van zaken in de wereld. Ze zeggen bijvoorbeeld 'omdat je weet dat het waar is'. Ze voerden ook aan: 'iemand heeft me het eerder verteld', en 'het is hetzelfde verhaal'. Dit lijkt op het hanteren van het coherentie criterium. Ze kunnen ook tot de bewering: 'het is waar' komen, omdat 'een ander het heeft gezien', 'anderen vertelden hetzelfde'. Dit lijkt op criteria als empirische steun en interbeoordelaarsovereenstemming. Er waren ook pragmatici: 'het helpt je als je het weet', 'als ik het weet, weet ik wat ik moet doen'. Er waren zelfs enkele jonge sceptici: 'dat weet je nooit' en intuïtionisten: 'ik weet het gewoon'; 'daarom'. De kinderen gebruikten ook kenmerken van de boodschapper als criterium voor validiteit: 'hij trekt zich nergens iets van aan, hij is zorgeloos, hij vertelt het zonder blozen, hij krijgt een rood gezicht, hij is serieus, hij vertelt het zonder te stotteren'. De laatste antwoorden kwamen meer voor bij de 8- dan 12-jarigen (Van Houdt<sup>98</sup>).

Onze studenten zijn verder dan deze 8- tot 12-jarigen. Ze leren het concept niet alleen door boeken te lezen en tentamens af te leggen. Waarom hebben we het er nooit over met hen? Er is onderzoek genoeg (Perry<sup>99</sup>) en in zijn navolging Kingston et al. en Hofer et al. De laatste bestudeerde de ontwikkeling van vrouwelijke studenten. Deze uitwerking van zelfbetrokkenheid lijkt zinvoller dan spreken over de pathologische categorieën die een student op zichzelf plakt naar aanleiding van de colleges klinische psychopathologie. Je valt immers altijd wel een beetje in een of twee DSM categorieën en hebt ook een plaats op een of twee categorieën van de *Child Behavior Check List* (CBCL).

*Waarheid en kennisbelangen:* Waarheid is aan waarden gebonden. Opvattingen van leken en professionals dienen een kennisbelang. Het belang kan beheersing/controle, het verwerven van een sociale positie, het verkrijgen van begrip of de emancipatie van de ander zijn. Er is ook een waarheid van de economische en sociale jungle. Dat belang ligt op steeds de loer in een neo-liberale utilitaristische, hedonistische samenleving. Als ik zeg: het ligt op de loer, impliceer ik, dat je dat belang moet vermijden. Dat beweer ik, als ik - dat zeggende - weggezet word als moraalridder of als 'sanft', 'soft'. Ik mag me afvragen: 'Moet ik U als wetenschapper serieus nemen, als u reclame maakt voor uw eigen winkel en uw werk in dienst staat van uw reputatie'? Moet ik geld geven aan de Maag-darm-lever Stichting ('het maag-lever-darm-alarm'), omdat geld uw noodzakelijk onderzoek mogelijk maakt naar oorzaken en effectieve medicijnen? En, aan de Hart- en Hersenstichting? 'Noodzaak' wordt in radiospotjes letterlijk genoemd. Voelt de bestuurlijke gezondheidszorg-elite die noodzaak dan niet? Moet het Koningin Wilhelmina Fonds mensen een berg laten beklimmen om 'noodzakelijk onderzoek' te laten doen? Zonder die tocht geen genezing? Of gaat het om belangen, zoals het overleven van onderzoeksinstellingen, werkgelegenheid en het instandhouden van

een structuur met een door de samenleving gewaardeerde en een goedbetaalde directie? De moraalfilosoof Williams<sup>100</sup> ziet sporen van dit soort belangen bij sociale wetenschappers: Macht en reputatie zijn: ‘... *one of the reasons, why, at the present time, the study of humanities runs a risk of sliding from professional seriousness, through professionalization to a finally disenchanted (betovering wegnemend, ontluisterend, ontgoochelend) careerism*’. Hij is bekend vanwege zijn ironische blik op het utilitarisme dat volgens hem uitgaat van een onpartijdige waarnemer die alwetend, belangeloos en zonder emotie is, maar verder een ‘normaal mens’. In Reves brieven en romans ‘kwam nooit een normaal mens voor’, zei hij naar aanleiding van het Ezelproces, aangespannen door senator Algra. Williams onderscheidt naast *accuracy*: de objectieve, valide beschrijving *sincerity*: oprechtheid. Dat houdt onder meer in het afwijzen van *free riders*, het anderen niet willen bejegenen vanuit eigenbelang en in staat zijn tot geven en ontvangen. Deze begrippen impliceren weerstand tegen *wishful thinking*, zelfbedrog en fantasie. *Sincerity* is het nastreven van objectiviteit. Dit lijkt op Royce’s et al.’s<sup>101</sup> begrip *authoritarianism*: we zijn afhankelijk van anderen om de aard en betekenis van gebeurtenissen, verschijnselen en menselijk gedrag te begrijpen en te verklaren. Die ander heeft iets te bieden als hij oprecht is en geen reclame maakt voor zijn gedachten en producten. De wetenschapper is geen lobbyist voor de wetenschap, opgevat als een onderneming. In die zin moet een universiteit *niet* ondernemend zijn. Burgers, werknemers, studenten gaan er vanuit dat de overheid, chefs/CEO’s, bestuurders en leraren/docenten waarachtig zijn en niet uit eigenbelang handelen. Je mag ze onwetendheid verwijten en dat is erg genoeg voor professionals. Williams argumenteert dat we de alledaagse waarheid met respect tegemoet moeten treden, maar daar niet kunnen stoppen. Waarheid en geldigheid kunnen immers gefabriceerd worden. De jeugdige en volwassen intuïtie dat iets niet klopt kan bedolven raken onder een zee van culturele regels en conventies van wetenschappelijke dwingende theorieën tot ideologieën, religies, sekten en winstbejag.

*Validiteit in psychologie en diagnostiek*: Al bestaat waarheid niet, we kunnen niet zonder een opvatting erover. Als je gekozen hebt, kun je vooruit. Je hoeft de ogen niet te sluiten voor andere opvattingen maar dat gebeurt gemakkelijk. Je kunt niet alle ballen tegelijk in de lucht houden. De psychologie heeft haar epistemologisch steunpunt gevonden in het logisch positivisme en de empirisch-analytische methode van onderzoek. In de onderzoeksacyclus domineert het vermijden van fouten. Er is weerstand tegen hermeneutiek, waaronder constructivisme en metafoorgebruik te scharen is. De kritische psychologie wordt niet serieus genomen omdat zij onleesbare essays heeft opgeleverd en actiegericht onderzoek aanbeveelt waarin onafhankelijke en afhankelijke variabelen verward worden.

Gegeven deze houdingen wordt het object van onderzoek recht doen een zaak van empirisch onderzoek bij representatieve steekproeven. Je zet de grillige, vage hermeneuten buiten de deur en stuurt kritisch psychologen naar de Sociale Academie oude stijl. Probleem opgelost. Dat is het eigenlijk ook wel voor onderzoekers. De keuze scheidt duidelijkheid. Je doet alsof je het weet, alsof je weet wat je moet doen, zegt de neo-Kantiaan Vaihinger<sup>102</sup>. Daar moet je niet badinerend over doen, zoals filosofen soms doen: ‘Wat jullie, psychologen als feiten

zien en objectief noemen, verraadt je naïviteit'. 'Er is geen berg feiten los van jullie kennen'. 'Je Cartesiaans-Sartreïaanse hoop op objecten/subjecten die zich ongeschonden aandienen, zoals ze zijn voor je transparante bewustzijn, is ijdel'. Het empirisch-analytisch onderzoek zit goed in elkaar. Het is verbonden met de kennistheoretische posities van empirisme en rationalisme. Het logisch positivisme is een slim en allerminst naïef empirisme/realisme. Het heeft de diagnostiek aan concepten en strategieën geholpen. Tests moeten inhouds-, construct en predictief valide zijn en experimenteel onderzoek moet de oorzaak aanwijzen. Correlationeel onderzoek moet significante en robuuste verbanden opleveren.

*Validiteitstheorie* is een groot woord voor de geschiedenis van opvattingen over geldigheid van beweringen op grond van test scores. Eerst was er predictieve validiteit: een geldige categorie, een *w a r e* score op een schaal of een aantal ware scores stellen iets voor als ze criteriumgedrag voorspellen. Daar is inhoudsvaliditeit bij gekomen: elk item uit een populatie moet evenveel kans hebben om in een instrument terecht te komen. Een technische formulering terwijl het er om gaat in hoeverre je het voor elkaar krijgt een inhoudelijk, samengesteld domein zuiver te representeren, bijvoorbeeld 'voldoende beheersing van de Engelse taal en letterkunde', 'kennis van de natuurkunde op BA niveau', enzovoort. Cronbach en Meehl<sup>103</sup> gaven een plaats aan onderzoek dat wel iets over een construct leerde maar niet onder de twee voorgaande viel. Hun uitwerking van 1955 is nog terug te vinden in het Cotan beoordelingssysteem van 2012. Campbell en Fiskes<sup>104</sup> multi-trait multi-method analyse van 1959 is daar nog bij gekomen.

Messicks opvatting over validiteit van 1989 is bedoeld als integratie. Zijn definitie is de basis van de APA Standards van 1999: *Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment*. Een ongemakkelijke definitie want wat valt er *niet* onder? Messick (persoonlijke mededeling, Groningen, september, 1992; hij overleed in 1996) zei dat de tekst waarin deze definitie staat, de moeilijkste was die tot dan hij geschreven had. Intussen verscheen een voorlopig nieuwe versie van de Standards in 2014. Het is interessant na te gaan of deze een nieuwe weg inslaan met validiteit.

*Research validiteit* is ook een groot woord voor geldige conclusies uit experimenteel onderzoek. De geldigheid wordt geborgd door (quasi) experimentele en correlatieve designs en de omstreden nulhypotesetoetsing. De laatste klopt wel, maar je hebt er niet veel aan. Bij ware experimenten wordt de oorzaak van de variantie in de afhankelijke variabele door randomisatie onomstotelijk aangewezen: validiteit gegarandeerd:  $p < .01$ ,  $< .05$  onder de nulhypothese. Campbell en Stanley hebben vier validiteitsconcepten uitgewerkt voor quasi-experimenten. Ze beschouwen deze als gedegenererde *true* experimenten. Ze halen alles uit de kast om hun validiteit te bewaken. Ze sporen bedreigingen op voor interne, construct, externe en statistische conclusie validiteit. Een *true* experiment kan degenereren tot quasi-experiment. Campbell en Stanley geven voorbeelden van *social experimentation*: verbeter de wereld, maar doe dat met een deugdelijk onderzoeksontwerp: waak voor onjuiste conclusies over

de oorzaak. Sociaal experimenteren is veldwerk waarin je de storende variabelen niet in de hand hebt en je de onafhankelijke niet willekeurig kunt doseren. De auteurs laten het nodige kunst- en vliegwerk zien om toch maar de interventie als oorzaak aan te wijzen. Er is telkens een grote lijst met bedreigingen die afgevinkt kan worden. Ze proberen ook correlatieve ontwerpen in te lijven bij quasi-experimenten, zoals het regressie-discontinuïteits ontwerp en tijdserie analyse.

*Toch schuurt de empirisch-analytische werkwijze in de diagnostiek* Voor een diagnosticus is het verhaal zelfs met valide testcores en deugdelijke (quasi) experimenten niet af. Het object van onderzoek is een cliënt, onmiskenbaar een subject met zijn bagage aan impliciete opvattingen, ongerefecteerde achtergrondkennis en gewoonten. Hij verwacht als subject behandeld te worden en zal zich vervreemd voelen als hij gereduceerd wordt tot een normproefpersoon. En, de diagnosticus ziet zichzelf niet als inwisselbare klerk, ook al volgt hij het protocol nauwgezet. Hij hoort het verhaal van de cliënt niet uit beleefdheid aan en vormt het om naar het beeld van zijn empirisch-analytisch gewonnen begrippen en onderzoeksbevindingen. Bij een arts-specialist kom je dat soms tegen: U kunt hem (m/v) vertellen wat u wilt; het gaat hem/haar niet om wat u zegt, maar om de uitslagen van onderzoeken en zijn kennis van ziekten en organen: 'de laboratoriumuitslag zegt meer dan uw duizend woorden'.

Empirisch-analytisch geschoolde experimentatoren gaan aan de individuele proefpersoon voorbij: binnenvariantie is *error*. Zij kunnen dat niet helemaal volhouden, want in een experiment vormen de proefpersonen zelf ook hypothesen over de bedoeling van de experimentele manipulatie en handelen er naar. Klassiek testtheoretisch begrepen betrouwbaarheid wordt bepaald in een steekproef. Iedere persoon wordt eenzelfde geschatte meetfout toegekend. Is het niet zinvol na te denken over het individualiseren van de meetfout? Waarom zou die voor iedere cliënt gelijk (moeten) zijn? Validiteit wordt ook in steekproeven bepaald. De score van een cliënt kan substantieel buiten de regressielijn liggen. Bovendien houdt de diagnosticus zich niet strikt aan de protocollen. Het HTM wordt niet helemaal als bedoeld uitgevoerd.

Er komt bij dat cliënt en diagnosticus communiceren. In communicatie gelden Grice's regels van relevantie en non-redundantie. Als een cliënt niet inziet wat zijn antwoorden op de Rorschach platen over zijn persoon/gedrag te vertellen hebben, dan is er het probleem van relevantie. Als de cliënt zegt dat hij steeds dezelfde stellingen/uitspraken moet beantwoorden in een vragenlijst, beroept hij zich op de regel van non-redundantie. Ook dit is een vorm van terugpraten. De opvatting over de relatie diagnosticus-client en die tussen experimentator en proefpersoon varieert van een niet-Ik en een Ik (proefpersoon en proefleider) tot twee Ikken (onderzoeker en mede-onderzoeker).

De diagnosticus moet de balans vinden. Je kunt niet zeggen tegen de cliënt: 'u weet het even goed als ik'. Maar ook niet: 'u hebt geen zicht op uw probleem, luister'. De cliënt als mede-onderzoeker klinkt wel sympathiek maar vertegenwoordigt een beperkt idee over waarheid/geldigheid: consensus. Beperkt, omdat het gaat over consensus tussen niet gelijkwaardige partijen: je doet 'alsof' je gelijk bent. Ooit een therapeut meegemaakt die zijn honorarium van uw Ziektekostenverzekering met u deelt? Deze cliënten zullen welbespraakt

en goed opgeleid zijn, maar ze zijn niet gelijkwaardig in kennis en behandeling, net zo min als U dat bent, wanneer U ziek naar de dokter gaat. Mondigheid heeft zo zijn grenzen.

*Subjectiviteit - Objectiviteit:* Bij tests ligt de nadruk op objectief scoren van antwoorden. Er is bij tests daarom terughoudendheid om redenen voor foute en goede antwoorden mee te wegen. Er sluipt gauw een subjectief oordeel in de scoring. In het (quasi) experiment is subjectiviteit een storende, vertekende factor voor de conclusie dat er een effect of samenhang is. In het ware experiment is subjectiviteit de binnenvariantie: individuele verschillen tussen de experimentele eenheden. In het quasi-experiment komt dat terug als de poging onjuiste interpretaties uit te sluiten over de relatie tussen de ingreep en het effect: interne validiteit.

*Objectiviteit als recht doen aan het probleem/de vraag van de cliënt.* Recht doen aan vraagt open te staan voor betekenissen van het gedrag van de cliënt. Dat bestaat onder meer uit zijn antwoord op taken, interviewvragen, testitems, zijn scores op afhankelijke variabelen in (quasi) experimenten. De betekenis van objectief scoren en van experimentele effecten is duidelijk, maar hoeft niet te gelden voor de cliënt. Hij is geen steekproef. Bovendien interpreteert hij de vragen en testitems. Dat geldt ook voor de onafhankelijke variabele die de experimentator manipuleert. De persoon heeft een vermoeden wat de proefleider wil met de onafhankelijke variabele, de ingreep, het programma, de interventie, de behandeling. De cliënt interpreteert het doel van de therapie, interventie, training of les en hij vraagt zich af wat de onafhankelijke variabele bij hem moet bewerkstelligen. Een tijdje is dat serieus genomen en werd *unobtrusive research* aanbevolen. Nu hoor je er nauwelijks meer over. Subjectief-objectief vormt een 'modderige tweedeling'<sup>105</sup>. Ze kan van alles betekenen en je moet uit de context opmaken waar ze betrekking op heeft. Objectief wordt soms democratisch opgevat als overeenstemming tussen iedereen of als het gemiddelde van allen. Een meritocratische interpretatie is overeenstemming tussen een aantal rationale, weldenkende, objectieve, terughoudende experts. Objectief kan ook verwijzen naar betrouwbaarheid en validiteit; subjectief wordt dan onbetrouwbaar en niet valide. Het klinisch oordeel is volgens Dawes et al.<sup>106</sup> bijvoorbeeld subjectief, impressionistisch en *onethisch*.

Weer een andere kleur van de tweedeling wordt zichtbaar als je objectiviteit opvat als een houding, stijl of vaardigheid. In dat geval zijn het gewaardeerde kenmerken, deugden van terughoudendheid, openheid, authenticiteit, betrokkenheid. Objectiviteit kan ook duiden op het vermijden van een tunnelvisie: probeer meer gezichtspunten bij het object van studie te betrekken, vermijd *foreclosure*, beperk je niet tot de eerste de beste hypothese. Daar is niets op tegen maar het vertraagt onderzoek en wordt nogal eens overgeslagen. Toch komen we er nooit vanaf: we hebben geen *regard survolante*.

*Conclusie:* De diagnosticus is geïnteresseerd in de waarheid, in een geldige, valide diagnose. Hij heeft het kennisbestand van de psychologie als ruggensteun bij zijn streven recht te doen aan het probleem van de cliënt. Dat bestand is vooral binnen het logisch positivistisch, empirisch-analytisch kader verworven en uitgebouwd. Dat is omvangrijk, maar hij kan er niet om

heen: hij moet 'tegen beter weten in', 'tegen de feiten in' alles weten. Aan het object van onderzoek recht doen vraagt ook dat de diagnosticus openstaat voor betekenissen van gedrag. Hij moet bij de cliënt achterhalen waar de schoen wringt.

Barendregt gaf er een voorbeeld van in een lezing (rond 1974, ik doe dit uit mijn herinnering) over fobieën. Hij probeerde deze te duiden vanuit verschillende psychologische stromingen van de geconditioneerde reflex, gewoonten, via trekken tot de gezins- en culturele omgeving. Een sluier kan voor een fobische cliënt zowel veroorzaker als verlichter van dat lijden zijn. Hij liet merken dat het probleem van de 'patiënt' betrekking heeft meer lagen dan de gedragsmatig psychologische. Een diagnosticus kan soms niet om de existentiële en spirituele laag heen. En, ik voeg toe: hij laat zich niet voortijdig opsluiten in een epistemologische stroming. Daarin wijkt hij af van de onderzoeker. Er is niet één basis en niet één weg.

Objectiviteit en waarheid hebben vele vaders, moeders, stromingen en draagvlakken nodig. Ze zijn gekenmerkt door kennis van zaken, redelijkheid, zuiverheid, deugdelijkheid, gerechtvaardigdheid en houdbaarheid van beweringen over de sociale werkelijkheid en het gedrag van de cliënt. Als je kiest voor objectiviteit als 'aan het object van studie recht doen', kun je je niet beperken tot de methodologische empirisch-analytische kenhouding. Diagnostiek en ook psychologie kunnen zich niet ontdoen van hermeneutiek. Dat houdt in dat ook hij de idiosyncratische betekenisverlening van de cliënt serieus neemt. Een minder prominente, meer gewantrouwd stroming is de kritische. Deze staat een andere opvatting van de cliënt toe: recht doen is hem emanciperen. Er hoort een andere methode bij: het verfoeide actie onderzoek; maar wees voorzichtig: er zijn nogal wat handelingswetenschappen met serieuze en gerespecteerde beoefenaars.

Diagnostiek is evenwichtskunst, een balans vinden tussen betrokkenheid en afstand, partijdigheid en onpartijdigheid. Deze balans kent deelnemers die gelijkberechtigd zijn, waarbij ieder er naar streeft waarachtig en authentiek te zijn. Het wil niet zeggen dat de deelnemers gelijk zijn in kennis en kunde. Als dat zo was, had je geen diagnosticus of therapeut nodig. Bij Hermans' Zelf Konfrontatie Methode<sup>107</sup> lijkt het bijna zover te zijn. Kouwer heeft het gesprek centraal gesteld als werkwijze om de cliënt als persoon te beschrijven. Beiden beschrijven op hun manier objectiviteit. Bij Hermans zijn het Ik - de diagnosticus - en het niet Ik - de cliënt – gelijkwaardig: twee Ikken dus. Bij Kouwer is dat niet precies omschreven, maar hij benadrukt een relatie, waar bij de cliënt terugpraat, waardoor beiden van opvatting en inzicht kunnen en mogen veranderen: diagnostiek is zo doende het voltrekken van een dynamisch, *emergent* proces door tenminste twee deelnemers.

## Referenties

1. Heidegger, M. (1970). *Wat is dat- Filosofie?* Utrecht: Telt Utrecht.
2. Pieper, J. (2007, de zoveelste bewerking, vertaling). *For love of wisdom: Essays on the nature of philosophy*. Münster: Ignatius.
3. James, W. (1907). *Pragmatism*. New York: Harvard University Press.
4. Rorty, R. (1980). *Philosophy and the mirror of nature*. Great Britain: Basil Blackwell Publishers.
5. Simon, H.A. (1991). *Bounded rationality and organizational learning*. Online beschikbaar: Permalink: <http://dx.doi.org/10.1287/orsc.2.1.125>.
6. Duijker, H.J.C., Palland, B.G. & Vuyk, R. *Leerboek der Psychologie* (1960, 2<sup>de</sup> druk) Groningen: Wolters.
- Duijker, H.J.C. (1959). Nomenclatuur en systematiek der psychologie. *Nederlands Tijdschrift voor de Psychologie*, 14, 89-116.
7. Multatuli, pseudoniem van E. Douwes Dekker. *Ideeën* (1862-1880, deel 3 tot 10). Amsterdam: Elsevier.
8. Popper, K. R. (1968). *The logic of scientific discovery*. London: Hutchinson.
9. Meehl, P.E. (1978). Theoretical risks and Tabular Asteristics: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Clinical and Consulting Psychology*, 46, 7, 806-814.
10. Bacon, F. (1620/2000). *Novum Organon*. Uitgegeven door M. Silverthorne & L. Jardine. Cambridge University Press: Londen/ Cambridge.
11. Heidegger, M. (1927; 6<sup>e</sup> druk 1962). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag. Nederlandse vertaling SUN Nijmegen door Marc Wildschut: *Zijn en Tijd*.
12. Kant, I. (1878/1961). *Kritik der reinen Vernunft*. (Critics on Pure Reason/ Kritiek der Zuivere Rede) Edited by Ingeborg Heidemann; P. Reklam Jun. Stuttgart. (Nederlandse vertaling beschikbaar).
13. Plato, Vertaling van zijn *Verzameld werk* (1980): Baarn.
14. Dawkins, R.A. (1995, 1996). *River out of Eden. A Darwinian view of life*. New York: Basic Books: Member of the Perseus Books Group.
15. Kuhn, T.S. (1962). *The structure of scientific revolutions*. International Encyclopaedia of Unified Science. Vol. 2. Chicago: The University of Chicago Press.
16. Reese, W.F. & Overton, W.F. (1970). Models of development and theories of development. In: L.R. Goulet & P.B. Baltes (Eds.): *Life-span developmental psychology: research and theory*. (pp. 115-145). New York: Academic Press.
17. Rawls, (1971; Revised Edition, 1999). *A Theory of Justice*. New York: Harvard University Press. Nederlandse vertaling beschikbaar.
18. James, W. (1902, 1982). *Varieties of religious experience*. Hammondsworth. In 1982 printed as Penguin Book (Nederlandse vertaling, 2010: James, W. *Vormen van religieuze ervaring: een onderzoek naar de menselijke aard*. (6<sup>de</sup> druk, Amsterdam: Abraxas).
19. Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, 114, 4, 522-536.

20. Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszin, T.W., & Reed, G.M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*, 2, 128-165.
21. Brunswik, E. (1952/1958). *The conceptual framework of psychology*. Chicago: Chicago University Press.
22. Groenier, M., Pieters, J.M., Hulshof, C.D., Wilhelm, P. & Witteman, C.L.M. (2008). Psychologists' judgment of diagnostic activities: deviation from a theoretical model. *Clinical Psychology and Psychotherapy*, *15*, 256-265.
23. Van Strien, P.J. (1975). Naar een methodologie van het praktijkdenken in de sociale wetenschappen. *Nederlands Tijdschrift voor de Psychologie*, *30*, 601-619.
24. Proudhon, P.J. (z.j.). *Wat is eigendom?* Vertaald door Z. Pennings. Utrecht: Uitgeverij IJzer.
25. James, W. (1907). *Pragmatism*. New York: Harvard University Press
26. Rorty, R. (1989). *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.
- Rorty, R. (1980). *Philosophy and the mirror of nature*. Great Britain: Basil Blackwell Publisher.
27. Bergner, R.M. (2011). What is behavior? And so what? *New Ideas in Psychology*, *29*, 147-155.
28. Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 2, 158-277.
29. Rotter, J.B. (1982). *The Development and Applications of Social Learning Theory. Selected Papers*. New York: Praeger.
30. Bandura, A. (1999). Social cognitive theory of personality. In L.A. Pervin & O.P. John. *Handbook of Personality Theory and Research*. (pp. 154-197). New York: The Guilford Press.
31. Schopenhauer, A. (1999). *De wereld als wil en voorstelling. Twee delen*. Vertaald door H. Driessen. Amsterdam: Werelduitgeverij.
32. Smedslund, J. (2009). The mismatch between current research methods and the nature of psychological phenomena: What researchers must learn from practitioners. *Theory & Psychology*, *19*, 778-794.
- Smedslund, J. (1997). *The Structure of Psychological Common Sense*. Mahwah, NJ: Erlbaum.
33. De Waal, F. (2015). *Bonobo en de tien geboden*. Amsterdam: Uitgeverij Atlas Contact.
34. Hegel, G.W.F. Vertaling van 'Phaenomenologie des Geistes' als: *De zelfverantwoording van de geest*. Door J.C.G. Hahn (1980). Uitgeverij Futile.
35. De Mul, J. (2014). *Kunstmatig van nature: Onderweg naar homo sapiens 3.0*. Stichting Maand van de Filosofie. Lemniscaat: Rotterdam.
36. Sartre, J.P. (1990). *De woorden*. Utrecht: Bijleveld.
- Sartre, J.P. *Gesloten Huis*. (2004). Utrecht: A.W. Bruna. (*Huis Clos*, 1947). Paris: Le livre de poche. (hieruit de beroemde zinsnede: 'u bent vrij, kies'!)
37. Schneider, E.D. & Sagan, D. (2005). *Into the Cool*. Chicago: University of Chicago Press.
38. Schrödinger, E. (1944). *What is life?* Cambridge: University of Cambridge Press.
39. Meehl, P.E. (1954). *Clinical versus Statistical Prediction*. Minneapolis: University of Minnesota Press.

40. Grove, W.M., Zald, D.H., Lebow, B.S. Snitz, B.E. & Nelson C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19-30.
41. Aegisdóttir, S., White, M.J., Spengler, P.M., Maugherman, A.S., Anderson, LA., Cook, R.S., Nichols, C.N., Langopoulos, G.K., Walker, B.S., Cohen, G. & Rush, J.D. (2006). The meta-analysis of the Clinical Judgment Project: The fifty-six years of accumulated research of clinical versus statistical prediction. *The Counseling Psychologist*, 34, 3, 341-382.
42. Kahneman, D. & Klein, G. (2009). Conditions for intuitive experience: A failure to disagree. *American Psychologist*, 64, 4, 515-526.
43. Holt, R.R. (1970). Yet another look at clinical and statistical prediction: or is clinical psychology worthwhile? *American Psychologist*, 25, 3, 337-354.
44. Van Dam, K. (1991). *Fixatie op fouten*. Lisse: Zwets & Zeitlinger.
45. *Diagnostic and Statistical Manual of Mental Disorders*. DSM-5 (5<sup>th</sup> Ed.) (2013). American Psychiatric Association. Washington DC: APA. Nederlandse vertaling zie referentie 85.
46. Klein, G. A. (1998). *Sources of Power: How people make decisions*. Cambridge. MA: MIT Press.
- Klein, G.A. (2003). *Intuition at work*. New York, NY: Currency & Doubleday.
47. Gigerenzer, G. & Gassmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.
48. Goldstein, D.G. & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 1, 75-90.
49. Barendregt, J.T. (1977). *Karakters van en naar Theophrastes*. Deventer: Van Loghum Slaterus.
50. Van Hoesel, A.F.G. (1955). *Zindelijk Denken: Foutieve denkwijzen en oneerlijke discussiemethoden*. Bloemendaal: H. Nelissen.
51. Derksen, Ton (2010). *De ware toedracht: een praktische wetenschapsfilosofie voor waarheidszoekers*. Amsterdam: Veen Magazines.
52. Toulmin, S. (1958, vernieuwde versie: 2003). *The uses of Argument*. Cambridge: Cambridge University Press.
53. Perelman, Ch. & Olbrechts-Tyteca, L (1958). *Traité de l'argumentation*. Brussel: Presses Universitaires. Dit is een boek waar de klassieke retorica benut wordt om zindelijk te argumenteren. Het is van 1958 en door de Universiteit Brussel uitgegeven. De laatste versie is van 2009. Dit klassieke werk is als pdf te downloaden van het internet. De eerste zin uit het boek toont het afstand nemen van het dualisme en de bewijsvoering van Descartes: 'La publication d'un traité consacré à l'argumentation et son rattachement à une vielle tradition, celle de la rhétorique et la dialectique grecques constituent une rupture avec une conception de la raison et du raisonnement, issue de Descartes qui a marqué de son sceau la philosophie occidentale de trois dernières siècles'.
54. Brown, N.J.S., Furtak, E.M., Timms, M. Nagashima, S.O. & Wilson, M. (2010). The Evidence-Based Reasoning Framework: Assessing Scientific Reasoning. *Educational Assessment*, 15, 123-141.

55. Wools, S. (2015). *All about validity: An evaluation system for the quality of educational assessment*. Dissertatie TU Enschede. Enschede: Ipskamp Drukkers.
56. Weinstein, M.C. & Fineberg, H.V. (1980). *Clinical Decision Analysis*. Philadelphia: W.B. Saunders.
57. Fernández-Ballesteros, R., De Bruyn, E, Godoy, A. Hornke, L, Laak, J. ter, Vizcarro, C., Westhoff, K. Westmayer, H. & Zaccagnini, J. (2001). Guidelines for the Assessment Process. *European Journal of Psychological Assessment*, 17,178-191.
58. Westhoff, K., Hagemester, C. & Strobel, A. (2007). Decision-Aiding in the Process of Psychological Assessment. *Psychological Science*, 49, 3, 271-285.
59. Kirk, R.E. (1968). *Experimental Design: procedures for the behavioral sciences*. Belmont, A Brooks/Cole Publishing Company.
60. O'Brien, W.H., Collins, A. & Kaplan, M. (2003). Case formulation. In: R. Fernández-Ballesteros (Ed.). *Encyclopedia of Psychological Assessment Vol. I* (pp. 164-167). London: Sage Publications.
61. Cronbach, L. J. (1955). Processes affecting scores on 'understanding of others' and assumed similarity. *Psychological Bulletin*, 52,177-179.
62. Funder D.C. & West, S.G. (1993). Consensus, Self-other agreement and Accuracy in Personality Judgment: An Introduction. *Journal of Personality*, 67, 4, 457-476.
63. John, O.P. & Robins, R.W. (1994). Accuracy and bias in self-perception: Individual differences in self- enhancement and narcissism. *Journal of Personality and Social Psychology*, 66, 206-219.
64. Letzring, T.D. (2008). The good judge: characteristics, behaviors, and observer accuracy. *Journal of Research in Personality*, 42, 6, 914-932.
65. Siegler, R.S. & Svetlana, M. (2002). A micro-genetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development*, 73, 793-809.
66. Barnes, K.L. & Moon, S.M. (2006). Factor structure of the Psychotherapy Supervisor Development Scale. *Measurement and Evaluation in Counseling and Development*, 39, 1, 130-140
67. Heyting, A. (Ed.). (1975). *L.E.J. Brouwer Collected works I: Philosophy and Foundations of Mathematics*. Elsevier: Amsterdam/New York. De wiskundige Lagendijk heeft een biografie - sort of - geschreven over deze beroemde Nederlandse intuïtionistische wiskundige.
68. Gotlieb, A. (2014). *Operatie 'werk Arthur de deur uit'*. Onder redactie van Joep Dohmen & Jeroen Wester. Amsterdam: Bertram+deLeeuw Uitgevers.
69. Thurstone, L.L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.
70. Hartshorne, H. & May, M.A. (1928). *Studies in Deceit*. New York: McMillan.
71. Mischel, W. (1968). *Personality and Assessment*. New York: J. Wiley.
72. Tupes, E.C. & Christal, R.C. (1961). *Recurrent personality factors based on trait ratings*. (Tech. Rep. No. ASD-TR-61-97). Lackland Air Force Base, TX: U.S. Air Force. In 1992 herdrukt en toen pas algemeen bekend geworden.

73. John, O. P. (1990). The 'Big Five' Factor taxonomy: Dimensions of personality in natural language and in questionnaires. In L.A. Pervin. *Handbook of Personality: Theory and Research*. New York: The Guilford Press.
74. McCrae, R.R., Costa, P.T. (1985). *The NEO-PI Manual*. Odessa, FL: Psychological Assessment Resources.
75. Witkin, H.A. & Goodenough, D.R. (1977). Field dependency and interpersonal behaviour. *Psychological Bulletin*, 84, 4, 661-689.
76. Allport, G.W. & Odbert, H.S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs*, 47, 1 (No. 211).
77. Cattell, R.B. (1943). The description of Personality: Basic traits resolved in clusters. *Journal of Abnormal and Social Psychology*, 58, 69-90.
- Cattell, R.B. (1965). *The scientific analysis of personality*. Hammondsworth: Penguin.
78. Schmitt, N. (2014). Personality and cognitive ability as predictors of effective performance at work. *Annual Review of Organizational Behavior*, 1, 45-65.
79. McCrae, R.R., Zonderman, A.B., Costa, P.T, Bond, M.H. & Paunnonen, S.V. (1996). Evaluating replicability of factor in the revised neo personality inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552-566.
80. Hofstee, W.K.B. & De Raad, (1991) Persoonlijkheidsstructuur: De AB5C taxonomie van Nederlandse eigenschapstermen. *Nederlands Tijdschrift voor de Psychologie*, 46, 262-274.
81. Heine, S.J. & Buchtel, E.E. (2009). Personality: the universal and the culturally specific. *Annual Review of Psychology*, 60, 3, 369-394.
82. McCrae et al. (2013). The inaccuracy of national character stereotypes. *Journal of Research in Personality*, 47, 831-842.
83. Rushton, J.P. & Jensen, A.R. (2010). The rise and fall of the Flynn effect as a reason to expect a narrowing of the Black-White IQ gap. *Intelligence*, 38, 213-219.
84. Lynn, R. (2011). *The chosen people: A study of Jewish intelligence and achievements*. Washington: Summil Publishers Augusta GA978-1-59368-036-7408.
85. *Handboek voor de classificatie van psychische stoornissen van 2014* (vertaling van de DSM-5, 5<sup>de</sup> Editie onder leiding van M. Hengeveld: Amsterdam: Boom).
86. Sackett, P.R. & Walmsley, Ph.T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, 9, 5, 538-551.
87. Holland, J.L. (1997). *Making Vocational Choices. A theory of vocational personalities and work environments (3<sup>rd</sup> Edition)*. Odessa FL: Psychological Assessment Resources.
88. Poropat, A.E. (2009). A meta-analysis of the Five-Factor model of Personality and Academic Performance, *Psychological Bulletin*, 135, 2, 322-338.
89. Gadamer, H. (2014). *Waarheid en methode. Hoofdpijnen van een filosofische hermeneutiek*. (Vertaling van 'Wahrheit und Methode' (1960) door Mark Wildschut. Nijmegen: SUN.
90. Roland Holst, A. (1967). *In ballingschap: Keuze uit eigen werk herzien en vermeerderd*. Den Haag: Bert Bakker/Daamen N.V. Het citaat is uit het gedicht 'Onweer' (pp. 66-67).
91. Vroman, Leo. (1964). *126 gedichten*. Amsterdam: Querido. (de aangehaalde regels zijn uit het gedicht 'Voor wie dit leest' (p. 95).

92. Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D.B. Harris (Ed.). *The concept of development: an issue in the study of human behavior*. Minneapolis: University of Minnesota Press. (pp. 125-148).
93. Raeff, C. (2011). Distinguishing between development and change: Reviving Organismic-Developmental Theory. *Human Development*, 54, 4, 4-33.
94. Heymans, P.G. (2002). Ontwikkeling van de persoonlijkheid en ontwikkeling over de levensloop. In J.J.F. Schroots: *Handboek van de volwassen ontwikkeling & veroudering*. Assen: Van Gorcum (pp. 251-299). Heymans, P.G. (2000). Talent for development: responding to contextual promises. In C.F.M. van Lieshout en P.G. Heymans (Eds.). *Developing talent across the life-span*. Sussex: Hove (pp. 151-167).
95. Peters, J. (1957). *Metaphysica. Een systematisch overzicht*. Uitgeverij Spectrum. Utrecht / Antwerpen.
96. *The Oxford Companion to Philosophy, New Second Edition*. Edited by Ted Honderich (2005). Oxford: Oxford University Press.
- Johannes Hirschberger (1964): *Kleine Philosophiegeschichte 4. Auflage*. Freiburg: Herder Bücherei.
- Hans Joachim Störig (2002, 26<sup>ste</sup> druk). *Geschiedenis van de filosofie*. Utrecht Spectrum. Vertaling van *Kleine Weltgeschichte der Philosophie*. Stuttgart: Kohlhammer-Verlag.
- Kenny, A. (1998). *A brief history of Western Philosophy*. Malden Massachusetts: Blackwell Publishers Inc.
97. Emily Dickinson<sup>97</sup> (Fr 1263, 1872; *Verzamelde gedichten*. Vertaling: P. Versteegen, 2011. Amsterdam: G.A. van Oorschot. (Het geciteerde gedicht is op het internet te vinden).
98. Houdt, A. van (1994). *Waarheid rondom de leugen van het kind*. Doctoraalscriptie Universiteit Utrecht, Vakgroep Ontwikkelingspsychologie.
99. Perry, J. (1970). *Forms of intellectual and ethical development in the College Years. A Scheme*. New York: Academic Press.
100. Williams, Bernard (2002). *Truth and Truthfulness, an essay in genealogy*. Princeton and Oxford; Princeton University Press.
101. Royce, J.R., Coward, H., Egan, E., Kessel, F. & Mos, L. (1978). Psychological Epistemology: A critical review of empirical research and the theoretical issues. *Genetic Psychology Monographs*, 97, 265-353.
102. Vaihinger, H. (1911, 1986). *Die Philosophie des Als Ob*. (The philosophy of the 'As If') (Nachdruck /1<sup>st</sup> Edition, 1927). Hamburg: Meiner Verlag.
103. Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
104. Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multi-trait-multi-method matrix. *Psychological Bulletin*, 56, 1, 81-105.
105. Gigerenzer, G. (2000). *Adaptive thinking: Rationality in a real world*. Oxford: Oxford University Press.
- Gigerenzer, G. (2008). *Rationality of Mortals: How people cope with uncertainty*. Oxford: Oxford University Press.

106. Dawes, R.M., Faust, D. & Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.

107. Hermans, H.J.M. (1974). *Waardengebieden en hun Ontwikkeling*. Amsterdam: Swets & Zeitlinger.

De recentste en (enige?) experimentele studie om trekken te veranderen is van Hudson, N.W. & Fraley, R. Chr. (2015) Volitional Personality Trait change: Can people choose to change their personality traits? (*Journal of Personality and Social Psychology*, vol. 109, 3 pp. 490-507). Het kan volgens de auteurs, als je maar wil. Het ging niet gemakkelijk, een interventie werkte niet, de andere wel.

### III Dynamiek van het kennisbestand

*Worden psychologische theorieën regelmatig herzien? Wordt kennis bijgewerkt, bij de tijd gebracht, tot 'kennis van nu' gemaakt? Kunnen we af en toe opruiming houden? Wordt een theorie niet vergeten en ze keert later terug? Moeten we op grond van ons voortschrijdend inzicht constructen toevoegen, nuanceren en herinterpreteren? Als we alles wat we uitgezocht hebben samenvatten, kunnen we dan afdoende antwoorden geven op specifieke vragen? 'Meta-analysis is meta-silliness', zei Eysenck ooit. Is dat te kort door de bocht? Kunnen we bevindingen zonder meer toepassen in de individuele diagnostiek? Veranderen gedragingen niet zo door de tijd heen dat de diagnose er haastig achteraan moet lopen, bijvoorbeeld bij het Flynn effect? Onderzoeksresultaten van diagnostiek en van therapie, interventie, 'treatment' worden samen benut in de Diagnose-Behandeling-Combinatie. Zijn DBCs het sluitstuk van succesvol toepassen van de kennis uit deze twee activiteiten?*

## 1. Herzien: vervangen, opruimen, erosie, vergeten, muteren en terugkeren

*Vervangen:* Kom je meer te weten door een theorie die niet voldoet te vervangen als gevolg van anomalieën in die theorie? Dat is Popperiaans uitzuiveren. De oude ruimt het veld voor een nieuwe, adequatere theorie na het verrichten van cruciale experimenten. Bij Popper was één zwarte zwaan genoeg om de uitspraak: 'Er bestaan witte zwanen' omver te gooien. Dit ontleende hij aan de Aristotelische logica. Hij stelde het cruciale experiment voor om de ene verklaring tegen de andere weg te strepen. Kennisverwerving als een serieel binair proces. Het past in de logisch positivistische, empirisch-analytische traditie. De Groot en Linschoten sloten daar bij aan, maar niet helemaal. De Groot had door zijn belangstelling voor denkpsychologie oog voor het hermeneutische proces van interpretaties bedenken en de creatieve uit-de-duimzuig-beginfase aan het begin van empirisch onderzoek: de *context of discovery*. En je kunt/moet na elke empirische cyclus weer van voor af aan beginnen met de nieuw ontstane vraag. Methodologisch was hij overwegend een empirisch-analyticus met nadruk op het voorkomen van vertekeningen in onderzoek en oordelen.

Zijn er voorbeelden van cruciale experimenten die een theorie omver blazen? Kan een theorie zo belast raken door met haar strijdige feiten dat zij uitgediend, niet nuttig meer is? Als we kijken naar resultaten van nulhypotesetoetsing - toch een procedure om het bestaan van een verband of effect vooralsnog niet te verwerpen - dan blijkt bijna iedere studie in de psychologie te slagen: 85% van de hypothesen worden aanvaard. In de psychiatrie is het percentage zelfs 92% (Fanelli<sup>1</sup>). Ze stelt vast dat tussen 1990 en 2007 het percentage studies met positieve resultaten met 22% is toegenomen. Wijst dit op inflatie van onderzoeksresultaten? Of moeten we er uit afleiden dat psychologen steeds veiligere of betere hypothesen formuleren? Of, worden ze beter in het stellen van vragen, van onderzoekbare vragen, opgesplitst in deelvragen? Publiceren ze resultaten ook als effecten, samenhangen ontbreken? Gaan ze ervan uit dat negatieve resultaten niets leren over ons gedrag? Je zou ook kunnen zeggen: die samenhang is er niet, of dat effect van deze onafhankelijke variabele kun je vergeten. Dat is interessante informatie, ook volgens Popper. Of stellen we te eenvoudige en tegelijkertijd te hoge eisen? We willen immers een definitief antwoord op een vraag.

*Opruimen:* Kunnen we de theorieën van bijvoorbeeld van Adler, Jung en Freud bij het oud vuil zetten? Hebben we de *sibling rivalry*, de *animus-anima* tegenstelling en het onbewuste driftleven van het *Es* genoeg uitgeplozen, of zijn we het zat? Raken we door nieuwe inzichten de oude, onjuiste kwijt als rammelende ballast? Is de thomistische beschrijving van situaties die emoties uitlokken ingehaald door Lazarus<sup>2</sup> met zijn 14 emotie-situatie koppelingen of door Frijda's monumentale boek<sup>3</sup>? Ze lijken nogal op elkaar. Zijn de vier persoonstypen van Hippocrates onjuist en overklast door het huidige temperamentsonderzoek?

Het ziet er naar uit dat schoonmaak moeizaam is. Onze handboeken blijven de studenten alle theorieën voorhouden, al is het onder het mom van het leren kennen van het historisch verloop van de theorievorming. Je kunt verwachten en begrijpen dat ze zeggen: 'Wat moeten we ermee; is alles ongedifferentieerd waar, wanneer kom jij wetenschapper/docent eindelijk

eens met heldere en onderscheiden ideeën? Docenten gebruiken meestal Amerikaanse *undergraduate* handboeken. Deze zijn bedoeld voor studenten die een cursus psychologie willen volgen om iets over zichzelf te leren. Mogen de docenten het zichzelf niet wat moeilijker maken en bijgevolg ook onze studenten?

*Erosie*: Soms is een theorie te dun, te zwak, ze *fades out*. Je hoort er niet meer over. Erosie treedt ook op als een construct zozeer bedolven, ondergesneeuwd raakt, dat je door de draadjes van verschillende dikte het construct zelf niet meer ziet. Alles hangt met alles samen, vertelde de Nijmeegse psychiater Prick op zijn druk bezochte colleges voor a.s. psychologen en medici. Je waagde het niet dat - in zijn bijzijn - te betwijfelen. Komt een dergelijk construct niet om onder zijn eigen gewicht van een teveel aan verbanden van wisselende sterkte? Zou dat het lot kunnen zijn van de BF? *'Komt ze niet om onder haar eigen dromen terugvlaag op haar zelve'*? Dit zegt Roland Holst mooi, of zo u wilt, ronkend.

*Vergeten*: Een opvatting, theorie, verschijnsel of effect wordt soms vergeten. Wie werkt Piagets stadia nog verder uit, wie onderzoekt de overgang van pre- naar concreet operationeel nog, wie verdiept zich nog in verschijnselen uit de Gestaltpsychologie, wie onderzoekt nog het Zeigarnik effect? Wie heeft het nog over veld-(on)afhankelijkheid (Witkin<sup>4</sup>). Hij had van één experiment, waarin een loodrecht geplaatste pijl in variabel frame geplaatst was, een hele persoonlijkheidstheorie gemaakt. De veld-onafhankelijken zijn niet erg geïnteresseerd in anderen en geven de voorkeur aan *non social situations*. Hebben we die resultaten niet meer nodig; zijn ze ingehaald door betere studies? Of ligt het in de kast onder het stof en kan het er later weer uitgehaald worden?

*Muteren en Terugkeren*: Een theorie ondergaat soms *for better or worse* een gedaanteverandering. Skinners functionele gedragsanalyse maakte de weg vrij voor *case descriptions*. In uitgewerkte vorm hebben ze de aard van een theorie over het gedrag van een persoon/cliënt. De Rorschachtest is terug van weggeweest in de VS en Sterns persoonsleer is weer populair in vooralsnog kleine kring (Lamiell<sup>5</sup>). En wat gaat er gebeuren in de opkomende economieën? Jung en Heidegger zijn bijvoorbeeld in kleine kring in India en China populair. Wetenschappers en practici komen er voor naar Nederland om door hogescholen en andere instellingen georganiseerde cursussen te volgen. Deze halen Amerikanen en Duitsers naar Nederland om dat onderwijs te verzorgen.

*Conclusie*: Vernieuwing in de zin van vervanging van een theorie door een betere is zeldzaam in de psychologie. Er zijn geen revoluties, waar koppen vallen. Er zijn geen cruciale experimenten. Steeds meer hypothesen halen weliswaar de streep maar of dat vernieuwing brengt in de zin van nieuwe verbanden, effecten, inzichten, kun/mag je betwijfelen. Dit type vernieuwing zou wel eens een wetenschaps sociologisch verschijnsel kunnen zijn, dat wil zeggen mode of de prestatie van een groep die kans ziet een thema en de waarheid over een thema te claimen. We blijven onze studenten alles voorhouden. Hun bezwaren zijn terecht. Inleidende boeken

bevatten regelmatig een verzameling losse onderwerpen. Er is wel wat erosie, soms verwaait een theorie, wordt ze vergeten of muteert ze. Zou de tijd heilzaam zijn en helen? Hoe zou deze stand van zaken kunnen veranderen? Of zijn we te ongeduldig? Er staat immers niet om de paar jaar een echte vernieuwer op. Valt er iets te leren van de geschiedenis van de filosofie en van die van de exacte wetenschappen? Daar zijn wel markeringspunten aan te wijzen. Of is de psychologie nog te jong om aan echte vernieuwing toe te zijn? Zou deze metafoor passend kunnen zijn: vat het vele onderzoek op als smalle lichtbundeltjes die op een groot gebied schijnen. Je ziet telkens een deel. Het opnieuw richten van de bundel laat steeds vanuit een ander perspectief hetzelfde zien: je ziet het verschijnsel van meer kanten. Je kunt geluk hebben en iets nieuws ontdekken. Er komt iets bij. Beide leiden niet tot het wegdoen van oude kennis, maar tot nuanceren en iets nieuws toevoegen.

## 2. Aanvullen: constructen toevoegen, nuanceren, herinterpreteren

Gedragconstructen zijn de molaire eenheden waar we het in de diagnostiek mee moeten doen. Voortschrijdend inzicht kan ertoe leiden dat een nieuw construct nodig is om een nog onvoldoende belicht facet van gedrag te beschrijven. Meer kennis over een construct kan inhouden dat het genuanceerd moet worden. Wat is de reikwijdte, is het te snel gelanceerd, of kun men er nog een tijd mee vooruit als onderzoeker en diagnosticus? Inzicht kan ook tot uitdrukking komen door een nieuwe interpretatie van een bekend construct. Het is mogelijk dat je op het verkeerde spoor zit met de interpretatie. Dat spoor kan gevolgd zijn vanuit een ondoordachte vooronderstelling. Sommige constructen bevatten de onbewezen connotatie dat het goed is voor U zo gekenmerkt te zijn of U zo te gedragen. Kun je die constructen niet omgekeerd lezen? Of, is dat vrijblijvende hermeneutiek waar logisch positivisten zich niet in herkennen? Geeft het meer inzicht en een genuanceerdere kijk op een gedrag of effect?

Als voorbeelden van nieuwe constructen heb ik voor Positieve Psychologie (PP) en Post Traumatische Groei (PTG) gekozen. Niet toevallig, omdat ze me - vooringenomen als ik ben - vreemd voorkwamen. Een psycholoog is niet aangesteld om als evangelist goed nieuws te brengen. Als voorbeeld voor 'nuanceren' kies ik het Flynn effect. Ik heb er moeite mee dat te accepteren: zo'n grote vooruitgang in IQ per 30 jaar? Het is populair en bijgevolg zo uitbundig onderzocht dat je door de bomen het bos niet meer ziet. Ik probeer er orde in te brengen door hetzelfde vraag-antwoord-procedé toe te passen als bij de Big Five. Hebben we daar nu een helder beeld van? Als voorbeelden van herinterpreteren beschrijf ik afwijkende lezingen van depressie, hechting en psychopathie. Bij de keuze voor deze constructen gold het Sartrianse: ik ben vrij, ik moet kiezen! Dat is geen gemakkelijke of vanzelfsprekende stap want het had evengoed anders gekund.

**Constructen** *Positieve Psychologie* (PP) is een recent toegevoegd samengesteld construct. Deze vernieuwing gaat gepaard met aansprekende bewoordingen. Ze wordt als *noch nicht da gewesen* aan de man gebracht. Redenen zijn mogelijk: optimisme en weerstand tegen negatieve interpretatie van gedrag als pathologisch en terughoudendheid om 'het project mens' als mislukking en het individu als *loser* te beschouwen. Daarnaast kan het ook voortkomen uit het afstand nemen van de medische lezing van gedrag. Daar ben je immers óf ziek óf oninteressant. PP wil een nieuwe kijk op de persoon bieden want er is te vaak, te gauw vanuit pathologie en ziekte naar zijn gedrag gekeken.

*American Psychologist* (Seligman<sup>6</sup>) heeft edities aan PP gewijd. De benaming is volgens mij een categoriefout. Het doel van onderzoek is niet om producten te leveren die een goed gevoel of mooi resultaat geven. Bevat PP nieuwe theorie en methode of is zij een uitdrukking van verlangen naar goed nieuws? Er wordt verwacht dat de ermee gepaard gaande houding voordelen oplevert, bijvoorbeeld voor de kwaliteit van leven, werk en relaties. Seligman et al. propageren dit concept en proberen het in te zetten om Amerikaanse soldaten weerbaar te maken tegen de verschrikkingen van de oorlog. Dit wordt in de *American Psychologist* beaamd door een stafchef, generaal Casey, Jr.<sup>7</sup>. Selectie en weerbaarheid van soldaten zijn vanaf het

begin thema's in de Amerikaanse psychologie geweest. Hier staan psychologen en overheid dicht bij elkaar.

PP klinkt goed in tijden van nood en als men de richting even kwijt is. Er wordt weinig nieuws toegevoegd aan het constructenarsenaal. Je komt met *quality of life*, BF, MMPI, copingstijlen en andere vragenlijsten een heel eind. Het construct staat dicht bij levensfilosofie: *eudamonia*, het goede leven dat maar niet wil komen. Ondanks het vele onderzoek in tijdschriften (meer dan 20.000 artikelen) wordt er geen nieuw facet aan de persoon, zijn denken en sociale context toegevoegd en er worden geen kenmerken verwijderd als niet valide. Dit is te vinden in hun eigen *Journal of Positive Psychology: dedicated to the furthering research and promotig good practice* (8:3, 207-221). De auteurs Rusk en Waters hebben 18.000 documenten gevonden die tot PP behoren en de impact factor voor PP was in 2011 2,64. Er zijn er 2300 artikelen onderweg.

*Post Traumatische Groei*: Een intuïtieve en voordehandliggende gedachte is dat trauma's U uit het lood slaan. U komt ze pas langzaam weer te boven en daarna bent U niet meer helemaal de oude. Een andere gedachte is dat trauma's U tot een weerbaarder persoon maken. Een trauma laat je niet gewond achter maar maakt nieuwe krachten los. Een positieve kijk op lijden is oud, denk het boek Job uit het Oude Testament van de christelijke Bijbel<sup>8</sup>. Het Boeddhisme is er van doortrokken. Schopenhauer<sup>9</sup> ontleende zijn kijk op de onervulbare Wil en het lijden als gevolg daarvan aan het Boeddhisme. Deze opvatting leeft ook in gematigde vorm in de psychologie. McAdams en McLean<sup>10</sup> zien mild en onontkoombaar lijden als mogelijkheid tot groei. Het is de volkswijsheid dat je groeit bij tegenslag: *and a pinch to grow an inch*. Personen die kans zien zin in hun lijden te ontdekken, tonen een hoger niveau van welzijn en rijpheid: *per aspera ad astra*.

Calhoun en Tedeschi<sup>11</sup> gaan een stap verder en spreken over *Post Traumatic Growth* (PTG). Dit gaat niet over een mild lijden maar over het overwinnen van een ernstig trauma. PTG staat voor de ervaring van een positieve verandering als resultaat van een heftige levenscrisis. Ze veronderstellen dat een dergelijke crisis positieve consequenties heeft, zoals de waardering voor het leven en sociale relaties, toeneming van persoonlijke kracht, verandering in prioriteiten en een rijker spiritueel leven. De in Jemen gegijzelde Nederlandse Judith Siegel en partner reageerden ongeveer zo op hun gedwongen verblijf. Of dit een duurzame respons of een momentane reconstructie van een absurde ervaring is, zal moeten blijken. Men gaat ervan uit dat deze groei tot uitdrukking komt in toenemende wijsheid en een adequater levensverhaal dan voorheen. Sommige auteurs verwelkomen dit concept en eindigen hun essays met de obligate opmerking: *more research has to be done* (*Psychological Inquiry*, 2004, 15, pp. 19-80).

Groei lijkt hier de omkering van wat er doorgaans gebeurt na een insnijdende levensgebeurtenis. De auteurs hebben het over een beperkt deel van die gebeurtenissen. Er zijn er vele en zeer verschillende. Het is niet te verwachten dat alle tot groei leiden. Therapeuten hebben het over 'een plaats geven' aan een trauma. Het levert geen groei maar het wordt vastgezet, raakt ingeklonken en verweert niet ernstig/schadelijk uit. Het leven kan doorgaan.

Vreugde putten uit lijden is dubbelzinnig: de *pursuit of happiness* uit de Amerikaanse

grondwet tegenover lijden als basis voor groei. Het vormt een contrast met wat een Europese (Ierse) dichter als Dylan Thomas<sup>12</sup> (1914-1953) zegt over tegenslag en dood: *...do not go gentle into that good night, Old age should burn at close of day, Rage, rage, against the dying of the light...* (onderlijning, jtl). Hij schreef deze regels terwijl hij zijn vader voor zich zag, een veteraan van WO I die vele jonge mannen ontmenselijkt had. Het herinnert ook aan een van de eerste documenten uit de Europese literatuur: Homerus' Ilias: het boek opent met: 'Godin, muze, vertel ons van de woede van de zoon van Peleus, Achilles...'. Deze woede wordt vaak gelezen als moed en bereidheid tot actie. Christenen probeerden de woede te neutraliseren en noemden het trots: *superbia*, vroeger een zonde, nu een gebrek, als je haar niet regelmatig toont. Er wordt voor gelijkmoedigheid gepleit bij tegenslag en frustratie.

Calhoun en Tedeschi gaan een andere kant op en wijzen op *voordelen* van lijden. PTG is nieuw construct maar met beperkingen: het is de beschrijving van resultaten van gebeurtenissen. Die werken verschillend uit op personen. Er is geen vaste relatie tussen gebeurtenis en reactie. Empirisch onderzoek moet laten zien dat er een robuuste relatie is tussen trauma en persoonlijke groei. Er zijn pogingen: Seligman en Fowler<sup>13</sup> passen het construct toe op de psychologische conditie van soldaten en veteranen (zie ook Larner & Blow<sup>14</sup>). Zij rapporteren over PTG onderzoek bij enkele van de 1,6 miljoen soldaten die actief zijn geweest in de *war against terror*. Ze formuleren een model van *meaning making* en groei bij veteranen in drie fasen: gedurende de tijd vóór de militaire inzet, gedurende de inzet en na de inzet in gevechten. De betekenis van de eerste fase kan onder druk komen te staan door ervaringen in de tweede fase en dat kan resulteren in een langdurige crisis in de derde fase. Het kan het begin zijn van een pathologisch proces waarin de gebeurtenissen uit het geheugen verdreven worden. Het kan ook - en dat zien ze bij een substantieel percentage veteranen - het begin zijn van *resiliency*, veerkracht en beheersing van herinneringen aan gebeurtenissen gedurende de gevechten. De laatste fase wordt door constructieve en coherente betekenisverlening gekenmerkt.

De auteurs benadrukken groei na *adversities*. Ze zijn verbaasd dat ze maar acht studies konden vinden (tot 2004) over groei als gevolg van moeilijke omstandigheden. De meeste studies gaan over het Post Traumatisch Stress Syndroom (PTSS). Nogal wat VS veteranen (ongeveer 10% - 15%) zeggen daaronder te lijden. Er wordt gesuggereerd dat er bij een deel van hen sprake is van *malignering* om een uitkering te krijgen. Het is waarschijnlijk dat trauma overheerst en niet groei. Ongeveer 10% tot 15% ondervindt langdurig en ernstig hinder van oorlogservaringen. Er was een moderator: de hogeren in rang toonden naar verhouding minder het PTSS.

*Er is tegengeluid*: Wortman<sup>15</sup> betwijfelt of de evidentie voor het bestaan van Post Traumatisch Groei (PTG) *overwhelming* (Calhoun & Tedeschi) is. Groei is een mogelijk gevolg naast andere. Ze wijst erop dat PTG de indruk kan wekken dat mensen kampioenen, sterk, onkwetsbaar en koel moeten zijn, als ze met tegenslagen geconfronteerd worden. Deze houding belast de overlevers. Als ze niet posttraumatisch groeien zijn ze *losers*, *wallowing their pain*, *making us more complacent toward evil or harm*.

Positieve Psychologie lijkt op de tegenpartij van Koot & Bie (z.j.): 'de tegenpartij is de partij voor jou en mij'. Zo'n partij kan gedurende enige tijd stemmen halen. Het construct kan in het dagelijks taalgebruik doordringen en bijna een eis of voorschrift worden voor hoe je te gedragen: 'Je moet niet zo negatief zijn'. In overlijdensadvertenties wordt de term 'positief' regelmatig gebruikt. De overledene blijkt een positief mens geweest te zijn. Wie

houden we over als zoveel positieve mensen ons ontvallen?

Logisch beschouwd kan er een *negatieve psychologie* geformuleerd worden maar die zal er niet komen. Ze bestaat in de literatuur: Griekse tragedies, christendom, gnosis en mysterie spelen van de Middeleeuwen, Spaans-Oostenrijkse barok, romantische *spleen*, Schopenhauers en Nietzsches pessimisme en nihilisme, decadente Franse verveling: *ennui* en existentieel lijden bij Pascal en Kierkegaard. In Nederland kun je denken aan de schrijvers W.F. Hermans, G. Reve en A. Grunberg. Deze literaire traditie steunt op een wijsgerige houding, bijvoorbeeld die van Schopenhauer (zie Thomas Manns<sup>16</sup> essay van 1938/1977; denk aan Reves bewondering voor Schopenhauer). Deze ademt melancholie, ontroostbaarheid en wanhoop en lokt een verlangen naar verlossing uit. Deze kunstenaars beschrijven dit levensgevoel. Ze willen met hun geschrijf ons lot niet verbeteren of ons leren het lot te dragen. Ze wrijven ons eerder het lot, dat trapt schopt en bijt onder neus als onontkoombaar. Nietzsche<sup>17</sup> schreef aan het eind van de 19<sup>de</sup> eeuw in zijn boek *Genealogie van de Moraal* over het verheerlijken van geweld tegen zichzelf en anderen. Hij noemde zelfkwalling een van de langdurigste en breedst gedragen tradities in de westerse samenleving en verwees naar de religieuze traditie van ascese (zie ook Sloterdijk<sup>18</sup>). Bij Nietzsche hangt het troosteloze, het ontroostbare samen met: 'God is dood! God blijft dood! Waar vinden wij troost, wij moordenaars aller moordenaars'. De uitroeptekens zijn van Nietzsche of van hen die het na hem overgeschreven hebben.

*Succes van PP en PTG*: Beide raken een gevoelige snaar en dat heeft gevolg voor het ontwerpen van behandelingen. Er is al een therapie die op aanvaarding van leed uit is: de *Acceptance and Commitment Therapy*. Er zijn congressen die door veel psychologen bezocht worden. Er zijn al sessies voor beginners en gevorderden. Hoe maak je een dergelijk onderscheid? Aan de TU Twente is er een leerstoel *Positive Psychology & Mindfulness*. Er is een handboek (nu met 20% korting: zie 'De Psycholoog' van mei, 2015) over levenskunst en positieve psychologie van Bohlmeijer et al.<sup>19</sup>. Een benaming voor de leerstoel *Mindfulness* zou in het Twents ...'k loat mie 'n kop nich gek maak'n' kunnen zijn. Dicht bij Duitsland liggend, zouden de auteurs *Die Apostel mit der sanften Vernunft* genoemd kunnen worden, geduchte concurrenten voor Immanuel Kants 'praktische' en 'reine Vernunft'.

De constructen PP en PTG zijn toegevoegd aan het arsenaal aan gedragsconstructen, maar brengen geen theoretische vernieuwing. Gedrag is niet inherent positief, negatief of traumatisch. McNulty en Fincham<sup>20</sup> laten bijvoorbeeld zien dat vier ogenschijnlijk positieve processen en gedragingen: vergevingsgezindheid, optimistische, positieve gedachten en vriendelijkheid bij gehuwden zowel het welbevinden kunnen bevorderen als schaden. In slechte huwelijken voorspelden deze processen versterking van de verstoorde relatie. Ze wijzen erop dat er naar de specifieke situatie en gedragingen gekeken moet worden om een oordeel op een positief-negatief schaal te geven. Hun boodschap is: Vermijd psychologische trekken en processen te labelen als inherent, onbetwistbaar en vanzelfsprekend positief of negatief. Met insnijdende gebeurtenissen, trauma's

kan het verschillende kanten opgaan. De beschrijving van ontwikkeling van Adolph et al.<sup>21</sup> past: lineair, versnellend, asymptotisch, stapsgewijs (stadia), S-vormig, variabel en onsystematisch, *stair-climbing*, U-vormig en omgekeerd U-vormig. Zoals hiervoor gezegd, worden deze veranderingen gewaardeerd als goed - slecht of passend - niet passend, normaal - afwijkend. Je hoeft er geen nieuwe naam te verzinnen of er draai aan te geven. *Parsimoniousness* is hier een deugd. We hebben al constructen genoeg, zie elk willekeurig psychologisch woordenboek.

**Nuanceren:** Ik neem het Flynn effect als voorbeeld. We gaan vooruit zonder er iets voor hoeven te doen: beid uw tijd, het wordt u in de schoot geworpen. Het effect wordt robuust genoemd. Door de tijd heen is er vooruitgang in IQ, schoolprestaties (*Scholastic Achievement Tests*: SAT), geschiktheden (*Aptitudes*) en Socio-Economische-Status (SES). Flynn<sup>22</sup> rapporteerde al voor 1980 over de winst in IQ scores over generaties. Hij was natuurlijk niet de eerste. Lynn<sup>23</sup> wijst erop dat deze studies al vanaf 1936 gedaan zijn door Runquist. Hij vond in 1936 verschillen in intelligentie: de mediaan van de scores op de *Minnesota College Aptitude Tests* nam tussen 1929 en 1932 toe: van 33.7 naar 42.3.

Het is moeilijk (voor mij) om in een zo grote vooruitgang en groei te geloven. Zijn we zoveel malen intelligenter dan Aristoteles? Weten we het intussen beter dan Einstein? Wordt hier niet iets verward? Technologische vooruitgang is er zeker: 30 jaar geleden had vrijwel niemand een mobiel of internet. En die vooruitgang gaat misschien nog revolutionairder dan de 10 á 15 punten per generatie, maar daar we hebben daar geen maat voor die vergelijkbaar is met het IQ. We krijgen de ene technologische revolutie na de andere te verwerken. Of kun je die vergelijking niet maken? Is het lot van de mens nog altijd hetzelfde en hebben we er in de loop van de tijd uiteenlopende verhalen bij verzonnen? Dat lot wijzigt misschien helemaal niet ondanks de stijging van IQ en prestaties. Verder is er weliswaar een gemiddelde stijging, maar het verschil tussen de slimmen en de niet-zo-slimmen neemt niet af. Misschien gaat dit het Piketty effect<sup>24</sup> heten en komt er meer onderzoek naar de kloof tussen rijk en arm, hoogopgeleid en laagopgeleid en gezond en ziek.

Hoe dan ook, het Flynn effect is gezocht, gevonden en onderzocht. Om mijn scepsis te voeden/ te bestrijden nam ik kennis van empirische studies. Dat zijn er veel en ze gaan over hetzelfde. Dat kan saai zijn. U kunt de tekst tussen de vraag en het antwoord overslaan als u mijn weergave vertrouwt.

*Vraag: Hoeveel gaan we vooruit?* Flynn rapporteerde in 1987 resultaten uit veertien landen, onder meer uit Europa, de VS, Japan, Australië en Nieuw Zeeland. Hij vond *massive gains* in SAT- en IQ-scores (*Raven Progressive Matrices*, *Wechsler*, *Otis* - een test voor rekruten - en verbale- en wiskundetests) over een periode van dertig jaar. De toename liep van 1.60 punten in Noorwegen tot 20.3 punten in Japan (1 SD = 15 punten). In één generatie was er winst tussen de 5 en 25 punten. Gemiddeld genomen wint men per generatie één SD (15 punten). In 1998 rapporteerde hij verschillen tussen domeinen: cultuurarme tests (Gfluid, Raven) laten winst van rond de 20 punten zien, per formaal IQ tussen de 10 en 20 en verbaal IQ < 10 punten. Tests die dicht bij onderwerpen liggen die op school geleerd worden, bijvoorbeeld rekenkundig

redeneren, algemene informatie en woordenschat, laten weinig of geen winst zien. Dat kan aan de vernieuwing en hernormering van de tests liggen.

*Antwoord:* Er is een vooruitgang van 1.0 SD per generatie. Cohen noemt dit een groot effect. Er is variatie over landen en de vooruitgang doet zich vooral op Gfluid voor.

*Vraag: Wordt vooruitgang gemodereerd door leeftijd, IQ niveau en sekse?* Een meta-studie van Trahan<sup>25</sup> et al. omvatte 285 studies (N = 14.031) met gegevens vanaf 1951. Twee intelligentie-tests leverden een gemiddelde winst op van 2.31 IQ punten per decade (Betrouwbaarheidsinterval: BI, 95% 1.99 - 2.64). Een vergelijking met Stanford-Binet en Wechsler IQ schalen laat per 10 jaar een winst zien van 2.93 (BI, 95% 2.30 - 3.50) bij 53 vergelijkingen (N = 3.951). De auteurs vonden geen steun voor leeftijd als moderator variabele. Ook het capaciteitsniveau bleek geen moderator: er deed zich geen toe- of afname voor in het aantal hoog- en laagbegeefden in twee generaties: Wai en Putallaz<sup>26</sup> bekeken 1.7 miljoen scores van 11-13 jarigen op de SAT en twee andere prestatietests: ACT en EXPLORE van 1981 tot 2010. Het effect was ook aanwezig bij de 5% hoogste scores. Het effect werd bij meisjes en jongens gevonden. Het was vooral zichtbaar bij de rekensubtests van de SAT, ACT en EXPLORE. Hun tabel 10 (p. 338) laat veranderingen per jaar zien. Deze zijn bij meisjes relatief groter op de SAT-Composite scores (0,07), SAT-Rekenen (0,04) en EXPLORE-Engels (0,09). Meisjes behielden hun scores op de SAT-taal, maar jongens gingen achteruit (-0,10) en jongens gingen ten opzichte van meisjes meer achteruit op ACT-Wetenschap.

*Antwoord:* Er is vooruitgang, in deze meta-studie iets minder dan Flynn rapporteert (0,75 SD). Leeftijd, (hoog) IQ en sekse zijn geen moderator variabelen.

*Vraag: Is het effect beperkt tot het Westen?* De Chinese adaptatie van de WISC liet een winst van 4.53 (*full scale*) zien tussen steekproeven van 1984 en 2006. Dat is 2.06 IQ punt per 10 jaar. Het verbale IQ won 4.27 en het performale IQ 4.08 (Liu et al.<sup>27</sup>). In Saudi-Arabië vonden Batterjee et al.<sup>29</sup> verschillen tussen 1977 en 2010 op Ravens Progressive Matrices. De gemiddelde score nam substantieel toe:  $d = 0,78$ , dat is 11.7 IQ punten in één generatie. Per 10 jaar was er een vooruitgang van 3.55 IQ punten.

*Antwoord:* Het effect wordt ook buiten westerse samenlevingen gevonden.

*Vraag: Welke oorzaken worden getoond of genoemd?* De winst mag ons dan toevallen, toevallig is het volgens Flynn niet. Hij noemt als oorzaken: onbekende omgevingsfactoren, ervaring met gestandaardiseerde tests, voeding, SES, verstedelijking, bestrijden en uitroeien van kinderziekten, historische rampen zoals oorlogen, verbetering van scholing, onderwijsprogramma's op de televisie en opleiding in het algemeen. Hij gaat er niet van uit dat het effect berust op genetische en biologische veranderingen. Deze nemen meer tijd in beslag dan één generatie. Sommige auteurs pleiten nochtans voor een genetische basis. Rushton en Jensen<sup>29</sup> verdedigen de genetische basis van de G-factor. Ze vermelden gegevens waaruit blijkt dat het WAIS-IQ-*verschil* tussen zwarten en blanken in de VS stabiel is gebleven tussen 1954 en 2008. De auteurs verdedigen de genetische bepaaldheid van het IQ en leiden

daaruit af dat programma's om verschillen tussen blank en zwart en SES te verminderen vergeefs zijn. De IQ verschillen waren in 1954 bij 17-jarigen 15 punten (1,0 SD), in 1965, 18 punten en 19 punten in 2008. Dit resultaat is vergelijkbaar met economische groei waarbij de verschillen tussen arm en rijk gelijk blijven en zelfs toenemen, terwijl allen er op vooruitgaan. Rushton en Jensen brengen hun resultaten in verband met het Flynn effect, maar hun studie ontkracht het Flynn effect niet. Er kan over generaties vooruitgang zijn bij zowel blanken als zwarten.

Mingroni<sup>30</sup> ging tegen Flynn's context verklaring in. Hij voert heterosis aan, een genetisch effect dat totstandkomt doordat exemplaren van verschillende genetische populaties mengen. Dit kan zich ook nog vanaf 1950 hebben voorgedaan in Europa en al eerder in de VS. De kleine geïsoleerde gemeenschappen verdwenen en veel jonge mensen trokken naar de steden. Veel jonge mensen gaan studeren en mengen met andere studerende. Ze gaan niet terug om de achtergebleven dorpschone (m/v) te ontmoeten. Dit doet zich ook in Nederland voor want sommige provincies ontvolken en vergrijzen. Mobiliteit is de oorzaak van genetische veranderingen en die veroorzaken op hun beurt veranderingen in intelligentie, geschiktheid en prestatie. Woodley<sup>31</sup> kritiseerde dit. Als heterosis een rol zou spelen, dan zou de bijdrage aan de IQ stijging sterk afgezwakt zijn. De winst zou te danken zijn aan ontwikkeling en idiosyncratische levensgeschiedenissen. De simulatie van Mingroni over 50 jaar met credits voor de verhoogde demografische mobiliteit kunnen volgens hem niet meer opleveren dan 3 IQ punten (0,20 SD) per generatie. Roivainen<sup>32</sup> geeft een argument voor een directe causale relatie van GPA en onderwijs naar IQ, dat wil zeggen van de economische en culturele omgeving naar IQ winst. Hij vermeldt onderzoek uit het Oost-Duitsland van voor de *Wende* in 1989. Tussen 1990 en 2006 bleken de scores van de rekruten uit de DDR met 0.5 IQ punt per annum gestegen. De winst was met een  $r$  van .89 gecorreleerd met het bruto nationaal product en met .78 met de verbetering van het onderwijs.

*Antwoord:* IQ-winst wordt beschouwd als gevolg van veel factoren, genetisch en omgeving: vroeg naar school, stimulering thuis, culturele veranderingen, individuele levensgeschiedenissen en de interacties tussen deze klassen van factoren.

*Vraag: Zijn er tegenvoorbeelden?* Als je twijfel hebt over een effect zoek je tegenvoorbeelden. Ze zijn er: Shayer et al.<sup>33</sup> gebruikten in 1975/1976 verzamelde gegevens van Piagetiaanse taken en vergeleken die met gegevens uit 2003. Er blijkt een terugval voor jongens ( $d = -1,04$ ) en voor meisjes ( $d = -0,55$ ). Dit zijn grote tot gemiddelde effecten. Ze hadden geen verklaring voor de terugval en vermeldden als mogelijke redenen: computerspelletjes, televisiekijken, afname van buiten spelen en van discussiëren met leeftijdgenoten, nadruk op lezen, schrijven en rekenen: de *back to basics* beweging in het basisonderwijs. In Pune, India is een differentiële aanlegtest (IAM: *Intelligence-Aptitude-Measurement*) ontwikkeld om ouders en leerlingen te helpen bij de keuze voor vervolgonderwijs. Meer dan 2200 leerlingen vanaf 12 jaar hebben deze gecomputeriseerde test afgelegd. De laatste vijf jaar is er een *afname* geconstateerd op de subtests voor rekenen en wiskunde (persoonlijk communicatie Dr. S. Watve, directeur van het Jhana Prabodhini Instituut in Pune dat de testafname verzorgt; 02-09-2013).

De resultaten kunnen niet meteen geïnterpreteerd worden als het echec van het Flynn effect. De steekproeven kunnen gewijzigd zijn. Er is een verschuiving naar hogere opleidingen in Maharashtra en onderwijsprogramma's veranderen. De kwaliteit van leraren en de behuizing van scholen verbeteren als gevolg van de economische groei. Watve suggereerde echter op dezelfde factoren als Shayer et al. Ze kende die studie niet.

Gignac<sup>34</sup> ging na in hoeverre het Flynn effect te bespeuren viel in de geheugensubtest van de WAIS: *Digit span forward* en *Digit span backward*: het nazeggen van een reeks willekeurige getallen in dezelfde en omgekeerde volgorde. Hij kon gegevens traceren over 85 jaar (1923-2008) bij 7.007 en 6.841 deelnemers. De werkgeheugenscores namen door de tijd heen niet toe: het Flynn effect deed zich niet voor. De waarden bleven gelijk: getallen nazeggen in dezelfde volgorde 6.56 plus en min 2.39 en in de omgekeerde volgorde 4.88 plus en min 2.58. Deze geheugenprestaties correleren hoog met Gfluid.

Woodley en Meisenberg<sup>35</sup> vonden een achteruitgang van 1,35 punt per decade in vier Nederlandse studies met de Differentiële Aanleg Test (DAT), de General Aptitude Battery (GATB) en de Amsterdamse Kinder Intelligentie Test (AKIT). De geboortecohorten waren van 1950-1990. Ook dit wijkt af van Flynn's samenvatting van 2009 dat de Gfluid tussen de 6,57 en 7,14 punten (plus en min 0.5 SD) toeneemt per generatie. De auteurs suggereren dat geschiktheidstests minder gevoelig zijn voor vooruitgang dan de Raven IQ test. Het kan liggen aan het soort test, maar het is ook mogelijk dat het effect wegvloeit door de gewijzigde samenstelling van de steekproef: de jongere studies hadden meer immigranten in de steekproef.

Recente mobiliteit, immigratie en veranderingen in de landen zelf, bijvoorbeeld het afnemen van het aantal kinderen hebben invloed op de toe- of afname van het IQ. Het zal steeds lastiger worden om een Flynn effect te vinden door snel opeenvolgende demografische veranderingen. *Antwoord: Er zijn tegenvoorbeelden.* Er is ook achteruitgang en dat doet zich voor bij Piagetiaanse taken. De dynamiek van landen maakt het steeds moeilijker een 'zuiver' Flynn effect te vinden en te duiden. In India is er waarschijnlijk achteruitgang op een Differentiële Aanleg Test doordat kinderen die niet goed Engels spreken door de welvaart op duurdere Engelstalige scholen komen. In andere landen is er achteruitgang door immigratie en vanwege het feit dat mensen een nieuwe taal moeten leren.

*Conclusie:* We kunnen, moeten het effect nuanceren: er is verbetering, verslechtering en gelijk blijven. Het gaat weer eens alle kanten op in de psychologie. Kunnen we het Flynn effect afserveren? Dat ook weer niet, want de taken verschillen en de samenstelling van de bevolking wijzigt.

Het Flynn effect is evenals de Big Five een voorbeeld van de *Werdegang* van een interessant effect: enthousiast binnen gehaald; overal gezocht en gevonden, genuanceerd voor facetten van intelligent handelen en het effect geldt voor niet voor alle cognitieve prestaties. Bij het zoeken naar het effect ging men uit van een betrekkelijk stabiele bevolking over minstens twee generaties. Dat is nu veranderd. Vergelijkbare steekproeven samenstellen om het Flynn effect te vinden in twee generaties zal op weerstand stuiten en moeilijk zijn.

**Herinterpreteren:** Toont een andere, nieuw bedachte *lezing* van een gedragsconstruct of effect dat we wijzer geworden zijn? Wat aanvankelijk als stoornis beschouwd is, kan gelezen worden als een adaptieve reactie.

Depressie is een voorbeeld: depressieve mensen kunnen rationele en aanvaardbare redenen aanvoeren voor hun gestemdheid. Er zijn gebeurtenissen die aanleiding kunnen geven tot depressie, bijvoorbeeld verlies van een geliefde, volkenmoord, burgeroorlogen, vluchtelingenstromen, onderdrukking van vrouwen, kindsoldaten, gebrek aan respect en pesten op de werkvloer, persoonlijk ingrijpend leed en het voortdurend brutaliseren, uitbuiten en onderdrukken van mensen door mensen. Een voorwaarde voor een depressie is dat de omstandigheden het zelf van de cliënt direct raken. Daar horen het wakker liggen van armoede in de derde wereld, de vernietiging van de natuur en de gestolen fiets meestal niet bij. Rumineren, het blijven doormalen en in een kringetje draaien met steeds dezelfde hinderlijke gedachten, wordt doorgaans als niet-adaptief beschouwd. Het kan echter ook een reactie zijn van een persoon die de tijd neemt om een probleem te doordenken. Hij geeft dat voorrang boven zaken die de gemiddelde burger bezighouden. De concentratie op het probleem reduceert de interesse in andere bezigheden en onderwerpen. Ze gaat samen met het vermijden van contacten om niet onderbroken te worden. Dit zijn symptomen van depressie: nadruk leggen op een probleem, rumineren, gebrek aan interesse: lusteloosheid en motorische inertie. De elementen kunnen ook gelezen worden als manieren van *coping*: omgaan met ernstige problemen.

Hechting is een ander voorbeeld. Het is een dominante typologie voor individuele verschillen op jeugdige en volwassen leeftijd. Het is een categorisering van verschillen, zoals eertijds de lichaamssappen bij Hippocrates. Het construct heeft de connotatie dat veilige hechting een voorwaarde voor het goede leven is. Het onderwerp is een imperium in de nederlandse pedagogie en dat hield en houdt belagers op afstand (Van IJendoorn et al.<sup>36</sup>).

*Attachment* was al een bron voor animositeit tussen Anna Freud (Sigmunds kleindochter) en de onderzoeksgroep van Bowlby<sup>37</sup>. Anna Freuds observaties van kinderen die uit de Nazi-kampen terugkeerden werden door Bowlby's groep onopgemerkt gelaten. Anna observeerde en beschreef en ging hermeneutisch te werk. Deze aanpak was niet populair evenals haar freudiaanse interpretaties. Daarnaast weegt mogelijk mee dat *attachment* in de Bowlby groep een imperium moest worden. Dat is gelukt. Het kreeg aanhang onder kinderpsychiaters, psychologen en pedagogen (zie bijvoorbeeld Stroufe et al.<sup>38</sup>). Overzichtsstudies claimen de bijdrage van veilige hechting aan het goede leven: een geordend sociaal leven met stabiele intieme relaties (Cassidy & Shaver<sup>39</sup>).

Het construct *nonattachment* bestaat ook. Het is uitgewerkt door Sahdra et al.<sup>40</sup> en houdt in dat men zich bevrijdt van *mental fixations*. Het *niet* gehecht zijn zou nuttig zijn voor adaptief functioneren. De oorsprong is de boeddhistische onthechting. De auteurs maakten een 30-item vragenlijst om *nonattachment* te meten. Voorbeelden van items met positieve lading zijn: 'Ik ben niet bezitterig over de dingen die ik heb; Ik voel geen behoefte om me aan slechte ervaringen te onttrekken in mijn leven; Ik aanvaard mijn *flows*; en met negatieve lading: 'Ik heb moeite om successen van anderen te waarderen als ze beter presteren dan ik'; 'Als de zaken niet lopen zoals ik wil, raak ik uit mijn gewone doen'. Het niveau van *nonattachment*

voorspelde adaptieve gedragingen, zoals reductie van gevolgen van lijden en psychologische rijpheid.

Psychopathie is een derde voorbeeld. Dit als stoornis gelabeld gedrag is gekenmerkt door zwakke impulscontrole, gebrek aan schuldgevoel en het ontbreken van empathie met hen die lijden. Niettemin kan men daar tegen in brengen dat deze kenmerken als adaptief geïnterpreteerd kunnen worden in specifieke situaties, bijvoorbeeld wanneer men geen vrees mag tonen en dominantie nodig is zoals leiderschap onder moeilijke omstandigheden. Lilienfeld et al.<sup>41</sup> probeerden deze veronderstelling te toetsen bij 24 presidenten van de VS tot aan George Bush. De auteurs ondervroegen historici over de presidenten en raadpleegden naslagwerken over presidentieel leiderschap om een index te verkrijgen over de mate van pathologie. *Fearless dominance* was gecorreleerd met goede prestaties, leiderschap, overtuigingskracht en crisismangement. De relatie bleef significant na correctie voor covariaten, onder meer intellectuele brillen. Daarin verschilden de presidenten aanzienlijk volgens de auteurs. Antisociaal gedrag was niet gecorreleerd met prestaties als president.

Zij bepaalden de rangorde van *Fearless dominance* van 24 presidenten. Het is aardig om p. 10, tabel 4 van hun artikel te bekijken. Enkele resultaten: Roosevelt (Theodore en Franklin) en J.F. Kennedy staan aan de *Fearless dominance* top, terwijl Taft, Adams en Coolidge onderaan staan. De auteurs beweren dat moed en diertheid geassocieerd zijn met psychopathie: mogelijk een verwaarloosde voorspeller van presidentiële prestaties en van generaals, ondernemers en bankiers. Ze stellen een controversieel construct voor: *successful psychopathy*. Ze aarzelen te zeggen dat VS presidenten psychopathische trekken vertonen. Die aarzeling verdwijnt overigens als het presidenten uit *scoundrel states* betreft, bijvoorbeeld Saddam Hussein, kolonel Gadhafi en president Assad.

Minder controversieel is een wat oudere studie van Winter<sup>42</sup>. Hij las de inauguratieredes van de presidenten van George Washington tot en met Bill Clinton (1993). Deze zijn door tekstschrijvers gemaakt maar ik neem aan dat de presidenten er mee instemden, 'zich er in herkennen'. Hij scoorde de teksten met behulp van de *Achievement*, *Intimacy* en *Affiliation* motieven van de *Thematic Apperception Test* (TAT) van Murray. Carter en Clinton kregen de hoogste score op *Achievement Imagery*, Harrison en Pollock de laagste. Kennedy en Bush (de oude) kregen de hoogste scores op *Affiliation* en *Intimacy Imagery* en Garfield, Th. Roosevelt en Taft de laagste. Het *Power* motief is interessant omdat het met *Fearless dominance* vergeleken kan worden. De hoogste scores voor *Power Imagination* waren voor Truman en Kennedy, de laagste voor Grant, Adams en Th. Roosevelt. De resultaten wijzen bij Th. Roosevelt op een complexe persoonlijkheid: hoge *Fearless dominance* gaat samen met lage *Power Imagination*.

*Conclusie:* De diagnosticus krijgt er steeds gedragsconstructen bij om het probleem/de vraag van de cliënt te beschrijven. De verandering doet practici soms goed: ze stemmen met de voeten en gebruiken ze regelmatig bij hun beschrijving. Het is weer eens iets anders. Er is geen radicale breuk met bestaande theorie en bekende constructen. Het succes van PP en PTG is begrijpelijk vanuit de behoefte aan goed nieuws en nieuw elan. Of het probleem van de cliënt er beter mee beschreven is, blijft een vraag, waarschijnlijk niet.

De tijd maakt ons slimmer, zegt het Flynn effect. Het effect is er wel, maar niet overal en altijd. En er zijn tegenvoorbeelden. De verklaring voor het effect bestaat uit de bekende omgevings- en genetische factoren. Daar kun je op individueel niveau niet veel mee.

Nieuwe interpretaties van constructen staan onder het regime van de hermeneutiek. Daarmee verschaft men zich een andere kijk, opvatting van op menselijk gedrag, bijvoorbeeld op lijden en depressie. De herinterpretaties laten zien dat je met sommige constructen alle kanten op kunt. Dat geeft de diagnosticus vrijheid. Het is geen uitgemaakte zaak wat een construct betekent voor het omschrijven van het probleem of de vraag van de cliënt. Kan dat wetenschappelijk gezien wel? Is het geen mode, of hebben we genoeg van een bepaalde morbide, negatieve lezing van gedrag? Is er wel genoeg vasthoudendheid bij onderzoekers en diagnostici met hun constructen? Leveren ze kennis niet wat gemakkelijk in voor een nieuwlichter?

### 3. Samenvatten: Meta-studies

*Helpt een samenvatting?* We kunnen denken dat er een afdoende antwoord op een vraag komt door alles over dat onderwerp bijeen te zetten: alles over dat construct, dit effect, die samenhang, dat verschijnsel. Behoud het goede en verwerp het onjuiste, valse, ondeugdelijke. Zo doende komt de waarheid vanzelf bovendien in de vorm van robuuste, herhaald aange-troffen bevindingen. Is waarheidsvinding mogelijk door alle empirische informatie bijeen te zetten en zo vast te stellen welke kennis staat als een huis? Kom je door meer empirisch en *fresh* onderzoek tot nieuwe, betere, meer uitgezuiverde inzichten? Onderzoek dat niets oplevert, gooi je weg en interventies die niets opleveren ook. Die zijn niet *evidence-based*. Geen omkijken meer naar. Inzicht vergroten is toch uitzuiveren, ballast wegdoen? Je kunt dit betwijfelen. Een samenvatting mocht bij Nederlands op de middelbare school al geen nieuws bevatten, maar wellicht zie je door samenvatten de belangrijke en robuuste effecten en verbanden wat beter.

*Helpt meer empirisch onderzoek over hetzelfde?* De logisch positivistische kenhouding heeft geleid tot een proliferatie van empirisch onderzoek. Het wordt in meta- en meta-metastudies bijeengezet. Eysenck wees dit af: *meta-silliness*. Lijkt me onterecht, want je leert zo de gemiddelde waarden en hun spreiding kennen van resultaten van soortgelijke onderzoeken. Daar is niets *silly* aan. En robuuste kennis verkrijg je door dezelfde resultaten te vinden in soortgelijke onderzoeken. Replicatiestudies zijn zinvol. Ze hebben opnieuw aandacht gekregen door fraude van een sociaal psycholoog. Bovendien zijn sociaalpsychologische *priming* studies op de korrel genomen. De ecologische validiteit van sociaalpsychologisch experimenteel onderzoek is naar verhouding het laagst (Mitchell<sup>43</sup>).

Het vinden van ongeveer dezelfde, vergelijkbare onderzoeksresultaten sterkt je in de conclusie over een effect of verband. Maar wat is een vergelijkbaar resultaat? Moeten de correlaties en effecten bij herhaling significant zijn? Of, mag je eisen stellen aan hoogte van de gemiddelden en de omvang van de variantie van de *d*- en *r*- waarde in meta-studies? De gemiddelden zijn feitelijk bescheiden tot medium en de omvang van de variantie liegt er niet om. Bedenk daarbij nog dat niet significante resultaten de tijdschriften doorgaans niet halen. Met een beetje welwillendheid kun je dat *outliers*, uitbijters noemen. In dat geval is het juist dat ze niet gepubliceerd worden. Maar weigeren tijdschriftredacties ook hoge *r*- en *d*-waarden of staan ze in dat geval juist te juichen? Er is bovendien al vanaf de jaren 20 bezwaar gemaakt tegen significantietoetsing. Volgens Cohen<sup>44</sup> vertelt het ons iets, dat we niet eens willen weten.

Een recente variant van deze benadering (de verzamelwoede van Bacons<sup>45</sup> mier) is *data mining*. Meta-studies en *data mining* zullen de komende tijd doorgaan. Het laatste betekent dat je zoveel mogelijk gegevens verzamelt van psychologische, sociologische, medische, demografische aard van een persoon of (doel)groep. Het wordt vooralsnog gebruikt om de consument te leren kennen. Je weet veel van een persoon of groep maar of je veel verklaart, begrijpt van het gedrag van die persoon - steeds vaker als consument - of groep is de vraag. Het is een vorm van *sophisticated* boekhouden van de ziel. Het is de Empiristische Hoop op de Zelfopenbaring van het Zijn. In deze werkwijzen zou geen westerse filosoof heil gezien hebben, van Plato via Descartes, Kant tot Merleau-Ponty en Heidegger. Pragmatisten als James en Rorty *omzeilen* dit vraagstuk. Psychologie-onderzoekers doen dat niet. Er is hoop op een definitieve meta-studie van de BF, uiteraard door Costa en McCrae en op de definitieve factoranalyse van intelligentie (denk

bijvoorbeeld aan Carroll<sup>46</sup>).

*Je hebt toch iets in handen met de resultaten van meta-studies:* Ja zeker, empirisch-analytisch pragmatisch telt het resultaat van meta-studies over een verschijnsel, effect, verband of construct. De effecten en verbanden zijn de feiten en die moeten daarna verklaard en geïnterpreteerd worden. Hoe ziet dat resultaat er in persoonsleer en intelligentie /cognitie onderzoek uit? We treffen bescheiden tot gemiddelde  $r$ - en  $d$ - waarden (Cohens vuistregel) met flinke ranges en SDs aan. Schmidt en Hunter<sup>48</sup> vonden bijvoorbeeld in hun samenvatting van 85 jaar onderzoek naar het voorspellen van werkprestaties  $r$ -waarden tussen .12 en .65. Ze krikten het gemiddelde op tot .50 door te corrigeren voor artefacten en administratieve fouten. Meyer en veel collega's<sup>48</sup> vonden in 125 meta-studies over het voorspellen van medische en psychologische criteria een range van  $r = .03$  tot .80. Het gemiddelde was .31 met een SD .19. Je hebt iets in handen maar het lijkt op de kinderhand die gauw gevuld is. Het stapelen van empirisch onderzoek leidt tot bescheidenheid. Het zuivert wel uit, maar je vindt er geen nieuwe verklaring/ interpretatie door. De regel: bestudeer alles, vat het samen in meta-studies en behoud het goede, werkt wel, maar beperkt. In een utilitaristisch kader gaat het om een zo groot mogelijke samenhang, een dito effect met een zo groot mogelijke winst voor een zo groot mogelijke groep. Daar passen meta-studies en *data mining* bij. Bij de laatste gaat het er vooralsnog om U te leren kennen als consument, niet als uniek psychologisch wezen.

*Conclusie:* Bij één cliënt wordt geen meta-studie verricht. *Data mining* is wel mogelijk - en dat gebeurt ook - bij een consument of een doelgroep. Meta-studies berusten op onderzoek van steekproeven om populatieparameters te schatten. Molenaar<sup>49</sup> heeft een poging gedaan om op individueel niveau gedrag te verklaren en te modelleren. Gedrag ontwikkelt zich volgens een idiosyncratisch individueel proces. Dat zorgt ervoor dat statistische technieken gebaseerd op interindividuele variatie (steekproefonderzoek) niet passen, want deze vereisen homogene populaties. Verder is gedrag volgens hem zelf-organiserend en dat resulteert in een idiosyncratische ontwikkeling van ieder van ons. Hij stelt technieken en procedures voor om bijvoorbeeld persoon-specifieke overerfbaarheid van gedrag te bepalen door intra-individuele variatie met verloop van tijd te meten. Hij legt elk individu aan de meetapparatuur en verkrijgt zo doende massa's data. De technische middelen voor registratie en data analyse zijn beschikbaar, maar dit type onderzoek vind je nauwelijks in de tijdschriften. Heymans<sup>50</sup> heeft voorbeelden uitgewerkt van zulk  $n = 1$  onderzoek. Hij volgde soms gedurende meer dan één jaar gedragingen van een persoon. Hij/zij verstrekt dagelijks/ wekelijks informatie. Eerst probeerde hij aspecten van het zelf van de persoon te identificeren om daarna een model te zoeken dat paste op de empirische *patterned change* van gedragingen, zoals morele cognities, persoonlijke belangen die zich manifesteren via emoties en het zelf-narratief.

De diagnosticus moet het vooralsnog met gegevens uit meta-studies doen en de resultaten zijn gemiddeld genomen bescheiden en variëren van studie tot studie aanzienlijk. Aan hem de taak uit te maken of de populatieparameters passen bij zijn cliënt.

#### 4. Toepassen: ons lot te verbeteren, of gaan we vanzelf vooruit?

Het toepassen van psychologische bevindingen is bedoeld om ons inzicht te vergroten, ons handelen efficiënter te maken en ons lot te verbeteren en in eigen hand te nemen. Je kunt ook geluk hebben want het Flynn effect lijkt ons in de schoot geworpen te worden. Het IQ gaat vanzelf omhoog in de loop van de tijd. Sommigen vinden het willen verbeteren van ons lot naïef: verbeteren begint met de beste bedoelingen, maar het kan ontsporen. Het *toepassen* van het kennisbestand van de psychologie dient tot meer welvaart, vrede en veiligheid. Studenten beginnen met dit idee aan de opleiding. Bij het Flynn effect wordt de verbetering van het IQ ons in de schoot geworpen. We hoeven maar een generatie te wachten en er staat weer een intelligentere groep jonge mensen klaar. De oude docent is al bijna 1,5 SD achteropgeraakt.

*Getuigt het streven je lot te verbeteren van naïviteit?* Er is niets tegen het verzachten van leed en het verbeteren van ons bestaan. Sommigen vinden verbetering van ons lot door de psychologie en diagnostiek naïef. Misschien is het dat ook wel. Er zijn goede en kwaadaardige voorbeelden te vinden: Sint Augustinus<sup>39</sup>, een bisschop in Noord-Afrika schreef in zijn *Bekenntnissen* in de 4<sup>de</sup> eeuw dat mensen verward zijn in experimenten: *nos autem in experimentis involvimur*. Hij was de *social experimentation* van Cook en Campbell ver voor. Hij had daarbij de controle op afgoderij op het oog. Zijn boodschap is het ware geloof te vinden en behouden. Dat betekende afstand doen van de heidense Romeinse cultuur.

De pedagoog Comenius<sup>51</sup> (zie Voigt, 1997) zei in 1639 de beroemde woorden: *Tempus est*, te vertalen met: 'Nu moet het gebeuren, Nu'. Een magazine van de jaren 80 en 90 heette: Ouders van Nu. Nu of Nieuw, altijd goed. Comenius' werk ademt het idee dat mensen volmaakt gemaakt kunnen worden en dat is de limiet van verbeteren. Ze hoeven er maar drie boeken voor lezen. Het dikste boek van God: de Schepping, de zichtbare wereld, het Boek dat de mens is, want die is naar Gods beeld geschapen en het Boek dat God de mens gegeven heeft: De Bijbel. Bij dat lezen moeten we niet aan de bereidheid van de hermeneuticus denken, maar eerder aan een bevel: Lees wat er staat! Niet: '*...de typemachines ratelen gekkenpraat; lees maar, er staat niet wat er staat...*'.

*Eigentijdse verbeteraars van ons lot:* Het denken over het verbeteren van ons lot is overgenomen door pedagogen en sommige filosofen. Levensfilosofie is een vak op de Humanistische Universiteit. In dit vak worden regels en houdingen aanbevolen die het bestaan en gedrag van de mens zouden verbeteren. Filosofen, pedagogen, filosofen en behandelaars overschatten effecten van hun boodschappen, regels en interventies. Ze geloven dat. Ze zijn er niet op uit om het empirisch te verifiëren. Ze tonen een aanstekelijk optimisme, zoals ook Theodore Roosevelt in 1897 op het tijdstip dat de VS een wereldmacht werd of al was. Hij zette zich in voor de *great uplifting of mankind*. Empirisch onderzoek van economen, onderwijskundigen en psychologen tempert die hoge verwachtingen, Bij de onderzoekers is dat minder het geval. Het empirisch-analytisch logisch positivisme kent zo zijn heilzame uitwerking. Ze

laten bijvoorbeeld zien dat (overheids)programma's om werkgelegenheid en schoolsucces te bevorderen zelden geëvalueerd worden. En, als het toch een keer gebeurt, zijn de resultaten niet om over naar huis te schrijven, denk aan de vele voorschoolse programma's (*nos autem in experimentis volvimur*) in Amsterdam (Van Calcar schoolbegeleiding en Kohnstamm: Proefcrèche) en Rotterdam (Slavenburg). Er werd verbetering van het niveau van schoolprestaties en intelligentie verwacht op de korte en middellange termijn. Intelligentie en cognitie zijn toch maakbaar, of niet soms?

*Ontsporing bij lotsverbeteraars:* Er zijn voorbeelden van goedbedoelde pogingen om ons lot te verbeteren die na een veelbelovend begin ontaarden. De communistische leider vóór Lenin, Trotski werd vermoord in opdracht van Lenin. In zijn autobiografie is hij een aartsverbeteraar. Zijn biografie laat een bruuft zien die niets en niemand uit de weg ging om zijn doel te verwezenlijken. Hij hield in 1932 een toespraak<sup>52</sup> voor Deense studenten waarin hij mensen lichamelijk en psychologisch incompleet: *half made, half developed* noemde. Training, discipline zou deze deplorabele toestand moeten veranderen. Onder Lenin en Stalin werd de Sovjetburger gemaakt. Mao Zedong leerde van Stalin hoe opvoedingskampen in te richten met als hoogte- of dieptepunt de culturele revolutie. Ook daar ontbrak uiteraard evaluatie onderzoek. Je zou het eens moeten wagen.

*Conclusie:* Vernieuwing als verbetering verschaft de psycholoog, diagnosticus, pedagoog en filosoof werk. Dat is niet te versmaden en niet te ridiculiseren. Wel met beide benen op de grond blijven staan omdat effecten van interventies overschat worden en iatrogeen kunnen zijn. Een voorbeeld van het laatste is het HALT programma voor jonge drinkers. Het cognitief *angehauchte* programma had geen effect. Het bleek dat jongeren daarna meer gingen drinken. De ons door de tijd in de schoot geworpen verbetering van intelligentie (het Flynn effect) is niet alomtegenwoordig; bij sommige taken is er ook achteruitgang, bijvoorbeeld Piagetiaanse (zie sectie 3). En, verbetering van ons lot kan vormgegeven worden vanuit een ontspoorde, cynische gedachtegang, zodat de meesten er slechter van worden. De diagnosticus verwacht geen eenduidig beeld en is bedacht op iatrogene effecten.

## 5. Toepassen: Diagnose Behandel Combinaties

De Diagnose (D) Behandel (B) Combinatie (C) is een werkwijze die belooft de behandeling efficiënter te maken dan tot nu toe door een afstemming te bewerken tussen elementen D en B. Er wordt bovendien gelet op de afzonderlijke kwaliteit van D en B. Wordt de belofte empirisch waargemaakt? Als diagnose wordt opgevat als opstap naar een behandeling dan ligt het uitzoeken van effectieve DBCs voor de hand.

Als er twee dingen zijn die door onderscheiden groepen worden uitgevoerd, is er naar westers gebruik onenigheid over wie het voor het zeggen heeft. Wat is D (de diagnose) waard? Draagt ze bij tot het kiezen van de beste behandeling (B) voor de individuele cliënt? Is er voldoende voorraad aan Bn en deugt wat op *stock* is? Een keten is immers zo sterk als de zwakste schakel. Zijn er al voldoende specifieke en effectieve DBCs?

*DBC's liggen voor de hand:* Menig universiteitsbestuurder maant aan te sluiten bij de actualiteit, de praktijk, het bedrijfsleven en kennis en kunde te valoriseren. De ivoren toren uit en de media in met goede en verantwoorde berichten, informatie, wetenschappelijke vondsten. Openbaar bestuurders vragen om maatwerk, bijvoorbeeld in de Zorg. Persoonsgebonden Budgetten (PGBn) worden gewijzigd en vertraagd uitbetaald vanwege angst voor fraude en omdat uitkeringsinstanties de administratie niet aankunnen. Ze waren bedoeld om individuele patiënten en cliënten beter te bedienen. Niet meer algemene regels uit Den Haag, want die leggen patiënten op een Procrustesbed. Je moet afstemmen. De zorgtoekenning, de organisatie en het uitbetalen moeten niet aan cliënten en mantelzorgers overgelaten worden vanwege belangenverstrengeling. Gemeenten kunnen zorgen voor afstemming en betere controle, want ze staan dicht bij de burger.

Het is begrijpelijk dat er in de medische en nu ook in de psychologische hulpverlening naar DBCs uitgekeken wordt. Het is een open deur dat er niet over- en onder gediagnosticeerd en over- en onderbehandeld moet worden. Alles met mate. De twee leden van de afkorting: Diagnose (D) en Behandeling (B) moeten daarom voor afzonderlijk deugen. De diagnose moet valide zijn en de behandeling moet effect sorteren. Daarna kan nagegaan worden of de combinatie tot synergie leidt. Vormen ze een koppel met een goede relatie? De behandelaar vraagt van diagnosticus zijn meerwaarde te bewijzen. De diagnosticus stelt voor niet meteen te beginnen met een behandeling, maar eerst te kijken wat er aan de hand is. Geen relatie zonder wederzijdse verwachtingen en eisen. Er is een tweedeling van diagnostici en behandelaars en die nemen elkaar graag de maat.

*We houden van tweedelingen:* Ik-niet-ik, *res extensa-res cogitans*, *ying-yang*, *mind-body*, goed-kwaad, god-duivel, man-vrouw, managers-de werkvloer, verleden-heden, theorie-praktijk, romantiek-rationalisme, ideaal-werkelijkheid, empirisme-rationalisme, context van ontdekken-context van rechtvaardigen, orde-chaos, klinische-statistische predictie, *nature-nurture*, links-rechts, *multitrait-multimethod* en *ga zo maar door*. En deze verdeelactie (soms een verdeel-en-heers actie) vertonen we gemakkelijk bij Diagnose en Behandeling. Dit wordt

in de kaart gespeeld door het feit dat er groepen te onderscheiden zijn die zich vooral met het een of met het ander bezighouden.

In werkelijkheid is elk onderscheid niet zo duidelijk, zelfs levend en dood materiaal vormen geen absolute tweedeling, want zuurstof en fotosynthese zijn nodig voor leven, maar ook dood materiaal bevat zuurstof, die nodig is voor leven en zuurstof kan vernietigend zijn. Voor de duidelijkheid houden we ze nog even uit elkaar. De schuld van Descartes, zijn *idées claires et distinctes* werken toch het gemakkelijkst. En, er is meestal ook wel een reden voor een onderscheid. Twee dingen zijn zelden identiek of onlosmakelijk gekoppeld.

*De relatie tussen Ds en Bn is niet één op één:* Misschien enkele in de medische zorg, maar zelden in de psychologische hulpverlening. Zoals Fodor<sup>53</sup> zei: het (elk) doel - verbetering in dit geval is multipel realiseerbaar. Bestuurders zien de DBC's wel zitten. Het is helder, je helpt cliënten en het is goedkoper. Minder foute diagnoses, minder ineffectieve behandelingen. Dat is het verhaal op papier. In methodologisch jargon: Haagse maatregelen en voorschriften vormen experimentele programma's die bestaan uit complexe ketens van acties: de onafhankelijke variabelen. Ze worden ongedifferentieerd op iedereen toegepast, die er - met of zonder diagnose - voor in aanmerking komt. Doel is maximalisatie van kwaliteit van zorg voor een zo groot mogelijk aantal cliënten: het utilitaristische principe. Gemeenten moeten maatwerk bieden. In het jargon: ze moeten rekening houden met moderatorvariabelen en interacties. De variabelen/covariaten zijn kenmerken van cliënten en hun sociale contexten. Kwade tongen beweren dat afschaffen van het PGB voortkomt uit wantrouwen ten opzichte van de cliënt en ten opzichte van commerciële bureaus, die de PGBs binnenhalen voor de zorgbehoeftigen. Anderen noemen het een bezuinigingsmaatregel. Er zijn, als altijd, veel interpretaties en redenen voor het doen van het goede en het streven naar het betere.

*Hoe zit het met de Ds?* De D staat voor diagnose en het diagnostisch proces. Deugen ze, zijn ze valide? Bestuurders steunen op beoordeling van tests volgens richtlijnen van de Cotan. Studietoetsen worden door het Cito gemaakt volgens de moderne testtheorie (IRT). Daar valt weinig op aan te merken. Kwaliteit is geborgd. Voor het diagnostisch proces is het hypothesetoesend model (HTM) aanvaard als valide procedure. Witteman et al.<sup>54</sup> zeggen het zo: het is '*...een leidraad voor het uitvoeren van psychodiagnostisch onderzoek in de praktijk*' en het '*...geeft inzicht in hoe verschillende stappen in het psychodiagnostisch onderzoek op een goede manier uitgevoerd kunnen worden*' (onderlijning: jtl). Alles geregeld met de diagnose. Behandelaars, zoals psychiaters, pedagogen en klinisch psychologen zouden geen probleem van diagnostiek sec hoeven te maken. Soms zeggen ze dat er geen test is voor een gedrags- of leerprobleem waar ze tegen aan lopen. Als de vraag groot genoeg is, komt die test er: marktwerking in testzaken. Er is vrede, of misschien beter gezegd onverschilligheid ten aanzien van aanbod en kwaliteit van het D-lid van de DBCs. Behandelaars maken zich niet zo druk om de D. Ze is bijna niet nodig, maar als je het zo graag doet, dan moet het maar. Je moet immers tests, studietoetsen en vragenlijsten voor persoonlijkheidsstoornissen gebruiken. Elk jaar is er even protest tegen de Cito-toets voor groep 8, maar dat gaat meestal snel over. Bestuurders houden

er aan vast. Er is niets beters, en het Cotan-oordeel is een expertoordeel. Het Cito zit een groot gebouw gevuld met vaklieden en heeft nauwelijks concurrentie.

De D component van de DBCs is hierboven uitvoerig besproken. De conclusie luidde steeds dat - gegeven de d- en r-waarden - bescheidenheid op zijn plaats is. Meta-studies ondersteunen deze bewering en het zit er niet in dat de resultaten op korte termijn verbeteren. De waarden in steekproeven zijn bescheiden tot gemiddeld en er is aanzienlijke variantie. De DBC is er voor de individuele cliënt, maar hij is geen steekproef.

*Onderzoek naar D gaat door:* Onderzoekers hebben uit de aard van hun missie redenen te over om stil te staan bij de kwaliteit van de elementen van diagnostiek: theorievorming, operationalisatie & meten, het instrumentarium en bij het diagnostisch proces. In de praktijk wordt de houding ten opzichte van de diagnose vooral bepaald door de kwaliteit van de instrumenten en nauwelijks door de wetenschappelijk staat van de twee overige elementen. Diagnoses leiden niet tot sombere gezichten bij DBC uitvoerders. Sommigen vinden ze niet zo nodig. Hulpverlening is het centrale doel: kom maar op met je diagnostische kennis.

*Is D onder bepaalde condities niet B?* Behandelen domineert zozeer dat een enkele auteur de diagnose al als *treatment* beschouwt: therapieën, programma's en interventie als *cookie monsters*. Het onderscheid tussen D en B is klein zeggen Poston en Hanson<sup>55</sup>. Ze zien de diagnose als een effectieve ingrediënt in de hulpverlening. Ze moet in dat geval gepersonaliseerd en in samenspraak met de cliënt uitgevoerd worden, denk aan Hermans' subject als mede-onderzoeker in de Zelf Konfrontatie Methode. Poston en Hanson vonden 17 studies waarin een dergelijke procedure gevolgd werd. Ze kwamen tot de slotsom dat de werkwijze effecten had die vergelijkbaar waren met klassiek uitgevoerde therapieën. Dit resultaat ondersteunt therapeuten die de tijdrovende diagnose met behulp van tests en het diagnostisch oordeelsproces willen overslaan. Dat mag en kun je doen want de relatie tussen diagnose en therapie is, als gezegd, niet zo hecht. Bij verschillende diagnoses wordt eenzelfde therapie voorgesteld en bij eenzelfde diagnose worden meer therapieën aanbevolen.

Er is ook kritiek. Lilienfeld et al.<sup>56</sup> beweren dat Poston en Hanson effecten van 'diagnose-plus' overschatten. Ze zouden studies vermelden die diagnoses met behandelingscomponenten combineren. Dat zijn geen pure diagnoses. Bovendien sluiten ze studies uit die geen significante resultaten tonen. Er zijn preciezer opgezette studies nodig om te laten zien dat een diagnose al een effectieve ingrediënt van de therapie kan zijn. Hermans en Poston & Hanson zien een glijdende schaal van diagnose naar therapie, terwijl Lilienfeld et al. het voorlopig als twee categorieën beschouwen.

*Hoe zit het met de Bn?* Bij Ds merkten we op dat het doel: weinig valse positieven en negatieven in bescheiden mate behaald wordt. Nu de Bn: B staat voor behandeling, interventie, therapie, ingreep, modificatie: *treatment*. Paul Meehl was behalve methodoloog en filosoof ook therapeut. Hij heeft bijna 10.000 uur cliënten behandeld. Toch heeft hij niet gepubliceerd over *techniques of psychological treatments* en hun effecten. Op de vraag 'waarom' antwoordde

hij: *'I didn't know, how'*. De huidige *evidence-based* brigade zou een lastige aan hem gehad hebben. Het is immers niet eenvoudig om de waarde en het effect van iets samengestelds als een behandeling te bepalen voor groepen en individuele cliënten. Niettemin willen we dat een behandeling effectief is.

Het experiment is de procedure om te bewijzen dat de onafhankelijke variabele - de behandeling - de oorzaak van de verandering in de afhankelijke variabele is, dat wil zeggen in het (on) gewenste gedrag van een individu of groep. Experimentele programma's dienen om ongewenst gedrag door therapie om te buigen, cognitie te stimuleren en beperkingen op te heffen of te verminderen. Het experimenteel ontwerp kan overtuigen dat de behandeling de oorzaak van de gewenste gedragsverandering is. Dit ontwerp is geliefd omdat het ons op het spoor brengt van oorzaken: *causae efficientes*. Je kunt die mechanistisch opvatten en dat snap je meteen: ik geef een klap en iets gaat stuk; ik doe iets concreets, iets doseerbaars en er verandert iets. Kern is dat het oorzaken zijn die we kunnen manipuleren. We kunnen controle uitoefenen op en over gedrag. Quasi-experimenten zijn gedegenereerde, 'alsof' experimenten waarmee we met enige omzichtigheid ook nog oorzaken op kunnen sporen.

Naast methodologische ernst spreekt uit het (quasi) experimentele ontwerp (Shadish, Cook & Campbell<sup>57</sup>) *optimisme* om de samenleving te verbeteren. De eerste publicaties over quasi-experimenten zijn uit de jaren 60. Deze auteurs nemen aan dat het onderzoek bijdraagt aan het verbeteren van het welzijn van mensen en hun omstandigheden. Ik noemde eerder twee tegenvoorbeelden: Als effect van begeleiding van criminele jongeren in inrichtingen wordt bijvoorbeeld beweerd dat er winst is geboekt, omdat bij gelijkblijvende criminaliteit, het geweld is afgenomen. Heeft begeleiding tijdens het verblijf ertoe geleid dat de criminelen hetzelfde bereiken, maar met minder moeite, met minder geweld? Het bureau HALT dat overtredingen met drank op bij jongeren aanpakt met een *treatment* behaalde geen succes. Jongeren gingen na de verplichte training meer drinken.

Het is zaak bij behandeling op te letten wat er precies gebeurt en wat er uit komt. Het kan tegenvallen: verwachtingen worden niet ingelost en het effect kan zich tegen je keren. Niettemin tonen behandelaars een aanstekelijk optimisme. Als gezegd worden programma's van de overheid om werkgelegenheid en schoolsucces te bevorderen zelden geëvalueerd. Misschien heeft de overheid de neiging met door haar goedgekeurde programma's door roeien en ruiten te gaan: een sterke minister, staatssecretaris. We wezen boven op *iatrogene* interventies. Recidives liegen er niet om en dat is het geval in bijna alle landen ter wereld, van Nederland tot China en India. Had Skinner dan toch gelijk toen hij voor afschaffing van straffen pleitte?

Sommigen vinden dit optimisme naïef. Er zijn goedmoedige en kwaadaardige voorbeelden van experimenteren. Beide gaan uit van maakbaarheid. Pedagogen en filosofen schrijven regels voor die het gedrag van de mens zouden verbeteren. Ze geloven dat het werkt of ze beschouwen het als *self-evident*: het is in zichzelf, met zichzelf en door zichzelf evident en goed. Dat is niet wat met *evidence-based* bedoeld wordt. Een voorbeeld is het Curriculum Schoolrijpheid van de pedagogen Dumont en Kok<sup>58</sup> uit de jaren 60-70. Toen er empirisch onderzoek werd verricht naar de opeenvolging van de cognitieve stappen en die voor een deel niet empirisch gevonden werd - de stappen in enkele domeinen vormden geen Guttmanschaal

- werd gezegd: 'Dat kan niet'. De auteurs waren niet bereid om hun logica te laten sneuvelen voor een empirische toets.

*En de Bn?* Een *treatment* bedoelt het goede te bewerkstelligen. Dat is boven iedere twijfel verheven. Empirisch onderzoek naar effecten van behandelingen matigt de verwachtingen bijna altijd. Ze zijn feitelijk bescheiden tot gemiddeld. Nu moet je niet op iedere slak zout leggen maar toch enkele feiten: Men juicht na succesvol experimenteel onderzoek. Het wordt gepubliceerd, er zijn significante effecten, redelijke effectgroottes en het publiek is enthousiast. Bij nader inzien kan de stemming wijzigen: de effectgroottes van om het even welke therapie zijn gering tot gemiddeld ( $r =$  gemiddeld .30 en  $d =$  gemiddeld  $< 0,50$ ).

Een voorbeeld: In studies van Cuijpers et al.<sup>59</sup> worden effecten van zeven behandelingen voor depressie vergeleken: cognitieve gedragstherapie, non-directieve ondersteuning, gedragsactivering, psychodynamische behandeling, probleemoplos- en interpersoonlijke therapie en sociale vaardigheidstraining. Ze verzamelden 53 studies over milde tot ernstige depressie bij volwassenen. De studies werden geselecteerd en alleen methodologisch degelijke werden opgenomen. Alle behandelingen hadden een zwak effect. Interpersoonlijke therapie was effectiever ( $d = 0,20$ ) dan ondersteunende ( $d = 0,13$ ). Het maakte niet uit of er sprake was van milde of ernstige depressie, want er was geen therapie die voor de ernstigere problematiek minder of meer effectief bleek. Dit resultaat is minder optimistisch dan Amerikaanse studies van bijvoorbeeld Weisz et al.<sup>60</sup>. Ze vermelden  $d$ -waarden tussen de 0,45 en 0,90. Het verschil met enkele Amerikaanse resultaten schrijven Cuijpers en collega's toe aan de publicatie *bias* van redacteurs van tijdschriften.

Er kan bovendien op gewezen worden dat succes gelijkgesteld wordt met de kans op de alfa fout:  $p < .01$ ,  $< .05$ , dus met significantie. Daar is het nodige al over gezegd. Alfa fouten zeggen niet veel. Daar had de ontwerper van de Anova, Fisher<sup>61</sup> al oog voor. Het is het zinvol bij productcontrole. Je kunt je niet veroorloven dat een auto een kans van 5% heeft op een ongeluk door mechanische gebreken. Er zijn redenen te bedenken voor het verschil in therapiesucces tussen Europa en de VS: therapietrouw, geloof in therapie, optimisme en striktheid van uitvoering.

Nog een voorbeeld: De Zwitserse onderzoekers Barth et al.<sup>62</sup> vergeleken zeven depressietherapieën: interpersoonlijke, cognitieve gedrags-, probleemoplos-, psychodynamische-, sociale vaardigheid therapieën, ondersteunende counseling en gedragsactivering. Ze deden een netwerk meta-analyse op basis van 198 studies bij meer dan 15.000 cliënten. De basisvergelijking: therapie-geen therapie toonde dat therapie helpt. Het effect was gering/bescheiden tot een enkele keer gemiddeld. De therapieën verschilden onderling niet: ze waren even effectief voor verschillende leeftijdsgroepen en hetzelfde gold voor verschillende uitlokkers van depressie, zoals ziekte, geboorte van een kind, verlies van een geliefde. De auteurs wijzen erop dat ze gemiddelden vermelden. De effecten hoeven niet voor iedere cliënt te gelden. Tot hier ging het over therapiesucces. Het succes van onderwijsexperimenten verschilt daar niet van. Verderop vermeld ik Hattie's omvangrijke meta-studie over effectieve ingrediënten voor onderwijsresultaten.

*Hoe houden de D's en de B's zich onder het Streng Methodologisch Geweten?* De diagnose komt er bij de op-behandeling-gerichten onder de DBC-beoefenaars matig af. Dat ervoer Kouwer al in de kritiek van de psychiater Kuiper<sup>63</sup> op zijn *Spel van de Persoonlijkheid*. Het maakt ze overigens eigenlijk niet zoveel uit. Een beperkt aantal verdedigt D, bijvoorbeeld Barendregt<sup>64</sup>. Een ander aantal stelt geen lastige vragen, onder hen docenten die D-onderwijs verzorgen. Niettemin zijn behandelaars van mening dat D zijn meerwaarde voor B moet bewijzen. Je doet al dat werk niet voor niets. Maar behandeling komt er in onderzoek ook niet zonder kleerscheuren af: de effecten zijn bescheiden en worden overschat. Dat wil niet zeggen dat gebruikers er geen heil in zien. Denk bijvoorbeeld aan door scholen gewaardeerde pestprogramma's die afgewezen worden, omdat onderzoek ontbreekt of geen resultaat laat zien. Zou hier sprake zijn van het placebo effect? En, zou dat niet bij onderzoek betrokken moeten worden? Hoe vreemd het ook voor ons causaal efficiënt verstand ook mag zijn: placebo's blijken regelmatig te werken, ook in de geneeskunde. Sommige huisartsen schrijven ze voor en met *meer* in hun hoofd dan de slogan: 'baat het niet, dan schaadt het niet'. Deze stand van zaken zou voor onderzoekers begiftigd met het Streng Methodologisch Geweten een uitdaging kunnen zijn. Ze zouden kunnen proberen uit te vinden, waarom, onder welke condities (covariaten) d- en r- waarden totstand komen en of de limiet van die waarden bereikt is.

*Alleen evidence-based B's toelaten:* Deze stand van zaken lokt acties uit om bij behandelingen het kaf van het koren te scheiden. Als je nu alleen *Evidence-Based Treatment Programs* (EBTP's) toestaat, zou dan de kou niet uit de lucht zijn? Dit heeft geleid tot een categorisering van therapieën. Ze verschillen met betrekking tot het empirisch bewijs voor hun effectiviteit. Chambles en Ollendick<sup>65</sup> onderscheiden: I 'Well-established/ Efficacious and specific II Probably efficacious en III Promising. Niet onderzochte zou IV kunnen zijn en Mislukte V. Placebo's hebben geen plaats in hun ordening. Misschien zou je de controle conditie in experimenteel onderzoek als zodanig kunnen opvatten. Ze verdelen de behandelingen over problemen, zoals gezondheid, eten, pijn als gevolg van kanker, hoofdpijn, migraine. De meeste behandelingen komen in categorie II terecht: mogelijk effectief; verbetering nodig, maar er is perspectief. De kinderombudsman stond waarschijnlijk de categorie *Promising* (III) toe voor programma's ter bestrijding van pesten. Het is mogelijk dat hij alleen programma's uitsloot, waar (nog) geen onderzoek naar gedaan is en die vooralsnog *face valid* (mijn IV) zijn. De V categorie zal niet gepubliceerd worden. Dat is jammer, je leert er immers van. De *evidence-based* eis heeft ertoe geleid dat medicijnen zoals ritalin (methylfenidaat) en antidepressiva gevestigd zijn in de zin van hun sporen verdiend hebbend. Ze zijn werkzaam. Cognitieve gedragstherapie en multi-systeem therapie komen af en toe voor in het rijtje van *evidence-based* producten. Meta-studies laten zien dat behandelingen resultaten opleveren in experimenteel onderzoek onder laboratoriumcondities. Succes is schaarser voor interventies in het veld. Zijn de I en II categorie interventies in het laboratorium onder ideale omstandigheden of in het veld uitgevoerd? Behandelingen moeten in concrete situaties, zoals gezin, schoolklas, de werkvloer geïmplementeerd worden. De effecten moeten generaliseren naar andere vergelijkbare gedragingen en situaties.

Er is het een en ander te doen over de *evidence-based* programma's. Als je het label I of II bemachtigd hebt, wil dat nog niet zeggen dat het overdraagbaar is naar nieuwe steekproeven, situaties en behandelaars. McHugh en Barlow<sup>66</sup> hebben bijvoorbeeld een overzicht gemaakt van EBPTs, hun implementatie en verspreiding in de VS. Ze stuiten op barrières voor de invoering van de EBPTs. Er is daarom training, supervisie en advies aangeboden aan gebruikers. Beschikbaarheid van EBPTs alleen is kennelijk niet genoeg. Kazdin<sup>67</sup> concludeert dat er al een lijst is met daarop 'goede' EBPTs. Toch krijgen sommige cliënten niet de behandeling die ze nodig hebben, omdat therapeuten hun favoriete behandeling geven. Hij beveelt mini-programma's aan voor behandelaars om zo de impact van de EBPTs te vergroten.

Barlow et al.<sup>68</sup> hebben al een *update* van EBPTs gemaakt. Ze zijn optimistisch over de effectiviteit van behandelingen, vooral farmacologische aangevuld met psychologische. Ze ontlenen hun optimistische kijk voor een deel aan resultaten van dieronderzoek. Twijfel blijft: was er een belang van de farmaceutische industrie? En, als het succes zo gemakkelijk te behalen is, waarom is het dan niet veel eerder gebeurd? Waarom neemt het aantal behandelingen niet af en blijven een paar goede over? Het tegendeel is het geval. Er komen er steeds meer bij.

De Psycholoog van november 2015 (advertorial) laat een voorbeeld zien. Naar aanleiding van een *survey* over wie de Nobelprijs voor de Psychologie verdient werd een persoon genoemd die als uitbijter werd opgevat. Ze werd genomineerd maar de jury vermoedde een lobby. Veel cliënten hadden juichende verhalen over een kandidaat met een nieuwe therapie *Past Reality Integration* geschreven. Getalsmatig was ze (Ingeborg Bosch) bij de beste tien. Weer een nieuwe therapie en uiteraard met een engelse naam. Eerst wordt vastgesteld dat de cognitieve gedragstherapie (de gouden standaard), medicatie en psychoanalyse niet leverden wat ze beloofden. Effectieve ingrediënten zijn specifiek (zie hieronder). Jonge behandelaars zouden het goed doen omdat bij hen het behandel-nihilisme nog niet heeft toegeslagen. De nieuwe therapie omvat gemiddeld 30 sessies en belooft 'cognitieve herstructurering, heropslag, gedragsverandering en toegenomene welbevinden bewerkstelligd via zelf-observatie, exposure en gedrags-(de)-conditionering, zonder medicijnen'. Blijvend resultaat voor een brede populatie (geen DBC dus). Nadruk ligt op de therapeut zelf, want hij kan nooit verdergaan dan waar hijzelf staat. Hij moet zijn blokkades opheffen, een open hart hebben, compassie met zichzelf en met de cliënt. Ik wacht een onafhankelijk onderzoek af en ben benieuwd of hoe dat methodologisch in het vat gegoten wordt. Er moet zowel bij de cliënt als de therapeut gemeten worden en moet longitudinaal zijn, gezien de claim van een blijvend resultaat. De prijs voor de 30 sessies is niet vermeld. Mijn voorspelling: het gaat als met alle andere interventies een d-waarde tussen de 0, 25 en 0.40. Deze predictie berust op de *past reality integration* van onderzoek. Als de therapie in het Duits gelabeld was, zou ze *Vergangenheitsbewältigung* heten. Een wel heel oud fenomeen na WO II en daar hebben Bosch c.s. niet aan gedacht.

*Naar werkzame ingrediënten van Bn zoeken:* Als er zoveel zijn en ze alle een beetje effectief (bescheiden, significante d-waarden) zijn, kunnen we dan niet de werkzame ingrediënten opsporen van deze interventies? Hoe werkt deze voedselmetafoor uit in onderzoek? Cuijpers en collega's suggereren dat succesfactoren specifiek zijn: ingrediënten maken alle soorten maaltijden (*treatments*) lekker, zoals vroeger bij de afhaalchinese. Ze passen niet precies bij een therapie.

Voorbeelden: Shirk en Karver<sup>69</sup> meldde op basis van 23 studies dat bij adolescenten de

kwaliteit van de relatie cliënt-therapeut het best de uitkomst voorspelde. Het gaat om een open, luisterende, niet-veroordelende houding en om uitleg geven in de taal van de cliënt. De relatie tussen de ingrediënten en het resultaat werd gemodereerd door een inhoudelijke factor: het probleem van de cliënt en daarnaast door factoren als timing en bron (wanneer gaf wie informatie), inschatting van de kwaliteit van de relatie en soort en bron voor het resultaat. Cuijpers et al. en Barth et al. wezen op enkele kenmerken van de interactie, zoals de cliënt aanmoedigen om buiten de therapie sessies om iets te ondernemen, het bieden van een begrijpelijk kader voor hun depressie, een verklaring geven voor wat de cliënt dwarszit en dat zo vertellen dat de cliënt het *frame* aanvaardt, zodat vanuit een gedeeld *frame* inzicht verschaft wordt en een behandelingsvoorstel gedaan kan worden. Met andere woorden, deze auteurs zoeken het verschil niet in de soorten therapie maar in werkzame onderdelen van om het even welke therapie. Welke de effectieve elementen in e-Mental Health zijn, moet nog uitgezocht worden, maar lijkt vooralsnog om algemene principes te gaan: bekwaam en accepterend communiceren. Ze lijken op de *unconditional positive regard* van de therapiegoeroe van de jaren 50-70: Carl Rogers. Niets nieuws onder de zon?

De voorbeelden laten zien dat effectieve ingrediënten conceptueel en hermeneutisch gecreëerd worden. Je moet goed kijken wat gemeenschappelijk is aan die interventies, wat hun onderliggende zin en betekenis is. Er is net zo min als bij de DSM-5 multivariaat onderzoek verricht dat de programma's en hun kenmerken reduceert tot enkele werkzame en onderscheidbare dimensies. Dat zou qua onderzoek een tour de force zijn en dat is niet te verwachten bij de gewoonte van het presenteren van *least publishable units*.

In theorie en in de ideale wereld zijn er volop EBTPs. Ze geschikt maken is vers twee, ze volgens het boekje uitvoeren is vers drie, de goede uitvoerders vinden is vers vier, de goede klinische groep bedienen is vers vijf en er zijn misschien nog meer verzen. Effectieve ingrediënten blijken niet specifiek voor een psychotherapie of begeleiding. Het zijn algemene kenmerken die te maken hebben met een open stijl, mensen ertoe brengen iets te ondernemen en bekwaam en adequaat communiceren. Hoewel gemakkelijk gezegd en op het oog eenvoudig toe te passen, zijn de feitelijke uitvoering van- en *timing* bij behandelingen niet simpel. Ze vergen ervaring en alertheid van de hulpverlener. Het is vergelijkbaar met de *performance* en *timing* van de cabaretier. Hij moet zijn publiek kennen en aanvoelen om het bij de les te houden en te amuseren: behandeling, interventie is kleinkunst. Onze cabaretier heeft heel wat *try outs* nodig voor hij met zijn voorstelling de boer op kan: de behandelaar als *stand-up comedian* van de hulpverlening. Dit is niet denigrerend bedoeld. Het drukt waardering voor vakmanschap uit.

Efficiëntie van onderwijsprogramma's en curricula is ook nagegaan. De vernieuwingslust van onderwijskundigen is niet te stuiten. Misschien lijden ze onder te snel voortschrijdende inzichten. Geldt ook hier dat de kwaliteit van de leraar, zijn enthousiasme en communicatievaardigheden de belangrijkste ingrediënten zijn? De Australische/Nieuw Zeelandse onderwijsonderzoeker Hattie<sup>70</sup> heeft in zijn boek van 2009: '*Visible learning: a synthesis of over 800 meta-analyses relating to achievement*' meer dan 50.000 studies uit 816 meta-studies samengevat. Eén van zijn vele conclusies: klassengrootte is een *minor factor* voor het prestatieniveau. De roep om kleinere klassen is niet te rechtvaardigen vanuit de gemiddelde prestatie. Om andere redenen kan klassenverkleining overigens aangenaam en nuttig zijn. Hattie noemt als belangrijkste variabelen: verwachtingen van leerlingen, afwezigheid van storende leerlingen,

stimuleren van gepast gedrag in de klas, kwaliteit en motivatie van de leraar, verbeteren van taalvaardigheid en -begrip, de voorafgaande prestaties van leerlingen, relatie tussen leraar en leerling, feedback aan de leerling over zijn prestaties en gedrag. Een samenvatting van zijn aanbevelingen staat op het internet: *tgl.pdf: Visible learning, Tomorrow's Schools, the Mindsets that make the Difference in Education*.

Ondanks de reddingoperaties van interventies in de vorm van tegemoetkomen aan de *evidence-based* eis en de zoektocht naar effectieve ingrediënten, kunnen we er niet meer van maken dan dat bij Behandeling/Treatment het glas zowel *halfvol* als *halfleeg* is. De stemming over het eerste lid, de D van de DBC was al gematigd: het kan vriezen en het kan dooien. Ik stelde vast dat de prestaties op het domein van de D bescheiden zijn. Niettemin blijkt uit vele nieuw ontworpen interventies dat er optimisme heerst (*Past Reality Integration*). Volgens sommigen is dat naïef. Zeggen dat het wanhoopspogingen zijn, gaat te ver maar de twee leden D en B spelen vooralsnog niet de sterren van de hemel.

*De DBC als maatwerk voor de cliënt:* De DBC is niet uitgevonden door de bestuurlijke elite maar het idee ligt zozeer voor de hand dat autoriteiten het gemakkelijk aanvaardden. Er moeten oude voorbeelden zijn. Een klassieke DBC is de opvatting dat psychoanalyse geschikt zou zijn voor hoogopgeleiden. In dit geval discrimineert de DBC. Gedragsmodificatie zou passen bij laagopgeleiden. Dat voelde niet goed. Er is vast wel een proefschrift van een bevlogen wetenschapper, misschien uit een *well to do family* (jaren 70!) die deze veronderstelling in twijfel trekt. Denk aan de cultuurcriticus/psychiater Daniels<sup>71</sup> (pseudoniem: Theodore Dalrymple). Hij besteedt uitvoerig aandacht aan het feit dat de Engelse onderklasse door de aanpak van praten en helpen - pampereen zegt hij - gestimuleerd wordt om niets te ondernemen. Die aanpak lijkt wel wat op de neo-analytische hulpverlening met dat vleugje rationaliteit: ik kan toch uitleggen, dat ze niet de hele dag voor de TV moeten liggen, werk moeten zoeken, de fles moeten laten staan en vooral goed naar zichzelf moet kijken.

Uit dit oude voorbeeld blijkt ook dat de hulpverlening *vooropliep* met combineren van diagnostiek en behandeling. De therapeut, onderwijs- en leerlingbegeleider voegde altijd al zijn behandeling naar kenmerken van de cliënt/leerling en zijn sociale context. Pedagogen maken een behandeling op maat na diagnose van het kind, zijn ouders en de school. Zij noemen het *Handelingsgerichte Diagnostiek*. De onderwijspsychologen benutten kennis over het functioneren van kinderen op domeinen als rekenen, taal, aandacht, motivatie, zelfconcept en emotionele en gedragsproblemen aangevuld met informatie over de opvoedings- en onderwijscontext. Zij noemen het: *Diagnostiek in de Leerlingbegeleiding*: Naar welke acties dat concreet verwijst is altijd moeilijk te omschrijven. In de praktijk van de uitvoering zijn er altijd vrijheidsgraden. De concrete praktijk is het resultaat van de combinatie van programma's x uitvoerders x situaties x leerlingen. Bovendien is er weinig onderzoek dat bewijst dat Diagnostiek + Behandeling tot gedragsverbetering leidt. Dat vergt een experimentele en controle groep of  $n = 1$  onderzoek met zijn eigen methodologie en bewijsvoering. De werkwijzen berustten niet op onderzoek maar op ervaringskennis, op praktijktheorieën zou Van Strien<sup>72</sup> zeggen.

*Het risico van face-validity bij DBCs:* Het maken van DBCs lukt wel. Er is creativiteit en fantasie voor nodig. Meestal is de behandeling samengesteld als een stappenplan en bevat verschillende

ingrediënten. Toch lopen DBCs het gevaar *face valid* te zijn. De therapie of training wordt vormgegeven en ingezet naar aanleiding van een waargenomen, voorondersteld effect en van een gelijkenis, passendheid (*fit*) van de behandeling met het ongewenste gedrag. Dat zijn vooralsnog hypothesen. *Face validity* is als de liefde op het eerste gezicht. Ze duidt op de overtuiging dat deze test/behandeling precies meet/bewerkstelligt wat je nodig hebt om het gedrag op een criterium te voorspellen/een gedrag te veranderen. Ze berust op een waargenomen of hermeneutisch geconstrueerde gelijkenis tussen het werk of een taak en de inhoud van de test/behandeling.

Voorbeelden: Er zijn instrumenten waarbij we denken en aanvoelen dat ze een concreet gedrag, prestatie voorspellen. In de jaren 50 van de vorige eeuw werden chauffeurs geselecteerd. Goede chauffeurs zijn belangrijk omdat ze duur materiaal vervoeren en op tijd bij klanten moeten afleveren. Er werd een chauffeursstoel geplaatst in een kamer met films die verkeerssituaties lieten zien. Individuele verschillen in gedrag op de stoel bleken niet gecorreleerd met het criterium: levertijd en aantal ongelukken. Het rijden in het verkeer was niet te vergelijken met de situatie in de stoel. Dit leerden we al uit Drenth's<sup>73</sup> *Testtheorie*. Vroeger maakten leraren zelf toetsen voor de leerlingen. De Schriftelijk Overhoren (SO-tjes) cijfers correleerden nauwelijks met tentamencijfers op de hogescholen en universiteiten. We herinneren ons de teleurstelling van Wissler<sup>74</sup>. Als leerling van McKeen Cattell deed hij onderzoek naar de relatie tussen test- en schoolprestaties van leerlingen ( $n = 90$  tot  $n = 252$ ) van een naburige *high school* en studenten van de universiteit van Columbia. De resultaten waren mager. Schoolprestaties correleerden  $r = .18$  met reeksen getallen onthouden,  $r = .08$  met de sterkte van de handgreep en  $r = .02$  met kleuren benoemen. De items vertoonden onderling weinig samenhang: de taken vormden geen ééndimensioneel en inhoudelijk goed te interpreteren construct. Hij stapte teleurgesteld over naar culturele antropologie. Nog een voorbeeld: In de jaren 60 en 70 werd een zogenoemde draadbuigtest gebruikt om het technisch inzicht op eenvoudig niveau te voorspellen. Een voorbeeld op papier moest met een soepele ijzerdraad op schaal nagemaakt worden. Het resultaat voorspelde het uitvoeren van eenvoudig technisch werk nauwelijks.

Dit zijn voorbeelden waarbij de constructeurs van diagnostische middelen, tests geloven dat ze het bedoelde criterium voorspellen: *quod non*. Binet en ook De Groot<sup>75</sup> (*Vijven en Zessen*) wezen op de onbetrouwbaarheid van schoolcijfers en oordelen van leraren. Deze voorbeelden zijn naar verhouding eenvoudig. Het gaat om het voorspellen van een criterium. Bij de behandeling gaat het om een voorspelling van complex gedrag na een ingewikkelde interventie. De Diagnose-Behandel-Combinaties zijn op het oog weliswaar aannemelijk, maar kunnen - evenals bij tests die een criterium zouden voorspellen - *face valid* zijn. Er is experimenteel onderzoek met controlegroepen nodig om resultaten van de complexe B op een aantal afhankelijke variabelen aan te tonen. Verschuieren en Koomen<sup>76</sup> en Braet en Bögels<sup>77</sup> voeren de redacties van boeken met zulke DBCs. Het eerste mikt op diagnose in functie van de toewijzing van zorg en op het verbeteren van de afstemming tussen onderwijsomgeving en individuele kenmerken van de leerlingen. Dit is het ATI onderzoek, waar Cronbach en Snow<sup>78</sup> in vastliepen. Het tweede bevat onder meer 42 protocollaire behandelingen voor kinderen en adolescenten met psychische problemen. Voor beide geldt dat bewijs voor effect van de behandelingen geleverd moet worden. Het enthousiasme van deelnemers aan DBC symposia en van uitvoerders van DBCs is niet genoeg.

*Face validity* kan ook de andere kant op werken. Er zijn tests/behandelingen met significante en voldoende predictieve validiteitscoëfficiënten/effecten. De autoriteiten kunnen deze echter opzij schuiven als niet terzakedoende, omdat zij het criteriumgedrag op school of bedrijf niet terugvinden in de tests/behandelingen. Een recent voorbeeld is een studie van Boelema<sup>79</sup>. Zij heeft gedurende enkele jaren het gevolg van wekelijks alcoholgebruik van adolescenten op concentratievermogen, impulscontrole en geheugen onderzocht. Ze vergeleek drinkende jongeren met minder- en niet-drinkende jongeren. Zij vond geen verschillen op de variabelen tussen de zes groepen die ze onderscheidde op basis van alcoholgebruik. Ze heeft haar data drie keer opnieuw geanalyseerd om zeker zijn van de uitkomst. Tegen het protest in van onder meer van artsen die het comazuipen op de TV brachten, bleef ze haar empirische resultaten trouw. Dat doet een onderzoeker en dat hoort zo. In een interview zei ze dat de instrumenten wellicht niet gevoelig genoeg waren. Ze had van mij mogen zeggen: 'Doet u het maar eens na met instrumenten die volgens u gevoeliger zijn; ik houd me aan mijn conclusie op grond van mijn onderzoek'. De algemene *face valid* overtuiging is dat alcohol hersencellen beschadigt bij adolescenten, *quod non* dus in dit onderzoek. Als we dit naar een DBC transponeren betekent het dat een behandeling wel het bedoelde effect heeft maar zij op *face valid* gronden niet gepikt wordt. Waarom zou ik iedere dag activiteiten ontplooiën, zoals een uur wandelen per dag, geen TV kijken na 23.00 uur, sociale contacten onderhouden, iets nieuws ondernemen, alcohol zeer met mate, enzovoort, als ik niet inzie wat ze te maken hebben mijn gedrag thuis en op het werk.

*Het ATI onderzoek van Cronbach en Snow:* Dit onderzoek is dé voorloper van de DBCs. Het gaat over kenmerken van de cliënt x zijn sociale context x interventie/behandeling. Er is theorievorming over individuele verschillen, ontwikkeling en de context. *Treatment* is in dit verband de manipuleerbare context in de vorm van behandelprogramma's. In zijn *Presidential Address* voor de Amerikaanse psychologenvereniging (APA) heeft Cronbach<sup>80</sup> correlatieel en experimenteel onderzoek omschreven als *The two Disciplines of Scientific Psychology*. Weer een tweedeling, maar ze schiep duidelijkheid. Later heeft hij ze in zijn boek met Snow van 1977 gecombineerd in het ATI systeem. Het idee is dat het effect van een behandeling afhangt van het niveau op een aantal individuele verschillen variabelen van personen en hun contexten. Het effect van een onderwijsprogramma hangt bijvoorbeeld af van de intelligentie, motivatie en creativiteit van de leerling en van de context: ouders, school, buurt, SES. Hattie heeft er intussen veel over bijeengezet.

Het ATI systeem kan ook op therapie worden toegepast met als gevolg dat de therapie aangepast wordt aan de individuele cliënt. De verwachting is dat afstemming tot betere resultaten leidt dan wanneer de gemiddeld genomen meest werkzame behandeling wordt ingezet. Cronbach en Snow vonden weinig empirische steun in hun boek vol voorbeelden uit onderwijs en therapie. Hoofdstuk 12 gaat bijvoorbeeld over *Personality x Treatment Interactions*. De behandeling van angst is het thema. In *Handelingsgerichte Diagnostiek* en *Diagnostiek en Leerlingbegeleiding* wordt niet dit standaardwerk over ATI en DBC verwezen. Synergie, interactie, afstemming tussen diagnose en behandeling zijn niet in een handomdraai bewerkstelligd. Het vergt veel  $n = 1$  en groepsonderzoek.

*Is er na Cronbach en Snow iets veranderd?* Lukt na 40 jaar die synergie, afstemming, integratie

wat meer? Ik neem als voorbeeld de drie theoretische oriëntaties van het thema I, sectie 1. Is er inter- en sub-disciplinair onderzoek? En, is het resultaat synergie? Als we de drie oriëntaties van psychologische theorievorming 2 x 2 combineren blijkt dat de Individuele Verschillen x Ontwikkeling Combinatie door de dominantie van de eerste vooral tot stabiliteits- en continuïteitsonderzoek heeft geleid. Typerend gedrag en intellectuele prestaties komen steeds als stabiel uit de bus. Stabiel verwijst naar absolute stabiliteit (gelijke gemiddelden over leeftijd), rangorde (de stabiliteitscoëfficiënten van tests en vragenlijsten), structureel (bijvoorbeeld vergelijkbare factorpatronen door de tijd heen), ipsatief (binnen een persoon dezelfde rangorde: weinig onderzocht) en processtabiliteit (overgangskansen van verschillende gedragingen blijven gelijk door de tijd heen: weinig onderzocht). Continuïteit wordt gereserveerd voor stabiliteit van gedragingen die er wel verschillend uitzien maar dezelfde betekenis hebben, denk aan de hechtingcategorie op jeugdige leeftijd en de kwaliteit van persoonlijke relaties in de volwassenheid. We leren keer op keer uit onze boeken hoe stabiel persoonlijkheid en cognitie zijn. Men is er niet op uit ons te leren hoe personen veranderen, terwijl veranderingscoaches volop werk hebben en succes claimen. Hoe kan dat?

Hoe gaat het met de combinatie Individuele Verschillen x Sociale Context? Integratie, synergie, afstemming? Het persoon-situatie debat is voorbij. De trektheoretici hielden vast aan transsituationele trekken en de situationisten probeerden de variantie in criteria te verbinden met variatie in situaties. Een beetje theorie brengt ze van meet af aan bijeen, zoals Eysenck al deed met situatie-specifieke responsen als onderste laag in zijn hiërarchisch persoonsmodel. De combinatie levert niet meer op dan dat een bescheiden hoeveelheid variantie in een criterium door situatie en trek gebonden kan worden.

De Ontwikkeling x Context Combinatie laat dominantie van de te manipuleren context zien. Piaget zag de sociale context voornamelijk als *aliment* voor de cognitieve ontwikkeling. In onderzoek werd geprobeerd die ontwikkeling te versnellen. Een voorbeeld is het bevorderen van de overgang van pre-operationeel naar concreet operationeel door training van deeltaken bij de conservatieproeven. Het gaat daarbij niet om Piagetiaanse structuren, maar om leer- en informatieverwerkingsprocessen. Die kunnen we immers - op papier - analyseren, isoleren, in een stappenplan onderbrengen en dan trainen. Dit onderwerp is nu van de onderzoeksagenda verdwenen. Het succes was beperkt en het onderzoek van korte duur. Loevinger<sup>81</sup> werd de psychometrie ontrouw en gaf zich aan de persoonsontwikkeling. Ze beweerde dat het effect van een training eerst zichtbaar kan worden als een bepaald ontwikkelingsniveau bereikt is. Een plausibele gedachte die wel enige steun kent volgens Westenberg. Het is geen populair onderzoeksthema en bijgevolg is er niet veel over bekend. Het is en blijft niettemin populair bij behandelaars. Er zijn congressen met de belofte goede DBCs te laten zien. U kunt er in de winter van 2015/2016 naar toe in een mooi Zwitsers oord. Een titel als 'MAATWERK, de weg naar *precision therapy* voor uw patiënten' laat zien hoezeer de DBC - in dit geval met de nadruk op de B - in de belangstelling staat. Ja, de ene patiënt is de andere niet. Wanneer psychologen het initiatief nemen tot een dergelijk congres ligt de nadruk op de D zonder de behandeling uit het oog te verliezen.

*Wat nu?* Je mag vaststellen dat voor het uitzoeken en onderzoeken van het effect van DBCs nog werk te doen is. De praktijk loopt voorop in de toepassing. Bestuurders zien DBCs wel zitten. Het ligt intuïtief voor de hand. Niettemin is maatwerk op basis van ATI onderzoek en

integratie van theoretische oriëntaties niet serieus van de grond gekomen. De dominantie van het experiment is gebleven zoals blijkt uit de nadruk op *evidence-based* behandelingen. Antipeestprogramma's zijn daar een voorbeeld van. De kinderombudsman wil wildgroei in deze programma's beperken. Er zijn er meer dan 100 van in Nederland. Er moeten een paar bewezen effectieve overblijven. Hij zegt daarbij te steunen op wat de wetenschap - zijn adviseurs - hem verteld heeft. Dat is één element van de DBC: de kwaliteit van behandelingen. Om effecten van de Bn te bewijzen is er in de empirisch-analytische traditie maar één weg: het *true experiment*. Sir Ronald Fisher heeft dat uitgelegd met veldjes groente. De binnenvariantie moet in de hand gehouden worden en de tussenvariantie moet op basis van de experimentele variabele het verschil maken. De binnenvariantie wordt groter naarmate de steekproef personen op allerlei variabelen verschilt. Men kan ze proberen weg te filteren met covariaten maar in de werkelijkheid kun je de groepen niet zo kiezen dat ze covariaatvrij zijn. En je doel is meestal niet iets te ontwerpen voor een zeer specifieke homogene groep. Je maakt het programma voor een grote gemêleerde groep. Je sluit geen meisjes of jongens, ouderen of jongeren, lage-hoge SES groepen uit en dat zijn de bekende covariaten. Het experiment is geschikt om utilitaristisch vast te stellen wat het beste programma is voor een zo groot mogelijke groep. We kennen de d-waarden en dat leidt tot bescheidenheid over de effecten van de Bn.

Misschien worden de DBCs te complex opgezet: te veel covariaten, risico en beschermende factoren en door duur en omvang te complexe behandelingen. Er kan onderweg zo veel uit de hand lopen dat je door de bomen het bos niet meer ziet. Lukt het met bescheidener opzetten om iets meer over een DBC te weten te komen? Misschien is een variant op adaptief testen een mogelijkheid om greep en zicht te krijgen op DBC's. Het wordt gedaan met 'Rekentuin' en 'Taalzee' van de UVA. Ze zijn te vinden op: <http://www.oefenweb.nl>. De afhankelijke variabele kan geschaald worden als een individuele verschillen dimensie van een groep of als een ontwikkelingsdimensie van een leerling. De behandeling bestaat uit veel, precies gekozen en uitgeteste opdrachten die passen bij de plaats op de latente trek of de ontwikkelingsschaal van de individuele leerling. Je hebt er een grote voorraad aan precies metende taken en opdrachten voor nodig. Dat is nog eens leerstof! Dit is wat Loevinger voor ogen stond. Als dit slaagt, kun je langzaam naar meer covariaten (diagnose) en samengestelde programma's (behandelingen) gaan.

Misschien is het mogelijk dat zorgvuldig opgezette en gemonitorde casestudies stukje bij beetje zicht geven op DBC's. Deze kunnen gegeneraliseerd worden naar soortgelijke gevallen. Je komt ze in tijdschriften niet tegen. Gigerenzer<sup>82</sup> noemt dit een verlies. Hij vertelde dat hij enkele artikelen in het *Journal of Experimental Psychology* van de jaren 20 tot 30 van de vorige eeuw had gelezen. Hij merkte op dat er iets verloren is gegaan: verschillende (statistische) methoden, precieze rapportage van individuele gevallen, zorgvuldige selectie van proefpersonen, nagaan of de sekse van de proefpersoon of proefleider van invloed was, enzovoort. Dit zijn precies de covariaten die er bij voorbaat uitgehaald worden om succes van een behandeling te bepalen. Ondertussen kun je practici, diagnostici en behandelaars aanmoedigen door te gaan met zoeken naar de succescombinatie bij individuele cliënten. Er valt van hen iets te leren als ze *vastleggen* hoe en waarom ze de DBC in een concreet geval zo vormgeven hebben, en wat het resultaat was op korte en middellange termijn. De werkgever/organisatie zou hen kunnen toestaan om één casus per maand zo volledig mogelijk uit te werken, te bespreken en

te rapporteren. Daar horen ook *niet* succesvolle DBCs bij. Misschien is er zelfs een tijdschrift te maken dat ze wil publiceren. Je moet de moed hebben om je niets aan te trekken van de

slagzin *Scientia non est individuorum*. Aan de hand van DBCs van individuele cliënten kan een kennisbestand opgebouwd worden. Natuurlijk is dit eerder geprobeerd en het resultaat was beperkt. Toch kun je pleiten om het met de kennis van nu nog eens te doen.

Barendregt vertelde op een bijeenkomst rond 1974 van persoonlijkheids- en klinisch psychologen hoe hij telkens een theorie formuleerde over het fobisch gedrag dat een bepaalde cliënt van hem vertoonde. Gezeten op een haastig door Hofstee aangesleepte stoel, gebogen over geblokt parket in de senaatszaal in Utrecht - zou hij een schaakbord waargenomen hebben - te midden van jonge en middelbare collega's, vertelde hij zijn verhaal. Hij bestreek daarbij de in thema I, sectie 1 genoemde bronnen en oriëntaties van theorievorming, van een pathologische trek tot een door de cultuur geïnduceerde fobie. Met de diagnose had hij wel vrede<sup>83</sup>. In 1974 verdedigde hij in *De Psycholoog* dat de diagnose niet overgeslagen moest worden bij therapie. Het lukte hem ook een theorie te formuleren voor de fobische patiënt. Hij deed ook precieze voorstellen voor behandeling: een DBC voor een cliënt *avant la lettre*. Hij was echter teleurgesteld over zijn voorspellingen van het criteriumgedrag: vermindering van de ervaren fobie, het kunnen uitvoeren van bepaalde handelingen door de 'patiënt' en het effect van zijn behandelingsvoorstellen. Er bleef naar zijn overtuiging veel te leren over. Hij suggereerde ook dat je toeval moest incalculeren. De slagzin *more research has to be done* was toen en is hier geen doodoener. Het is de vraag of dat het geijkte empirisch-analytische steekproef onderzoek is.

De overtuiging en de intuïtie van het bestaan van synergie tussen diagnose en behandeling blijven, ook al heb je niet veel bewijs in handen. Het verdient uitgezocht te worden. Je kunt daarbij buiten het *true experiment* gaan. DBCs zijn in de medische wetenschap aanvaard en daar spiegelt een diagnosticus zich aan. DBCs zijn ook aantrekkelijk om de bijna marxistische gedachte dat we veel weten maar er weinig mee doen. Is kennis er om de wereld te veranderen, je gedrag te verbeteren? Ja, een beetje en het lukt misschien ook wel een beetje.

*Conclusie:* De diagnosticus moet zijn plaats bepalen en soms veroveren in het DBC debat. De DBC gaat over de *relatie* tussen Ds en Bn en over de relatie tussen diagnostici en behandelaren. Beide(n) hebben te maken met hun kennisbestanden. Deze dwingen tot bescheidenheid over de resultaten en leiden soms tot twijfel. Bij *daten* moet je echter niet te veel twijfelen om op stoom te komen. Je moet de ander de ruimte geven. Als de behandelaar geen spoor van twijfel vertoont over zijn behandeling en verwacht dat de diagnosticus voedsel verstrekt om zijn behandeling te verbeteren, hoef je geen synergie te verwachten. Dit was de houding van de psychiater/behandelaar Kuiper ten opzichte van Kouwer, de diagnosticus/twijfelaar naar aanleiding van zijn boek *Het Spel van de Persoonlijkheid*. Als de diagnosticus zelf te zeer twijfelt over zijn diagnose kan hij de *date* vergeten. Je gaat niet met een potentiële *loser* in zee. Als de behandelaar nauwelijks gelooft in zijn interventie, voert hij die niet of kwansuis uit (behandel-nihilisme). Hij heeft nochtans enig - niet al te naïef - geloof nodig om op gang te komen en te blijven. Hij doet er van alles aan: interventies maken en benutten, zoeken naar effectieve ingrediënten, uitkijken voor iatrogene effecten en ontsporing en *true experiments* en klassieke  $n = 1$  studies uitvoeren om meer aan de weet te komen.

Om met de metafoor van relatie te besluiten: D en B beoefenaars hebben kans van slagen als

elk erkent niet perfect te zijn, elk niet de dominante partij wil zijn op basis van schoonheid en *good breeding* (pikorde in de wetenschap). Elk streeft synergie na zonder volkomen in elkaar op te willen gaan, maar ook niet strikt hun eigen weg willen gaan. Ze kunnen af en toe door één deur gaan. Ze hoeven zich niet mooier voor te doen dan ze zijn.

## Referenties

1. Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States data. *PLoS (Public Library of Science) One*, 5, 4, e 10271. DOI: 10.1371/journal.pone.0010271.
2. Fanelli, D. (2010). 'Positive' results increase down the hierarchy of the sciences. *PLoS (Public Library of Science) One*, 5, 3, e10068. DOI: 10.1371/journal.pone.0010068.
3. Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics*. DOI: 10.1007/s11192-011-0494-7. (published online: 11 September 2011).
4. Lazarus, R.S. (2006). Emotions and interpersonal relationships: Towards a person-centered conceptualization of emotions and coping. *Journal of Personality*, 74, 1, 9-46.
5. Frijda, N. (1986). *The Emotions*. Cambridge: Cambridge University Press. Nederlandse vertaling beschikbaar.
6. Witkin, H.A. & Goodenough, D.R. (1977). Field dependency and interpersonal behaviour. *Psychological Bulletin*, 84, 4, 661-689.
7. Lamiell, J.T. (2010). Why was there no place for personalistic thinking in the 20<sup>th</sup> century psychology? *New Ideas in Psychology*, 28, 1, 135-142.
8. Lamiell, J.T. & Laux, L. (2010). Reintroducing critical personalism: an introduction to the special issue. *New Ideas in Psychology*, 28, 1, 105-109.
9. Seligman, M.E.P. & Csikzentmihalyi, M. (2000). Positive Psychology, an Introduction. *American Psychologist*, 55, 1, 5-11.
10. Seligman, M.E.P. & Steen, T.A. (2005). Positive Psychology: Progress, empirical validation of interventions. *American Psychologist*, 60, 5, 410-421.
11. Seligman, M.E.P. & Fowler, R.D. (2011). Comprehensive soldier fitness and the future of psychology. *American Psychologist*, 66, 1, 82-86.
12. Casey, Jr., G.W. (2011). Comprehensive soldier fitness: a vision for psychological resilience in the U.S. Army. *American Psychologist*, 66, 1, 1-3.
13. *De Bijbel*. Auteurscollectief (2004). *De Bijbel*. Querido Jongbloed. *Het boek Job* (pp. 773-835).
14. Schopenhauer, A. (1999). *De wereld als wil en voorstelling, twee delen*. Vertaald door H. Driessen. Amsterdam: Werelduitgeverij.
15. McAdams, D.P. & McLean, K.C. (2013). Narrative Identity. *Current Directions in Psychological Science*, 22, 3, 233-238.
16. Calhoun, R.G. & Tedeschi, L.G. (2004). Posttraumatic Growth: conceptual foundation and empirical evidence. *Psychological Inquiry*, 15, 1, 1-18.
17. Thomas, Dylan (2003). *Collected Poems*. London: Phoenix.
18. Seligman, M.E.P. & Fowler, R.D. (2011). Comprehensive soldier fitness and the future of psychology. *American Psychologist*, 66, 1, 82-86.
19. Larner, B. & Blow, A. (2011). A model of meaning-making coping and growth in combat veterans. *Review of General Psychology*, 15, 3, 187-197.
20. Wortman, C.B. (2004). Posttraumatic growth: progress and problems. *Psychological Inquiry*, 15, 1, 81-90.

16. Mann, Th. (1977). Schopenhauer. In: G. Hoffmans (Ed.) *Über Arthur Schopenhauer*. Zürich: Diogenes Verlag (appeared originally in Stockholm, 1938).
17. Nietzsche, F. *Genealogie van de Moraal*. (2010). Amsterdam: De Arbeiderspers.
18. Sloterdijk, P. (2009). *Du sollst dein Leben ändern. Über Antropotechnik*. (You must change your life. About Anthropotechnique). Suhrkamp: Frankfurt am Main. Nederlandse vertaling beschikbaar: *Je moet je leven veranderen*.
19. Bohlmeijer, E., Bolier, L., Walburg, J.A. & Westerhof, G. (2013). *Handboek Positieve Psychologie. Theorie, Onderzoek en Toepassingen*. Amsterdam: Boom.
20. McNulty J.K. & Fincham, F.D. (2011). Beyond Positive Psychology? Toward a contextual view of psychological processes and well-being. *American Psychologist*, doi: 10.1037/a0024572.
21. Adolph, K.E., Young, J.W., Robinson, S.R. & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review*, 115, 3, 527-543.
22. Flynn, J.R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F. & Turkheimer, e. (2012). Intelligence: New findings and Theoretical developments. *American Psychologist* (January 2, 2012) preprint.
23. Lynn, R. (2011). *The chosen people: A study of Jewish intelligence and achievements*. Washington: Summil Publishers Augusta GA978-1-59368-036-7408.
- Lynn, R. (2013). Who discovered the Flynn effect? A review of early studies on secular increase of IQ. *Intelligence*, <http://dx.doi.org/10.1016/j.intell.2013.03.004>.
- Lynn, R. & Vanhanen, T. (Eds.). (2002). *IQ and the wealth of nations. Human evolution, behavior, and intelligence*. Westport: Praeger Publishers.
24. Piketty, Th. *Kapitaal in de 20ste eeuw*. Vertaald uit het Frans door L. van den Berg, M. Kaas, H. Klootwijk, D. Pieters & M. Sarkar. 6<sup>de</sup> druk. Amsterdam: de Bezige Bij.
25. Trahan, L.H., Stuebing, K.H., Fletcher, J.M. & Hiscock, H. (2014). The Flynn Effect: A meta-analysis. *Psychological Bulletin*, doi.org/10.1037a0037173.
26. Wai, J. & Putallaz, M. (2011). The Flynn effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces. *Intelligence*, 39, 443-455.
- Wai, J. Putallaz, M. & Makel, M.C. (2012). Studying intellectual outliers: Are there sex differences and are the smart getting smarter? *Current Directions in Psychological Science*, 21, 6, 382-390.
27. Liu, J., Yang, H., Li, L., Chen, T. & Lynn, R. (2012). An increase of intelligence measured by the WPPSI in China, 1984-2006. *Intelligence*, doi: 10.1016/j.intell.2012.01.008.
28. Batterjee, A.A., Khaleefa, O., Khalil, A. & Lynn, R. (2013). An increase in intelligence in Saudi Arabia, 1977-2010. *Intelligence*, 41, 91-93.
29. Rushton, J.P. & Jensen, A.R. (2010). The rise and fall of the Flynn effect as a reason to expect a narrowing of the Black-White IQ gap. *Intelligence*, 38, 213-219.

30. Mingroni, M.A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other Trends. *Psychological Review*, 114, 3, 806-829.
31. Woodley, M.A. (2011). Comment: Heterosis doesn't cause the Flynn effect: a critical examination of Mingroni, 2007. *Psychological Review* 118, 4, 689-693.
32. Roivainen, E. (2010). European and American WAIS III norms: Cross-national differences in performance subtest scores. *Intelligence*, 38, 187-192.
- Roivainen, E. (2012). Economic, educational, and IQ gains in Eastern Germany 1990-2006. *Intelligence*, 40, 571 - 575.
33. Shayer, M., Ginsberg, D. & Coe, R. (2007). Thirty years on - a large anti-Flynn effect? The Piagetian test Volume & Heaviness norms 1975- 2003. *British Journal of Educational Psychology*, 77, 1, 25-41.
34. Gignac, G.E. (2015). The magical number 7 and 4 are resistant to the Flynn effect: no evidence for increases in forward or backward recall across 85 years. *Intelligence*, 48, 85-95.
35. Woodley, M.A. & Meisenberg, G. (2013). In the Netherlands the anti-Flynn effect is a Jensen effect. *Personality and Individual Differences* (2013). <http://dx.doi.org/10.1016/j.paid.2012.12.022>
36. Van IJzendoorn, M.H., Bakermans-Kranenburg, M.J. (2014) Confined quest for continuity. The categorical versus continuous nature of attachment. *Monographs of the Society for research in Child development*, 79, 3: 157-167. Doi: 10.11/mono.12 120. (Voor kennismaking met het uitgebreide oeuvre over *Attachment* van Van IJzendoorn c.s. zie het Internet).
37. Bowlby, J. (1982). *Attachment*. New York: Basic Books.
38. Sroufe, L.A., Egeland, B. Carlson, E.A. & Collins, W.A. (2005). *The Development of the Person: The Minnesota study of risk and adaptation from birth to adulthood*. New York: The Guilford Press.
39. Cassidy, J. & Shaver, P.R. (Eds.) (2008). *Handbook of attachment: Theory, research and clinical applications (2<sup>nd</sup> edition)*. New York, NY: Guilford.
40. Sahdra, B.K., Shaver, P.R. & Brown, K.W. (2010). A scale to measure nonattachment: A Buddhist complement to Western research on attachment and adaptive functioning. *Journal of Personality Assessment*, 92, 2, 116-127.
41. Lilienfeld, S.O., Wood, J.M. & Garb, H.N. (2006). Why questionable psychological tests remain popular. *Scientific Review of Alternative Medicine*, 10, 6-15.
42. Winter, D.G. (1996). *Personality analysis and the interpretation of lives*. New York: McGraw-Hill.
43. Mitchell, G. (2012). Revising truth or triviality: the external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, 7, 2, 109-117.
44. Cohen, J. (1988). *Statistical power analysis for the social sciences (revised edition)*. New York: Academic Press.

45. Bacon, F. (1620/2000). *Novum Organon*. Uitgegeven door M. Silverthorne & L. Jardine. Cambridge University Press: London/ Cambridge.
- Bacon, F. (1625). *The Essayes or Counsils, civil and morall, of Francis Lord Verulam, Viscount St Alban*. London: John Haviland. Nederlandse vertaling: *Essays van Francis Bacon*. Amsterdam: Boom (1978). Het essay over Doortastendheid is interessant.
46. Carroll, J.B. (1993). *Human Cognitive Abilities. A survey of factor analytic studies*. Cambridge: Cambridge University Press.
47. Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 2, 262-274.
48. Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszin, T.W. & Reed, G.M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist*, 56, 2, 128-165.
49. Molenaar, P.C.M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, 2, 4, 201-218.
- Molenaar, P.C.M., Sinclair, K.O., Rovine, M.J., Ram, N. & Corneal, S.E. (2009). Analyzing developmental Processes on an individual level using non-stationary time series modeling. *Developmental Psychology*, 45, 1, 260-271
50. Heymans, P.G. (1994). Ongepubliceerde rapporten over meerdere casus: *De symbolische structurering van een gevreesde levensperiode: 154 dagen uit het leven van Lex. Biografische Revalidatie: Hoe worden verslaafden door biografische therapie geholpen bij hun zelfgekozen ontwikkelingstaak (december, 1996)*. Vakgroep Ontwikkelingspsychologie, Universiteit Utrecht.
- Dissertaties met n = 1 studies: Klein Wassink, M. (1999). *Betekenis in ontwikkeling*. Utrecht
- Dissertatie: (een jonge vrouw wordt gedurende 301 dagen intensief gevolgd).
- De Rooy, A. (2012). *Finding benefits in adversity: symbolic tools in adult development after life disruption*. Dissertatie Universiteit Utrecht.
51. Comenius, J. A. (1997, gepubliceerd in 1639). *Der Weg des Lichts, Via Lucis*. (Ingeleid en vertaald door U. Voigt). Hamburg, BRD.
52. Trotsky, L. (27 November, 1932). Geciteerd in T. Rütting: *Pavlov und der neue Mensch. Diskurse über Disziplinierung in Sowjetrussland*. München, 2002 (p. 179). (Pavlov and the new man: Arguing about discipline in the USSR).
53. Fodor, J.A. (1974). Special sciences (Or the disunity of science as a working hypothesis). *Synthese*, 28, 97-115.
54. Witteman, C., Van der Heijden, P & Claes, L. (2014). *Psychodiagnostiek: het onderzoeksproces in de praktijk*. Zoetermeer: De Tijdstroom.
55. Poston, J.M. & Hanson, W.E. (2010). Meta-analysis of psychological assessment as therapeutic treatment. *Psychological Assessment*, 22, 203-212.

56. Lilienfeld, S.O. (2011). Public scepticism of psychology. Why many people perceive the study of human behavior as unscientific. *American Psychologist*, DOI: 10.1037/a0023963. Pp. 1-19.
57. Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for generalized causal inference*. New York: Houghton Mifflin Cy.
58. Dumont, J.J. & Kok, J.F. (1973). *Curriculum Schoolrijpheid Deel I*. Den Bosch: Malmberg.
59. Cuijpers, P., Donker, T., Van Straten, A. Yuan, L. & Anderson. G. (2010). Is guided self-help as effective as face-to-face therapy? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, 21, 1-15.
- Cuijpers, P., Van Straaten, A. & Andersson, G. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*, 76, 6, 909-922.
60. Weisz, J.R., Weiss, B., Han, S.S., Granger, D.A. & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited; a meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117, 3, 450-468.
- Weisz, J.R., Doss, A.J. & Hawley, K.M. (2005). Youth Psychotherapy outcome research: A review a critique of the Evidence Base. *Annual Review of Psychology*, 56, 2, 337-363.
- Weisz, J.R., McCarty, C.A. & Valeri, S.M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological Bulletin*, 132, 1, 132-149.
61. Fisher, R.A. (1955). Statistical methods and scientific induction. *Journal of the Royal Statistical Society*, 17, 1, 69-78.
62. Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S. et al. (2014). Comparative efficacy of seven therapeutic interventions for patients with depression. A network meta-analysis. *Plos Med*. 10 (5) e 1001454 doi 10.1371 Journal pmed 1001454.
63. Kuiper, P.C.: Barendregt, J. (1965). 'Ingezonden' en P.C. Kuiper, 'Ingezonden'. *Nederlands Tijdschrift voor Geneeskunde*, 109, 1736.
64. Barendregt, J. (1974). De relatie van diagnostiek en therapie binnen het fobieën project. *De Psycholoog*, 9, 295-308.
65. Chambles, D.L. & Ollendick, T.H. (2001). Empirical supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-761.
66. McHugh, R.K. & Barlow, D. (2010). The dissemination and implementation of Evidence-Based Psychological Treatments. *American Psychologist*, 65, 2, 73-84.
67. Kazdin, A.E. (2008). Evidence-based treatment and Practice; new opportunities to bridge clinical research and practice, enhance the knowledge base and improve patient care. *American Psychologist*, 63, 3, 146-159.
68. Barlow, D.H., Bullis, J.R., Comer, J.S. & Ametay, A. (2013). Evidence-Based Psychological Treatments: An update and a way forward. *Annual Review Clinical Psychology*, 9, 1-27.
69. Shirk, S.R. & Karver, M. (2003) Prediction of treatment outcome from relationship variables in child and adolescent therapy: a meta-analytic review. *Journal of Consulting and Clinical Psychology*, 71, 3, 452-464.

70. Hattie, J. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. (Van internet te downloaden).
71. Daniels, A.M. (pseudoniem: A. Dalrymple) 2012: *Leven aan de onderkant*. Houten: Spectrum.
72. Van Strien, P.J. (1975). Naar een methodologie van het praktijkdenken in de sociale wetenschappen. *Nederlands Tijdschrift voor de Psychologie*, 30, 6-1-619.
73. Drenth, P.J.D. & Sijtsma, K. (2006). *Testtheorie: Inleiding in de theorie van psychologische tests en zijn toepassingen* (4<sup>de</sup> vernieuwde druk). Houten: Bohn, Stafleu van Loghum.
74. Wissler (1901). 'The correlation of mental and physical tests'. *The Psychological Review, Monograph Supplement*, 3, p. 6.
75. De Groot, A.D. (1967). *Vijven en zessen*. Groningen: Wolters.
76. Verschuieren.K., & Koomen, H. (2007) (redactie). *Handboek diagnostiek in de leerlingbegeleiding*. Antwerpen-Apeldoorn: Garant.
77. Braet, C. & Bögels, S.M. (2014). *Protocollaire behandelingen voor kinderen en adolescenten met psychische problemen*. Amsterdam: Boom.
78. Cronbach, L.J. & Snow, R.E. (1977). *Aptitudes and Instructional Methods*. New York: Wiley.
79. Boelema, S. (2014). *Alcohol use in adolescence. A longitudinal study of its effect on cognitive functioning*. Academisch proefschrift Universiteit Utrecht.
80. Cronbach L.J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
81. Loevinger, J. (1976). *Ego Development: Conceptions and Theories*. San Francisco: Jossey Bass.
- Loevinger, J. (1957). Objective tests as instruments of psychological theories. *Psychological Reports Monographs Supplement*, 3,635-694.
- Loevinger, J. (1997). Stages of Personality Development. In: R. Hogan, J. Johnson & S. Briggs (Eds.). *Handbook of Personality Psychology*. New York: Academic Press (pp. 199-208).
82. Gigerenzer, G. (2008). *Rationality for mortals. How people cope with uncertainty*. Oxford: Oxford University Press.
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.
83. Barendregt, J. (1974). De relatie van diagnostiek en therapie binnen het fobieën project. *De Psycholoog*, 9, 295-308.

## IV Omgeving van diagnostiek

*Bij een praatje - presentatie geheten - van een 'executive' van een bedrijf of instelling wordt je op het hart gedrukt dat je de omgeving moet kennen waarin je opereert. Je moet de 'grote en kleine spelers' kennen. Dit is geen voetbalmetafoor: te banaal. Het gaat om serieuze zaken. Meestal is het veld de markt of in kostschoolformulering: de tucht van de markt. De omgeving voor diagnostiek is hier bepaald als het publiek en als impliciete en alternatieve theorievorming. Ze morrelen af en toe aan diagnostiek zoals die volgens de regels van de kunst bedreven wordt. Misschien dagen ze uit en voegen ze iets toe, wie weet? De diagnosticus heeft daarnaast te maken met testexperts en psychometrici die over zijn schouder meekijken en hij moet zijn positie bepalen ten opzichte van wetenschapstheoretische stromingen.*



## 1. Expliciete theorie tussen impliciete leken- en alternatieve theorie

Mijn voorkeur voor driedelingen is niet religieus, cultureel of crimineel geïnspireerd. Het is geen verborgen hint naar de Drievuldigheid, Driehoeksverhouding of Bermuda Driehoek. Hij is wel hardnekkig. Drie bases voor psychologische theorievorming zijn onderscheiden: alledaagse, impliciete van leken: burgers en boeren, officiële, expliciete van vakbroeders en -zusters en alternatieve van dwarsliggers en buitenlui.

*Impliciete theorie:* Deze kun je van tafel vegen en dat doet Millon<sup>1</sup>. Hij wuift en passant ook nog de DSM-5 weg. Deze auteur ziet theorievorming van leken over pathologie als 19<sup>de</sup> eeuws bijgeloof. Men kan zich naar mijn mening houden aan Merleau-Ponty's<sup>2</sup> opvatting van de *expérience vécue* als basis voor begrijpen en kennen van gedrag. Sommige bèta's, bijvoorbeeld wiskundigen zijn het daar niet mee eens. Ze fulmineren tegen de geleefde ervaring als bron voor inzicht in wiskunde en benadrukken het leren van formules en operaties door te oefenen en nog eens te oefenen. De visie van het Freudenthal instituut wordt bijvoorbeeld door wiskundige Van der Kraats bestreden. Lakoff en Núñez<sup>3</sup> verdedigen daarentegen de geleefde ervaring en de realistische wiskunde in hun boek: *Where mathematics come from? 'How the embodied mind brings mathematics into being'* (onderlijning jtl). Misschien hebben ze het over verschillende zaken. We beginnen in ieder geval allemaal met de geleefde ervaring. Dat is de sensomotorische periode die Piaget zo minutieus beschreef. Kijk naar een baby of peuter: zo beginnen ze: bewegend, waarnemend, lerend! Leren en doorgronden van wiskunde vergen daarnaast en later lang en veel rekenen, regels leren en oefenen en nog eens oefenen.

*Sceptische leken:* Dautzenberg<sup>4</sup> is een voorbeeld van het wantrouwen van het lekenpubliek: *Vijftwintig jaar geleden kwam ik voor het eerst in aanraking met de wetenschap (als student), daarna beroepsmatig (als adviseur, onderzoeker, journalist en schrijver). Sindsdien zie ik de universiteiten als hoeders van fantastische constructies, als saaie en bovenal bureaucratistische toverkastelen waarin niettemin lekker geheimzinnig gedaan wordt over weinig tot niets. De wetenschap heeft de samenleving absoluut verrijkt - met illusies, raadsels en enkele praktische handvatten.* Gelukkig dat laatste nog wel. Je kunt er een *argumentum ad hominem* tegen aangoeien: 'Wie denk je wel dat je bent; je bent niet eens gepromoveerd' of zeggen, dat je serieus met wetenschap bezig bent en dat er nog veel uitgezocht moet worden. Dat bevredigt niet. Er hoort een antwoord op te komen. Jammer dat er weinigen op ingaan en denken dat het wel overwaait. Je zou zeggen doe het een en laat het ander niet. Maar zo gaat het niet want we houden van eenvoudige tegenstellingen en lezen anderen graag de les. Het langs je af laten glijden van deze schimpscheut is een vorm van wegstijgen.

*Expliciete theorie* is een gedurende enige tijd geldend, dominant paradigma volgens Kuhn<sup>5</sup>. Dat zou vooral op exacte wetenschappen van toepassing zijn. Kuhns definitie van paradigma is waarschijnlijk een constructie achteraf om een mogelijk grillig verloop in categorieën onder te brengen. Exacte wetenschap beoefenen is immers gedrag en dus kan het daar ook rommelen.

In de psychologie zijn er geen paradigmashifts in de zin dat een theorie het veld ruimt voor een andere. Er zijn wel verschillende perspectieven op wat gedrag is, wat de mens en zijn gedrag zijn en waarvoor ze dienen.

*Alternatieve theorie* ontstaat in de psychologie uit onvrede met bestaande theorie en methodologie. De expliciete theorie laat gedragskenmerken liggen, zet ze tussen haakjes, ontkent hun betekenis en soms hun bestaan. In de exacte wetenschap doet dit zich ook voor, denk aan de opvatting van licht als deeltjes of als golven. Daar komen ze er echter meestal wel uit zodat een paradigma gedurende een tijd dominant kan zijn tot een ander het opvolgt. Ze kunnen soms ook naast elkaar blijven bestaan. Of een paradigmashift nu mode is of van een nieuw inzicht getuigt is in de psychologie moeilijker vast te stellen dan in de exacte wetenschap. Of daar mode ook van mode sprake is kan ik niet beoordelen maar het zal wel een beetje. Exacte wetenschap is ook mensenwerk.

Een recent voorbeeld van alternatief denken in de fysica is de verklaring/opvatting van Verlinde<sup>6</sup> over het ontstaan van het heelal en de aard van de natuurwetten. Het wijkt af van het gangbare scheppingsverhaal zonder God, zoals Robbert Dijkgraaf dat op TV zo prachtig over het voetlicht heeft gebracht. Voor de aarde is de zwaartekracht centraal en verklaart het gedrag van objecten. De oorsprong is de *Big Bang* theorie. Kennelijk was er daar voor niets; toen een knal, en er was iets, dat in razend tempo uitdijde. Dat is een realistische voorstelling om 'al wat is' te beschrijven/verklaren. Het begint met iets werkelijks, hoe klein ook; het is bijna niks, maar toch niet niks en het werd heel veel, misschien wel oneindig veel. Het uitdijen houdt misschien nooit op, al kunnen we ons daar geen voorstelling van maken. Verlinde heeft een formele kijk ontwikkeld. Hij beschouwt de natuur en uiteindelijk ook de mens als een oneindige veelheid van mogelijkheden en informatie. Het klassieke idee van de zwaartekracht plaatst hij tussen haakjes. Het is volgens hem een entropische kracht die veroorzaakt wordt door veranderingen in de informatie die verbonden is met de positie van materie. Alternatieve theorieën komen en gaan. Een enkele is een blijvertje en wordt opgenomen in de expliciete. Ons denken, kennen, waarnemen en voelen houdt het kennelijk niet bij één verklaring of l'ezing. Soms ziet men stagnatie, maar er komt toch steeds weer iets nieuws, want men heeft het gevoel, het idee uitgepraat te zijn over een thema. Het is ook overal en altijd wat.

*Hoe om te gaan met de bronnen?* De onderzoeker kan zich bij één bron houden. Dat is de expliciete theorievorming. Daarnaast *onderzoekt* hij impliciete theorievorming; dat is immers *gedrag* van leken. Zijn onderzoek laat meestal zien hoe gebrekkig die is in het licht van het expliciete kennisbestand. Meestal blijft hij bij één - zijn - bron. Dat publiceert het gemakkelijkst. De diagnosticus kan dat niet. Hij moet de drie ballen tegelijk in de lucht houden. Het kennisbestand van de expliciete past hij toe op één cliënt. Om impliciete theorie kan hij niet heen omdat de cliënt niet als een onbeschreven blad bij hem aanklopt. Alternatieve theorievorming is epistemologisch en in het gebruik van gedragsconstructen vrijmoediger dan de expliciete. Ze laat naast de dominante empirisch-analytische kijk op gedrag de hermeneutische en kritische toe. Deze verruimen de mogelijkheden het probleem/gedrag van zijn

cliënt te begrijpen en een rijke en voldoende beschrijving te construeren van het probleem, de vraag, al staat nergens hoe dat precies moet. Zijn aanpak houdt het midden tussen 'rupsje nooit genoeg': wanneer is de beschrijving nu voldoende en de 'Franse slag': 'ik zie, weet het al!' Protocollen zijn gewild en behulpzaam want ze bevatten het door schade en schande wijs worden.

*Er is overmaat, want al het oude blijft:* In de psychologie is er een overmaat, een teveel aan theorieën en methoden. Studenten en sommige docenten worden er horendol van. Ze blijven vermeld worden in hun tekstboeken, ook al wordt postmodernistisch beweerd dat ze relictten uit het verleden zijn (Mroczek & Little<sup>7</sup>). Nieuwe psychologische theorievorming verdringt oude niet omdat ze fout of onjuist is. Als het goed gaat, verfijnt ze en voegt nieuwe informatie, beschrijvingen, interpretaties en verklaringen toe. Dit houdt in dat onderzoek en waarheidsvinding in de psychologie het verschuiven van perspectief en het verleggen van de horizon is; letterlijk zoals satellieten, onze verlengde ogen, steeds verder het heelal in kijken. Ze bieden nooit het definitieve verhaal. Daarvoor moet je naar goeroes, tovenaars en heilbelovende instituten. Ze verbreedt en verdiept kennis van gedrag. Empirisch onderzoek laat steeds weer zien dat we factoren over het hoofd zien en - als we ze beet hebben - ze over- of onderschatten. Bovendien kunnen we de goede mix tussen de factoren maar niet vinden. Door die overmaat zie je door de bomen het bos niet meer. Dat is een basis voor begrijpelijke klachten van studenten en practici.

*Waarom alternatieven?* Alternatieve theorieën en methoden vergezellen en reageren op expliciete theorieën, constructen en methoden. Er wordt iets toegevoegd of afgewezen, omdat interessante en relevante inhoud en/of methoden ontbreken. Oppositie tegen bestaande theorie komt ook voort uit de overtuiging dat zij van verkeerde veronderstellingen over de aard van de mens uit gaat. Soms er is verzet tegen iedere vorm van theorie. Een eerste motief voor alternatieve theorievorming berust op de vaststelling, overtuiging dat er al wel veel theorieën zijn, maar de ware nog niet. Voorbeelden zijn Gardners<sup>8</sup> intelligentie-theorie, Czikszentmihalyi's<sup>9</sup> (1988) opvatting over persoonlijkheidsgroei (*flow*) en Seligmans en Czikszentmihalyi's<sup>10</sup> *Positive Psychology*. Dit motief doet zich ook bij religies voor: De Paulus<sup>11</sup> van het Nieuwe Testament legde de basis voor het christelijk geloof door het los te maken van het Griekse veelgodendom. En, als dat van pas kwam, weer te verbinden met het Joodse Oud Testamentische geloof. Augustinus (4<sup>de</sup> eeuw) schreef een theologie door zich af te zetten tegen het Manicheïsme en de Romeinse godsdienstige praktijken van zijn tijd. Mohammed deed hetzelfde. Hij incorporeerde elementen van bekende religies en voegde nieuwe elementen toe. Dit alternatieve denken kun je naïef noemen. Vooral aanhangers - misschien niet de ontwerpers - denken te geloven en te weten dat ze geloven wat ze geloven. En niet naïeve gelovigen/volgelingen zijn weliswaar bereid chronisch te twijfelen, maar hun houding is precies de brandstof om een systeem of religie levend te houden. De Nederlandse Remonstranten zijn daar een voorbeeld van. Ze staan twijfel toe, maar maken tegelijkertijd *in prime time* reclame voor hun geloof op de radio. Allemaal goed bedoeld, daarover geen misverstand.

Een *tweede* motief voor alternatieve theorievorming is de observatie en overtuiging dat de oude niet toereikend zijn, overboord gegooid moeten worden, omdat ze inhoud *overwaarden*. Het gaat immers maar door en de ene opvatting, regel en inhoud buitelt over de andere heen, denk aan de regels van bedrijfscoaches en de pestprogramma's. Sommige psychologen en psychometrici zeggen dat de betekenis van gedragingen slechts bestaat in *the eye of the beholder*. Ze geven de voorkeur aan heldere, onderscheiden, lege structuren: functies en formules. Dit motief kan onschuldig uitwerken als een meta-religie, een iets-isme of een abstracte psychologische meta-theorie, bijvoorbeeld: alle gedrag is beslissen; alle gedrag is te verklaren met de mathematische speltheorie. Abstract wil zeggen: los van elke inhoud of alle inhoud overstijgend. Deze theorie brengt ideeën voort waar aanhangers zelf niet in hoeven te geloven. Het kunnen gelovers zonder geloof of theoretici zonder theorie zijn. Er is immers niet iets concreets.

Er is een *kwaadaardige uitwerking* van dit tweede motief denkbaar. Het doet zich voor als wetenschappelijke of bestuurlijke systemen/doctrines, religies of goede doelen organisaties gebruikt worden om je willekeurige, levensvreemde, niet-kloppende en niet te realiseren inhouden in de maag te spritsen. Een collega vertelde me eens dat 'de Jap' in de interneringskampen onbegrijpelijke straffen uitdeelde voor iets dat je bij God niet als een vergrijp kon zien. Precies dat: inhoud doet er niet toe, als je maar boven iemand kunt staan. *Het Slot* van Kafka<sup>12</sup> heb ik ook zo gelezen: landmeter Joseph K. is wanhopig op zoek naar inhoud en naar een concreet persoon: de kasteelheer maar hij krijgt geen van beide(n) te pakken. Ondertussen beheersen ze wel zijn leven. Deze roman was onvoltooid; symbolischer kan het niet. Het opleggen van zo'n systeem wordt op een slimme manier gedaan. Je hebt het niet meteen door. De burger is immers niet zo maar om de tuin geleid. Toch worden uitingen als 'Psychologische Theorie'; 'Onderzoek heeft aangetoond dat ...' een enkele keer gebruikt om mensen dingen te laten kopen die ze niet nodig hebben, schulden te laten maken, geld af te troggelen of om militair succes te verkrijgen in dubieuze oorlogen. Historische voorbeelden zijn: R.K. pausen - bijvoorbeeld de Medici - die de christelijke religie aan hun laars laptten en oorlogen voerden om hun bezit te vergroten. Franse bisschoppen maakten afspraken met Islamitische leiders om katholieke collega's in Oostenrijk te vernietigen. Deze uitwerking wordt gekenmerkt door een dogmatische organisatie, een structuur die illusies schept, programma's maakt om het geloof, de theorie te verbreiden, denken en onderzoek te promoten of financiële producten te verkopen. Het wordt gedaan voor aanzien, status, macht en rijkdom van de leiders: je hebt er toch maar een mooie grote bank neergezet door je slimme acquisities. Dit is gechargeerd, want meestal is het een glijdende schaal en een zich langzaamaan loszingen van de werkelijkheid.

Protagonisten van dit soort systemen weten te winnen in de jungle van egoïsten. Weer G. Rudyard Kipling: 'Wat is de wet van de Jungle' vroeg Bagheera. 'Eerst slaan en dan waarschuwen'. Ze accepteren geen kritiek, denk aan Hubbard, Markies de Sade, sommige CEO's, bankdirecties en dictators. Ze zijn het contact met wat er in de samenleving gebeurt kwijt en hebben dat ook niet nodig. Ze zetten immers iets neer dat groter is dan zichzelf, groter dan de mensen, groter dan de samenleving, groter dan gewone stervelingen kunnen bedenken. Ze doen het in dienst van iets 'Abstracts' en daar mogen koppen bij vallen. In milde vorm doet dit zich wellicht voor in de psychologische zelfhulpindustrie en

managementtrainingen. Barendregt<sup>13</sup> noemde dit 'De zielenmarkt'. De inhoudten volgen elkaar in een zodanige sneltreinvaart op dat het lijkt dat inhoudelijke acties en interventies er nauwelijks toe doen. Er zijn veel programma's en onderzoekers willen dan ook orde op zaken stellen door onderzoek naar hun effectiviteit en werkzame ingrediënten.

*Alternatieve methodologie:* Kritiek op expliciete methodologie bestaat bijvoorbeeld uit bezwaren tegen het gebrek aan aandacht voor het individu. Er wordt gewezen op veronachtzaming van  $n = 1$  analysetechnieken, kwalitatieve methoden, zoals gesprek, interview, de dialoog met de cliënt en longitudinale tijdreeksen bij één persoon. Het constructivisme - een specifieke uitwerking van de hermeneutische epistemologische stroming - stelt voor om tests en vragenlijsten in de prullenbak te gooien want ze dienen om groepen te onderscheiden van andere en om bestaande verhoudingen in de samenleving te consolideren. De experimentele onderzoeksmethodologie maakt volgens sommigen de cliënt tot een ding te midden van andere dingen, tot de normproefpersoon van de kritisch psychologen. Zij lappen op hun beurt het onderscheid tussen onafhankelijke en afhankelijke variabele aan hun laars.

*Conclusie:* Impliciete theorie is voor de diagnosticus van belang omdat de cliënt met zijn eigen theorie de spreekkamer binnenkomt. Het publiek bekijkt psychologie en diagnostiek wantrouwend van ex-minister van onderwijs Plassterk: 'de bèta's weten tenminste waar ze het over hebben' tot polemisten die psychologie/diagnostiek als wetenschap en haar toepassingen laken. Alternatieve theorie is onstuitbaar en bevat soms een aanzet tot vernieuwing. Iedere vernieuwing is echter geen succes. Houd maar eens uit elkaar wat mode is en voorbijgaat en wat je helpt het probleem, de vraag de cliënt beter te beschrijven. Wat is een zinvol nieuw construct of verzameling constructen: DBCs, Emotionele Intelligentie, DSM-5, EMR, de Polygraaf? Ze zijn zo populair dat de afkortingen al voldoende zijn. 'De letters G O D, ik combineer er sleutelwoorden mee', zei de dichter Achterberg ooit. En sommige bringers van nieuwe theorie interesseren zich niet voor die theorie. Het gaat hen niet om de inhoud, waar U zich druk over maakt en ook niet om Uw motieven. Ze willen levensvreemde, in onze ogen onnodige structuren realiseren, zoals je bij sommige religies, sekten en zelfs instellingen ziet. De diagnosticus kan niet om alternatieve theorie heen. Hij heeft te maken met alle drie bronnen. Voor hem is de verhouding van de drie in tegenstelling tot de onderzoeker *niet* hiërarchisch. Er is geen nummer één. Omdat hij ook niet alles tegelijk kan, zal één bron wat meer in het vizier zijn, maar dat neemt niet weg dat hij alle drie recht probeert te doen in een kritische dialoog. Daar hoeft geen bron het te winnen en er mag onenigheid blijven bestaan. Toch moet hij de boel bij elkaar houden. Zijn conclusie is niet als de conclusie van een artikel volgens de vereisten van expliciete theorievorming in een empirisch-analytisch kader. Ieder zijn vak.

## 2. Psychometrie en Diagnostiek

De practicus, hulpverlener, gedragswetenschapper wordt met de neus op methodologie en psychometrie gedrukt. Dat is in de exacte wetenschappen nauwelijks het geval en in de geesteswetenschappen minder. Komt ervan als je tussen de alfa's en de bèta's vertoeft. Als de interiorisatie van de methodologie slaagt, vormt zich het Streng Methodologische Geweten. Dit is nog geen DSM categorie. Je kunt je er niet voor laten behandelen, maar wie weet wat de toekomst brengt.

Psychometrie herinnert de diagnosticus er aan dat hij instrumenten moet gebruiken die op een formeel meetmodel berusten. Bovendien is de diagnosticus als gebruiker van het psychologisch instrumentarium niet vrij. De overheid vraagt sedert enige tijd om tests en vragenlijsten te gebruiken die een keurmerk van de Cotan hebben. En, als hij tot behandelen overgaat of een behandeling aanbeveelt moet hij zich er van vergewissen dat de interventie *evidence-based* is, zoals uitgesproken door een erkenningscommissie.

*Het Streng Methodologisch Geweten: Aan Barendregt<sup>13</sup> werd eind jaren zestig gevraagd om een boek over methoden voor klinisch psychologen te schrijven. Hij gaf de opdracht een mooie draai en actualiseerde de karakters van Theophrastes, een leerling van Aristoteles. Eén laat het Streng Methodologisch Geweten zien: De bijgelovige bidt tot de god van de methodologie. Hij beschouwt nutteloos onderzoek als volmaakt. De vreze des heren (de methodologie) is het begin van alle wijsheid. Hij bidt de hele dag, en als hij een erectie krijgt kijkt hij eerst naar het apparaat of de penuitslag voldoende is. Zo niet, dan beperkt hij zich tot de bewondering van het meisje. Eén ander laat de zorg voor het instrument zien: De test- en instrumentenpoetser sleutelt en sleutelt aan de instrumenten en maakt van het middel een doel. Hij verandert bestaande apparatuur, tests, al voor hij ze gebruikt heeft. Hij ontwerpt instrumenten zo dat ze op hele groepen toegepast kunnen worden, ook al gebeurt dat nooit. Elke bevinding laat hem koud, zelfs een belangrijke, totdat de uitkomst elektronisch geregistreerd kan worden.*

*Diagnostiek en testtheorie en tests: De diagnosticus kan niet om de klassieke en moderne testtheorie heen. Daarnaast houdt hij zich aan de Cotan-testbeoordelingen. Deze commissie stimuleert gebruik van valide tests en de overheid weegt haar beoordelingen mee. Het gekrakeel over tests en interventieprogramma's is daar een voorbeeld van. Je kunt niet zo maar een test of programma pakken en kijken of gebruikers er vrolijk van worden. Ze zijn bij voorkeur predictief valide en tonen in experimenteel onderzoek significante resultaten, dat wil zeggen verbeteringen in typerend gedrag of prestaties. De vereisten voor tests zijn uitvoerig beschreven in zeven rubrieken (Cotan<sup>14</sup>) en er zijn in Nederland meer dan 500 tests en vragenlijsten beoordeeld op de criteria uit de rubrieken. Het uitvoeren van een psychologisch experiment is vaak beschreven en is een onderdeel van de BAMA opleiding in de sociale wetenschappen.*

De verhouding tussen diagnostiek en psychometrie is goed voor milde animositeit. Inhoudelijke theoretici & professionals en psychometrici hebben zo hun eigen kijk op vooruitgang in de psychologie. Psychometrici zijn soms verbaasd dat hun Item Respons modellen zichzelf

niet verkopen. Ze zeggen dat inhoudelijke theoretici en diagnostici semantische interpretaties verbinden aan items en relaties veronderstellen tussen de gedragingen. Ze hebben of kennen echter geen formeel model dat kan toetsen of die veronderstelling deugt. Ze schrijven items en verzinnen opdrachten uit de losse pols. Ze tellen de goede antwoorden of het aantal keren 'ja zeggen' op, alsof meer altijd beter/hoger is. Zo menen ze te kunnen zeggen dat de een intelligenter is of extravarter dan de ander. Dat kan niet, omdat niet bewezen is dat er één of meer schalen, één of meer categoriesystemen aan ten grondslag liggen. Misschien zijn de vragen en opdrachten wel appels en peren die niet opgeteld kunnen worden. De inhoudelijk theoreticus en de diagnosticus maken zich er niet druk om zolang ze 'gelijk krijgen', hetzij op basis van *face validity* hetzij op basis van empirisch bewijs in de vorm van correlaties en d-waarden.

*Relatie tussen psychometrie en diagnostiek:* Een eerste uitwerking van een relatie is elkaar mijden of beweren dat je niets met de ander te maken hebt. De relatie is een non-relatie. Het past niet: ieder kan zijns weegs gaan. Sommige exacte wetenschappers en een enkele psycholoog noemen meten van gedrag onmogelijk. Het meetprobleem bestaat dus niet. De tweede uitwerking is de opvatting dat ze een te overbruggen kloof vormen. Bij de derde uitwerking herkennen we een bekend relatieprobleem. De een stelt de ander in gebreke en vice versa. In de vierde uitwerking omzeilt de psycholoog de relatie en onttrekt zich aan een antwoord op wie het beter weet. Daarnaast zie je ook soms de 'kata holos' of 'totus' gedachte van integratie: het én-én in plaats van het óf-óf. Je mag, kunt dat naïef noemen. Het werkt meestal niet.

*Gedrag is non-kwantitatief:* De vlieger van het meten gaat alleen op voor objecten en natuurkundige verschijnselen: voor *res extensa* en niet voor *res cogitans*. In de 21ste eeuw zegt Saint-Mont<sup>15</sup> het bijna precies zo: *...psychological phenomena are meaningful, relational, non-extensive, interactive, socio-culturally and historically constituted phenomena with moral and political significance. All these attributes of psychological phenomena are non-quantitative.* Dit is duidelijk, maar psychologen kwantificeren met gemak: dit is je score op intelligentie, dit is je *Big Five* profiel, dit is je prestatie op de Cito-toets, dit is je MMPI profiel. Als iets bestaat, bestaat het in een hoeveelheid en kun je het meten: de lengte, breedte, hoogte van Descartes: *res extensa* dus. Als het resultaat van kwantificeren van een kenmerk van een object, verschijnsel of van gedrag herhaalbaar is (betrouwbaar/stabiel) en zinvolle informatie oplevert (valide) dan beweer je iets van betekenis over het gedrag van een cliënt. We hebben zo toegang tot gedrag en we kunnen nagaan of het invariant is bij herhaling en bij gebruik van andere procedures en of we er iets mee kunnen dat ons aanbelangt: begrip, inzicht verwerven, voorspellen, controleren, beslissen. Ik ga niet mee in de redenering van Saint-Mont. Hij beroept zich ook nog op een fysicus die de beperkingen van zijn modellen ziet, als je de werkelijkheid ermee wil beschrijven en verklaren. Jaynes<sup>16</sup> zegt dat mathematische strengheid niet juistheid van resultaten garandeert. Hij beweert dat naarmate er meer nadruk ligt op abstracte wiskunde er minder aandacht wordt besteed aan de geldigheid van de premissen in de echte wereld. Zo beschouwd, heb je er misschien niet zoveel aan in de

psychologie en de diagnostiek als empirische activiteiten. Daar gaat het er om een geobserveerd geheel van relaties formeel af te beelden. Het vastleggen van gedrag is de eerste stap. Na beschouwer Saint-Mont, een echte bèta, een ingenieur: Trendler<sup>17</sup>. Echte wetenschap is exacte wetenschap. Daarbuiten zijn er enkel onderhoudende verhalen van en over mensen. Aan het laatste maken psychologen zich schuldig. Hij beweert dat meten in de psychologie onmogelijk is. Meten vereist kwantificeren onder complexe condities. Het axioma is dat omvang van een grootheid, kwantiteit groter, kleiner of gelijk moet zijn dan, aan een andere grootheid: concatenatie. Die term kreeg ik als Nijmeegse 2<sup>d</sup>e jaars student in het college 'Variantie Analyse' van Roskam te verwerken. Ik heb zijn handgeschreven en daarna gestencilde collegedictaat nog. Op een congres vertelde ik hem dat ik er niets van begreep. 'Maar nu toch wel', zei hij, en, 'Ik ben er van overtuigd dat mijn stelling over meten juist is'. Psychologische kenmerken zijn onvoldoende te controleren en te manipuleren om na te gaan of aan dat axioma voldaan wordt. Psychologische experimenten bevatten geen manipulaties, zoals in de natuurkunde en leveren geen Newtoniaanse wetten op. Als psychologen dat pretenderen, bijvoorbeeld door uiteindelijk 100% variantie te zullen verklaren en correlaties van  $r = 1.00$  te bereiken, doen ze alsof.

Trendler vergelijkt het thema theorie-meten met het *mind-body* probleem. Daar bestaat ook geen oplossing voor: *In general, faced with the measurement problem, psychology is left with the following alternatives: discuss the topic from here to eternity, as Markus and Borsboom<sup>18</sup> suggest, seriously attempt to find out if psychological attributes are measurable (e.g., Kyngdon<sup>19</sup>, or look for other solutions....'* (pp. 18). En: *The measurement issue in psychology is not ontological but epistemological in nature. That is, in reality psychological attributes might be quantitative, but if they really are, we do not know ('ignoramus') and will never know ('ignorabimus')* (p. 19). Deze Latijnse werkwoorden voeg ik toe; ze staan in de titel van zijn artikel. En verder: *Ironically the result of the effort of making psychology more scientific by modeling it after physics is that it has become a perfect pathological science in the sense that it now perfectly mimics natural sciences. Ten slotte: Psychologists perform experiments, analyse data, test hypotheses, publish results in journals, organize conferences, establish psychological institutes, award prizes and so on...., they act like physicists, but the essential difference is while physics accumulate real replicable knowledge, no such consequence is to be found in 'Newtonian' psychology.* De citaten liegen er niet om. Deze opvatting is voor een deel in 1982 al naar voren gebracht door Roskam<sup>20</sup>. Hij baseerde zich ook op Newton: psychologie moet universele gedragswetten opsporen. Hij toonde zich later in zijn loopbaan een aanhanger van Rasch.

Er rest psychologen, diagnostici volgens Trendler niets als te 'doen alsof'. Maar, waarom mag dat niet? Kant argumenteerde in zijn *Kritik der reinen Vernunft* dat we het object (en het subject) *an sich* niet kunnen kennen, maar wel kunnen doen alsof volgens zijn *Kritik der praktischen Vernunft*. Je kunt je het leven gemakkelijk, of zo u wilt, moeilijk maken, door een elegant, in zichzelf helder en geliefd model te kiezen, waar geen werkelijkheid aan beantwoordt. Nu is de werkelijkheid misschien soms niet om aan te zien, maar velen hebben niets anders en daaronder bevinden zich diagnostici.

*De kloof overbruggen:* De psychometricus Guttman<sup>21</sup> wilde het debat over meten en over meten tegenover theorie beëindigen. Dit blokkeert volgens hem de communicatie tussen psychometrici enerzijds en theoretici & diagnostici anderzijds. Het is een rem op instrumentontwikkeling. Volgens Guttman is het een taak van de meettheorie om structurele hypothesen te formuleren en in een facetontwerp onder te brengen. Een voorbeeld van een dergelijk facetontwerp is een studie van Stouthard en Peetsma<sup>22</sup> over het meten van het tijdsperspectief. Ze definiëren eerst de extensie, dat wil zeggen het korte en lange termijn perspectief. Ten tweede noemen ze relevante objecten, bijvoorbeeld schoolloopbaan, beroep, persoonlijke ontwikkeling en sociale relaties. De items moeten de facetten: extensie en objecten weerspiegelen. Het geheel wordt in kaart gebracht met behulp van een karteringszin, die naast de facetten steekproeven bevat (bijvoorbeeld jong en oud), een houding (bijv. positief en negatief) en reacties op de objecten (affectief, cognitief en waarneembaar gedrag). De structuur wordt getoetst met confirmatieve factoranalyse. Het formuleren van een karteringszin vraagt conceptuele arbeid en het schrijven van passende items is tijdrovend. Jammer, dat het niet vaker gebeurt. Guttman hield rekening met de semantische inhoud van een theorie én met eisen van een meetmodel.

Wilson<sup>23</sup> treedt in het voetspoor van Guttman maar verwijst niet naar zijn werk. Hij reflecteert op zijn 30-jarige loopbaan als psychometricus en wil recht doen aan inhoudelijke theorie én statistische modellen. Hij neemt een *splendid isolation* tussen de twee waar. Als voorbeeld noemt hij zijn *Saltus* (sprong) model dat ontwikkelingsstadia beschrijft. Het is een speciale instantie van een confirmatief gemengd IRT model met lineaire restricties op reeksen itemmoeilijkheidsgraden voor verschillende populaties: bijvoorbeeld leeftijdsgroepen. Het model had weinig invloed. Het is mogelijk dat stadiummodellen niet (meer) populair zijn of de ingewikkelde vergelijkingen afschrikken. Het is ook denkbaar dat zo'n model iets oplevert dat erg voor de hand ligt: een ouder subject is meestal als persoon rijper en een stapje verder in de cognitieve ontwikkeling.

De afnemende belangstelling voor stadia is te zien aan het werk van Van der Maas<sup>24</sup> van de UVA. Het eerste enthousiasme voor de chaostheorie om bijvoorbeeld Piagetiaanse stadia te beschrijven is verdwenen. Het levert niet zo veel op: je vindt niet zo gauw van die Guttman-schaal-achtige sprongen, overgangen. Bovendien zijn er niet-lineaire-dynamische modellen, die dezelfde verschijnselen soepeler, dat wil zeggen probabilistisch modelleren. Het argument dat het *Saltus* model te complex zou zijn, bestrijdt Wilson want moderne technieken zijn nog veel ingewikkelder, al komen de resultaten met een druk op de knop te voorschijn.

Wilson formuleert vier vragen om psychometrie en inhoudelijke theorie samen vorm te geven. Ten eerste dienen beide een antwoord te geven op wat het is dat we meten. Dat is werkelijk gedrag dat op de ene of andere manier geordend c.q. te ordenen is: het is geen bulktransportachtige brei. Daarvoor moeten psychometrici, theoretici en professionals samenwerken. Ten tweede moeten manieren ontworpen worden om te observeren en zich zo toegang tot een construct, dat wil zeggen de vormgeving van wat iets is, te verschaffen. Dit vraagt dat zij gezamenlijk klassen en typen items, taken, vragen ontwikkelen. Ten derde vraagt het scoren van resultaten van observaties een gezamenlijke inspanning. Ten vierde moeten de gescoorde

gegevens gemodelleerd worden. Dit vereist de transformatie van antwoorden op de vragen in te schatten parameters. Dat is een kerntaak van psychometrici.

Wilson geeft een voorbeeld aan de hand van zijn instrument, dat hij *Understanding of Data Display* noemt. Het is gemaakt is voor middelbare scholieren. Hij ontwerpt eerst een *map* (een wegenkaart) met constructen om de eerste vraag te beantwoorden. Dit lijkt op Guttmans facetontwerp. Items worden geschreven om toegang tot het construct te krijgen: de empirische specificatie/realisatie. Wilson erkent dat het beantwoorden van ieder item de inzet van verschillende vaardigheden vraagt, maar er moeten items gevonden worden die vooral de bedoelde vaardigheid in een steekproef uitlokken. De items worden 'goed-fout' gescoord en hij weegt de argumentatie voor het antwoord mee in de scoring. De derde vraag gaat over scoring en codering. Als men een reeks categorieën voor ogen heeft kan men nagaan of ze gerepresenteerd en onderscheiden zijn. Dit kan het toekennen van antwoorden aan een stadium zijn of het bepalen van een positie op een ordinale of intervallschaal. Het antwoord op de vierde vraag is het maken van een model: *scaling* en *IRT modeling*. De mate van passing van de gegevens in een model kan bepaald worden met stressmaten.

Hij ziet uit naar samenwerking tussen psychometrici en theoretici. Ze zouden gezamenlijk kunnen publiceren in inhoudelijke en psychometrische tijdschriften. Een vrome wens, geuit aan het eind van een loopbaan. Hij zegt het zo (p. 20): *...we (psychometrici) need to transcend the statistical modeling box of our thinking that currently proscribes what is seen as serious work*. De auteur wil af van de verwijten dat psychologische theorie *sloppy* is en statistiek en testtheoretische modellen een mathematisch circus zijn.

*Je moet je leven veranderen*: In een relatie vraagt de een de ander en vice versa soms, af en toe om een beetje te veranderen. *Du sollst dein Leben ändern! Man, word eens volwassen!* De een stelt de ander in gebreke. Kyngdon beweert dat het in de psychologie niet schort aan experimentele apparatuur en controle, maar aan psychologische kwantiteiten en de beschrijvende theorieën van de systemen, waarin die kwantiteiten fungeren. De *sketchy and incomplete theory formation* staat meten in de weg. Sijtsma<sup>25</sup> wil meten en theorie recht doen. Als ex-voorzitter van de Cotan is hij vertrouwd met het handwerk van testconstructie. Psychometrici komen vaak pas in zicht *nadat* de items geschreven en de data verzameld zijn. Dat is jammer, maar het is de realiteit. Items kunnen niet zonder meer opgeteld worden. Er moet een schaal aan ten grondslag liggen volgens een modern testtheoretisch model. Er is pas sprake van meten als items passen in een functie, bijvoorbeeld in een van de Item Respons Functies. Dit is het geval bij tests, vragenlijsten en attitudeschalen die onder het regime van de IRT geconstrueerd zijn. Het Cito publiceert alleen instrumenten die aan een zelf ontworpen *relaxed IRT model* voldoen. Sijtsma komt tot de slotsom dat meetproblemen in de psychologie niet opgelost worden, omdat goed ontwikkelde theorie ontbreekt. De relatie meten-theorie lijdt onder zwakke psychologische theorievorming. In 2006 beweerde hij<sup>26</sup> dat toetsing met IRT de beste persoonlijkheids- en intelligentietheorie aan het licht zou brengen. Is er sprake van voortschrijdend inzicht, van teleurstelling, van beide? In zijn artikel uit 2012 wordt niet verwezen naar Guttmans facet analyse.

*Een andere weg, een omweg:* Smedslund<sup>27</sup> ziet niet te nemen hindernissen voor de logisch positivistische, empirisch-analytische psychologie. Deze veronachtzamen kenmerken van gedrag die we onmiddellijk begrijpen. Gedrag van objecten, verschijnselen en van dieren verschilt wezenlijk van dat van een mens. Hij is immers open voor een onbegrensd aantal mogelijke invloeden dat zijn gedrag kan bepalen. Zijn gedrag is onomkeerbaar want de mens is geen *tabula rasa*, heeft een geheugen en leert van zijn ervaringen. Hoeveel en wat precies is lastig te bepalen, maar je moet het wel proberen te achterhalen of te schatten. Hij is onderdeel van een gedeeld systeem van betekenissen door afkomst, gezin, school, buurt en cultuur. Hij is uniek door zijn niet-gedeelde ervaringen. Het is niet ironisch bedoeld, als Smedslund zijn artikel de titel geeft: *What reseachers must learn from practitioners*. De kenmerken klinken plausibel en de diagnosticus kan er zijn voordeel mee doen. De psychologie moet zijns inziens aansluiten bij sociologie, taalkunde en culturele antropologie. Hij acht de logisch positivistische epistemologie van de psychologie en de methodologische nadruk op vertekeningen en fouten eenzijdig, vreemd en soms niet plausibel.

Zoals vaker hebben partijen in een relatie beide een beetje gelijk. De vier opvattingen zijn niet te integreren. Er is geen mediatrix/mediator die dat voor elkaar krijgt. Toch is dat een behoefte want we houden van eensgezindheid en eenvoud in relaties. Als we erkennen dat het niet lukt, is er geen andere mogelijkheid dan meer ballen in de lucht te houden. Meten is niet het voorportaal tot universele gedragswetten. Newton is onbereikbaar voor de diagnosticus. Hij kan wel schalen gebruiken die aan model-eisen voldoen. De itemschrijvers moeten er vrede mee hebben dat de helft van hun noeste arbeid tevergeefs is. Overigens geen slechte score in de psychologie. Theoretici erkennen hun beperkingen en weten dat ze genoodzaakt zijn te reduceren. Ze krijgen nooit het hele veld in beeld: ze beschikken niet over een *regard survolante*. Het Hegeliaanse *Ganze* halen ze niet en dus ook niet het Hegeliaanse *Wahre*. Ook als men de appels en peren optelt, verkrijgt men een somscore die correleert met criteriumgedrag. Als men items verwijdert die niet aan een model voldoen, stijgt de predictieve validiteit feitelijk gering en soms in het geheel niet.

*Conclusie:* De diagnosticus kent bovengenoemde verschillen en benut zijn vertrouwde instrumentarium. Hij zal zich herkennen in de hindernissen die Smedslund beschrijft. Hij weet dat gedrag historisch en cultuurgebonden is en afhangt van de sociale context, niet solide, maar vluchtig en fragmentarisch is. Dat verhelpen de vele studies die stabiliteit van gedragskenmerken in steekproeven melden geenszins. Het probleem, de vraag van een cliënt is ook een functie van zijn unieke adaptatie (individuele verschillen, persoonlijk profiel), zijn ontwikkeling (historisch gebonden, oscillerend en voorbijgaand gedrag door de tijd heen) en zijn specifieke sociale en fysische context.

## Referenties

1. Millon, Th. (2012). On the history and future study of personality and its disorders. *Annual Review Clinical Psychology*, 8, 1-18.
2. Merleau-Ponty, M. (1945. 1958 translation by Colin Smith). *Phenomenology of Perception*. London and New York: Routledge Classics. Er is een Nederlandse vertaling: *Fenomenologie van de Waarneming*.
3. Lakoff, G. & Núñez, R.E. (2000). *Where mathematics come from? How the embodied mind brings mathematics into being*. New York: Basic Books.
4. Dautzenberg, A.H.J. (2013). *Rafelranden van de moraal, Novelle*. Antwerpen Amsterdam: Atlas Contact.
5. Kuhn, T.S. (1962). *The structure of scientific revolutions*. International Encyclopaedia of Unified Science. Vol. 2, 2. Chicago: The University of Chicago Press.
6. Verlinde, E. (2011). The origin of gravity and the laws of Newton. *Journal of High Energy Physics*, April, 2001:29 Open Access. [http://link.springer.com/article/10.1007/JHEP04\(2011\)029](http://link.springer.com/article/10.1007/JHEP04(2011)029).
7. Mroczek, D.K. & Little, T. (Eds.) (2006). *Handbook of Personality Development*. L. Erlbaum Associates Publishers. Mahwah, New Jersey; London.
8. Gardner, H. (1993). *Creating Minds*. New York: Basic Books.
9. Csikszentmihalyi, M. & Csikszentmihalyi, I.S. (Eds.) (1988). *Optimal experience: Psychological studies of flow in Consciousness*. New York: Cambridge University Press.
- Csikszentmihalyi, M. & Rathunde (1998). The development of the person: an experiential perspective on the ontogenesis of psychological complexity. In: *Handbook of Child Development*. W.W. Damon (Ed.) 5<sup>th</sup> Edition. Vol. 1. New York: J. Wiley and Sons.
10. Seligman, M.E.P. & Csikszentmihalyi, M. (2000). Positive Psychology, An introduction. *American Psychologist*, 55, 1, 5-14.
11. Paulus, voorheen Saulus: In: *Het Nieuwe Testament Onderdeel van de Bijbel: Auteurscollektief, geschreven tussen de 8ste eeuw voor Christus (BCA) en 60 na Christus*. De brieven van Paulus (pp. 2135-2338) zijn geschreven ongeveer 50 jaar na Christus.
12. Kafka, F. (1997). *Het Slot*. Amsterdam: Athenaeum-Polak van Genneep.
13. Barendregt, J.T. (1982). *De Zielenmarkt. Over psychotherapie in alle ernst*. Boom: Meppel/Amsterdam.
- Barendregt, J.T. (1977). *Karakters van en naar Theophrastes*. Deventer: Van Loghum Slaterus.
14. Evers, A. , Lucassen, W., Meijer, R.R. & Sijtsma, K., (2010). *Cotan beoordelingssysteem voor de kwaliteit van tests*. Geheel herziene versie mei, 2009; gewijzigde herdruk mei 2010. Utrecht: Nederlands Instituut van Psychologen.
15. Saint-Mont, U. (2012). What is measurement all about? *Theory and Psychology*, 0(0) DOI: 10.1177/09593354311429997.
16. Jaynes, E.T. (2003). *Probability Theory: The logic of science*. Cambridge, UK: Cambridge University Press.

17. Trendler, G. (2009). Measurement theory, Psychology and the Revolution that cannot happen. *Theory and Psychology*, 15, 5, 579-599.
- Trendler, G. (2013). Measurement in Psychology: A case of *ignoramus et ignorabimus?* A rejoinder. *Theory and Psychology*, DOI: 10.1177/0959354313490451 (pp. 1-25).
18. Markus, K.A. & Borsboom, D. (2011). Reflective measurement models, behavior domains, and common causes. *Doi: 10.1016/j.newideaspsych.2011.02.008*.
19. Kyngdon, A.(2013). Descriptive theories of behavior may allow for the scientific measurement of psychological attributes. *Theory and Psychology*, DOI: 10.1177/0959354312468221 (pp. 1-24).
20. Roskam, E.E.Ch.I. (1982). Hypotheses non fingo. *Nederlands Tijdschrift voor de Psychologie*, 37, 331-359.
21. Guttman, L. (1971). Measurement as a structural theory. *Psychometrika*, 36, 3, 329-347.
- Guttman, L. (1978, August). Recent structural laws of human behavior. *Paper read at the 9<sup>th</sup> International Sociological Congress, Uppsala Sweden*.
22. Stouthard, M.A.E. & Peetsma, Th.T.D. (1999). Future-time perspective: Analysis of a facet-designed questionnaire. *European Journal of Psychological Assessment*, 15, 1, 99-105.
23. Wilson, M. (2013). Seeking balance between the statistical and scientific elements in psychometrics. *Psychometrika*, DOI: 10.1007/s11336-013-9327-3.
24. Van der Maas, H. & Molenaar, P.C.M. (1992). Stage-wise cognitive development: an application of the catastrophe theory. *Psychological Review*, 99, 395-417.
25. Sijtsma, K. (2012). Future of Psychometrics: ask what psychometrists can do for psychology. *Psychometrika*, 77, 1, 4-20.
26. Sijtsma, K. (2011). Psychometrics in psychological research: Role model or partner in science. *Psychometrika*, 71, 3, 451-455.
27. Smedslund, J. (2009). The mismatch between current research methods and the nature of psychological phenomena: What researchers must learn from practitioners. *Theory & Psychology*, 19, 778-794.

## V Hoe kom je vooruit?

*Hoe komt wetenschap vooruit en hoe een wetenschapper? Dat hoeft niet samen te vallen. Slechts enkelen kunnen de wetenschap een stap verder brengen. Je komt bij zulke grote denkers en onderzoekers uit dat je liever je mond houdt. De in 2015 aangetreden KNAW voorzitter spreekt in het NRC over belangrijke en beantwoorbare vragen stellen. De minister van Onderwijs stemt daar mee in. Laat het prestigieuze tijdschrift Nature nu een dergelijk initiatief genomen hebben. Het legt ons een tiental vragen voor dat op antwoord wacht. Wetenschap heeft baanbrekend onderzoek nodig. Dat wordt nu gelezen als de kunst, vaardigheid, het talent om in hoge impact tijdschriften te publiceren. Vier Nederlandse bètahoogleraren doen uit de doeken hoe dat moet. De promotoren, in Duitsland zo charmant Doktorväter genoemd, zijn degenen die onderzoek en kennis vooruit helpen door jonge wetenschappers op het goede pad te zetten. Ervaren publicisten zeggen op een congres wel eens hoe dat in zijn werk kan gaan.*

## 1. Vragen stellen

Als een psycholoog het niet meer weet, zegt hij met een ernstig gelaat: 'Wat denkt u er zelf van?' De *Big Five* is zo tot stand gekomen. Ze zijn het kleinste gemene veelvoud van wat de gemiddelde weldenkende burger zegt dat adjectieven zijn om personen mee te karakteriseren. Dat kan ik ook, dacht de minister van Onderwijs. Ze stelde een commissie onder leiding van Rinnooy Kan en De Graaf in om echte onderzoeksvragen op te sporen. Zo handelend, schakelt zij de burger in om vragen te stellen aan wetenschappers. Dan hoor je wat de mensen willen weten en kun je een lijst met prioriteiten voor de Nederlandse wetenschap samenstellen. Er is een jury samengesteld die de goede vragen er uit haalt. Er zijn meer dan 11.000 vragen ingediend. Je zou naar de kwaliteit van de vragen kunnen kijken. Als de jury dat kan beoordelen, weet ze al wat kwaliteit is. Waarom stelt ze die zelf dan niet? Je zou kunnen denken dat wat het vaakst gevraagd wordt, prioriteit verdient. Een democratisch standpunt, dat wel, maar is het ook een goed criterium? Frequentie is een lastige maat voor de bepaling van de kwaliteit van vragen. Alleen al omdat Nederland vol zit met belangengroepen die er hun stempel op proberen te drukken. Als die pressiegroepen hun vragen niet terugzien, zeggen ze dat de selectie en analyse van de 11.000 vragen niet deugen. Worden zo zinnige vragen gesteld?

De actie lijkt op het gedicht *Impasse* van Nijhoff: de minister verkeert in een impasse. Ze ziet de burgers bezig in hun bedrijf en denkt: Ik vraag het vandaag: waarover wil je dat de wetenschap schrijft? Het gedicht kent twee einden: 'een nieuw bruiloftslied'. Dat zou de minister wel willen, maar het dichterlijk mooier einde is waarschijnlijker: 'ik weet het niet'. We kunnen afwachten welk antwoord de minister krijgt. Komt de wetenschap, bezig in haar bedrijf uit de ivoren toren? Moet je de wetenschapsbeoefenaar eigenlijk niet zijn gang laten gaan? Kan er ook subsidie naar onderzoek dat de gewone burger niet noemt of zelfs niet wil? Waarom is niet gevraagd welk onderzoek gestopt moet worden? Waarom moet het Koningin Wilhelmina Fonds bedelen en enkele enthousiastelingen een onmogelijke berg laten beklimmen? Als dat probleem echt opgelost moet worden dan is er toch de overheid, wij allemaal dus, maar dan vertegenwoordigd door authentieke en belangeloze onderzoekers.

*Tien vragen van het tijdschrift Nature*: Er is eerder geprobeerd vragen voor de wetenschap te formuleren door auteurs die het voor elkaar krijgen om in *Nature* (470, 18-19 (2011)/doi: 10.1038/470018a) te publiceren. Er werd een tiental vragen gesteld waar psychologisch onderzoek antwoord op zou moeten geven. Ik ga ze langs, antwoord en geef commentaar:

(1) *Hoe kan bevorderd worden dat mensen meer op hun gezondheid letten?* Je zou een wederopstanding van de psychofysica kunnen bepleiten om mensen direct te informeren over belangrijke *cues* voor hun gezondheid die ze niet zien en voelen. De informatie betreft gevolgen op de korte, middellange en lange termijn. Dit ligt voor het grijpen gegeven de technologische ontwikkelingen. Zou er weerstand zijn bij dokters: de mensen, patiënten geheten, snappen

dat niet? En aarzeling bij leken: wil/moet ik dat allemaal weten? Gedragsbeïnvloeding, vooral preventie in de gezondheidszorg is niet eenvoudig.

(2) *Hoe ontwerpen we effectieve en betrouwbare instituties, zoals regeringen, overheden en banken?* De reflex bij problemen met instituties is vaak meer toezicht, maar dat werkt niet. Moeten we niet op zoek naar andere en diepere, onderliggende mechanismen, structuren, die overheden en regeringen gebruiken c.q. nodig hebben om hun werk te doen en zich in stand te houden? Moeten ze niet hun angst voor de vrije burger laten varen? Denken ze niet te zwart-wit en zien als enig alternatief anarchie? Zijn ze niet te bang voor controleverlies? Onder-tussen moet het beleid wel consistent blijven, vooral als die burger steeds verschillende wegen in slaat. Meer openheid om de ideeën op te pikken zou niet misstaan. In het moderne China lijkt alles soepel te verlopen, maar alles is gericht op controle op iedereen, en iedere seconde. En, hoe is het in de VS? Niet zoveel anders waarschijnlijk, al lijkt de controle geïnternaliseerd. Het meest vrije én gebonden land, dat je kunt indenken, zei een collega tegen mij die er 20 jaar gewerkt heeft. Een institutie, structuur is leeg en neutraal, heeft een tijd zijn verdienste, maar neigt ertoe de realiteit van alledag te overleven en zichzelf in stand te houden.

(3) *Hoe kan de mensheid zijn collectieve kennis en wijsheid vergroten?* Je denkt hierbij al gauw aan de revolutionaire mogelijkheden van de sociale media. Men kan zich afvragen of de artikelen in de huidige psychologietijdschriften bijdragen aan het verspreiden van die collectieve wijsheid. Is de manier van verspreiden, ontleend aan de exacte wetenschappen, wel geschikt? Zijn artikelen niet geschreven voor een paar collega's? Hoeveel artikelen worden door velen gelezen? Zouden *Open Access* tijdschriften helpen? Of gaat daar een pikorde gelden die niet met het baanbrekende karakter van onderzoek te maken heeft? Deze open tijdschriften vragen geld voor publicatie. Geld stuurt zeker, maar welke kant op, en wie beschikt er over voldoende geld?

(4) *Hoe de vaardigheidskloof te dichten tussen de zwarte en blanke Amerikaanse bevolking?* Dit is een Amerikaans vraagstuk, maar komt in een of andere gedaante in alle landen ter wereld voor. Er zijn en zullen altijd kenmerken gevonden en aangegrepen worden om een 'wij-zij' verdeling te maken. Individuele verschillen zijn een feit, variantie is een feit, maar hoe ermee om te gaan? Is dit een uitdaging voor de wetenschap? Zij kan de variatie nauwkeurig vaststellen, maar ze niet verminderen of opheffen. Sociale experimenten leiden niet tot de ene afdoende aanpak. Is dit niet eerder een politiek en maatschappelijk vraagstuk? Het verminderen van de variantie in capaciteiten en misschien zelfs geluk is voor een deel geslaagd door onderwijs, maar neemt ook weer snel toe als er een crisis is: denk aan de groter wordende *Gini index* in veel landen of aan de grafieken van Piketty<sup>1</sup>. De *Gini index* meet de ongelijkheid tussen waarden van een frequentieverdeling. Het is een maat van de (on)gelijkheid van variaties en wordt gebruikt om de ongelijkheid van inkomens in een bepaald gebied te meten. Een coëfficiënt van nul (0) drukt uit dat één persoon of instelling beschikt over alle inkomsten en een waarde van 1.00 op de percentiel schaal drukt uit dat iedereen hetzelfde inkomen heeft.

Wellicht geldt dit ook voor het verschil tussen de boven- en middenklasse tegenover de onderklasse. Waarom willen of kunnen we die kloof niet dichten? Er zullen nieuwe gemaakt worden, als er een overbrugd is.

(5) *Hoe kunnen we informatie van individuele personen bundelen om de beste beslissingen te nemen?* Dit is goed uitgezocht. Er is veel bekend over informatie-integratie en de meest winstgevendende beslissing. Kahneman<sup>2</sup> heeft duidelijk gemaakt dat er omgevingen zijn waar je niets anders kunt doen als gokken. Dat is niet alleen in het casino, maar ook op de beurs. Er zijn altijd mensen die tijdig waarschuwen voor calamiteiten, maar ze worden zelden gehoord. Hoe dat te verbeteren? Wat gebeurt er met onafhankelijken die de tunnelvisie van direct betrokkenen betwisten? Komen zij er evenals als klokkenluiders slecht van af? Denk aan het verhaal van Arthur Gotlieb<sup>3</sup> die misstanden bij zijn werkgever de Nederlandse Zorgautoriteit (NZa) bekend maakte. Hij werd zodanig getreiterd dat hij zelfmoord pleegde. Waarschijnlijk is de vraag duidelijk, maar het probleem zal nooit opgelost worden, al moet je daar contrafactisch naar streven. Je zou kunnen leren van grote projecten waar het wel goed gegaan is. Dat zouden interessante case studies zijn. Maar als iets goed gegaan is of gaat, verliest de onderzoeker c.q. de subsidieverstrekker zijn belangstelling meestal. Als het goed gaat, kijken we er niet meer naar om.

(6) *Hoe verklaren en begrijpen we de menselijke capaciteit om kennis totstand te brengen?* Dit is een filosofisch-psychologisch vraagstuk. Er zijn al vele aanwijzingen voor antwoorden van Plato, Aristoteles, Descartes, Hume, Kant, Popper, Piaget tot de huidige postmodernisten. De cognitieve psychologie van Piaget<sup>4</sup> gaat over dit onderwerp, maar raakt in vergetelheid. Verdient die opnieuw aandacht? Psychologen hebben de beperkingen van menselijk kennen en oordelen breed uitgemeten door ze te vergelijken met resultaten van rationele modellen. Er is onderzoek naar creativiteit. Er zijn voorbeelden van instituties waar dat gelukt is, van het Eindhovens Natuurkundig Laboratorium tot Silicon Valley. Het zoeken naar een speld in de hooiberg is werkelijkheid geworden in grootscheepse programma's als *Head Start*. Ze werken niet. Toch maar *casestudies* en zoeken naar condities waar mensen in vrijheid iets mogen maken? Of kan het anders: gewoon een opdracht? Denk aan de ontwikkeling van de atoombom in de VS onder leiding van Oppenheimer. Dat was een opdracht. Denk aan de landing op de maan. Dat was een doel van de NASA op aandringen van de president. En, heeft Mr. Jobs van Apple niet zo gewerkt? Hij had precies voor ogen wat hij wilde maken. Een visionair is soms nodig voor nieuwe kennis en kunde.

(7) *Waarom verdienen vrouwen minder dan mannen?* Daar zijn veel en gedetailleerde antwoorden op gegeven. Ze gaan van sociale discriminatie tot biologische verschillen. Waarschijnlijk hebben psychologen, sociologen en antropologen al aardig wat antwoorden klaarliggen. Ze spreken echter niet met één mond. Het antwoord is misschien eenvoudig in het Westen: De meerderheid pakt wat ze pakken kan in een neoliberale-utilitaristische maatschappij en daar zijn mannen vooralsnog beter in dan vrouwen.

(8) *Hoe wordt biologisch sociaal en omgekeerd?* Is dit een ervaring van het *mind-body* vraagstuk? Zijn er al niet aanvaardbare antwoorden gegeven vanuit evolutionair biologisch theoretiseren? Was dat niet Piagets levenswerk? De primatoloog De Waal<sup>5</sup> zou er wel raad mee weten. En, is het wel een goede vraag? Installeert of reanimeert dit niet voor de zoveelste keer een dualisme, een *muddy dichotomy*?

(9) *Hoe kunnen we ons wapenen tegen black swans.* Dit gaat om infrequente gebeurtenissen die we niet zien aankomen maar desastreuze gevolgen hebben. Het is ook op te vatten als het ene tegenvoorbeeld van Popper dat een hele theorie omvergooit. Waren die er maar in de psychologie: ze bestaan niet! De psychologische kant van dit vraagstuk is uitgewerkt, bijvoorbeeld in de signaal-detectie theorie. De menselijke waarneming is niet zo ingericht dat zeldzame en desastreuze verschijnselen het gewicht krijgen dat ze verdienen. Er is en blijft onvoorspelbaarheid, maar dat kan voor een deel teruggedrongen worden door het vergroten van ons waarnemingsvermogen met apparaten. Seismologische apparatuur, satellieten en raketten doen niet anders dan onze ogen diep in de bodem te brengen en aan de rand van het universum. Is die rand een werkelijk verschijnsel, of een grens, gecreëerd door ons vergrote oog. We kunnen ons wel aardig wapenen, maar dat vraagt vaak dure maatregelen en grote investeringen voor zaken die op een *dit* moment *geen gevaar* vormen. Langetermijndenken is lastig en je bent afhankelijk van modellen om dat denken af of uit te beelden. Die kloppen nooit helemaal; altijd onzekere parameters. Hoe lastig is het om de klimaatverandering serieus te nemen. Bovendien moet je daarvoor soms tegen belangen (de lobbyisten van de oliemaatschappijen), maar ook van wetenschappers, ideologieën en religies in gaan. Dat zijn lastige opposanten. Er zit voorlopig geen antwoord in dat hout snijdt. Onze eigen aard zit ons in de weg.

(10) *Waarom blijven sommige gedragingen en processen zo lang in stand, ook al willen we er van af?* De auteurs in *Nature* bedoelen onder meer geweld. Amerikaanse militairen slaan er over de hele wereld op als er een conflict is: 'Eerst slaan dan waarschuwen' luidt, als eerder gezegd, Bagheera's Wet van de Jungle. Misschien is duidelijker wat we *wel* willen (succes, de baas zijn) en wat we *niet* willen (armoede, lijden en waar we bang voor zijn: geweld, misbruik). Dit is het 'wat' van de diagnostiek: de vragen, waar de cliënt mee komt, en die onderzoek uitlokken. Psychologen kunnen dit aardig beschrijven en verklaren, maar dat betekent niet dat ze het kunnen voorkomen. Spreekt hieruit niet de veronderstelling van de maakbaarheid van de samenleving volgens egalitaire principes? Dat is tot nu nergens gelukt, denk aan de sociale experimenten in de voorbije Sovjet Unie en het huidige Noord-Korea.

*Conclusie:* De tien vragen zijn niet origineel en liggen voor een deel buiten de wetenschap. Ze reflecteren waar het weldenkend publiek en beschaafde regeringen zich druk over maken. En, sommigen in die samenleving en in enkele regeringen hebben daar geen boodschap aan. De vragen dagen niet uit om nieuwe gedachten over wat gedrag is te formuleren en gaan niet over waarom we ons gedragen, zoals we doen. Ze gaan over (on)gewenste gevolgen van ons gedrag; niet over *het gedrag zelf*. Nu maar afwachten wat de oproep aan alle burgers oplevert?

Is een goede vraag een speld in de hooiberg, de kleinste gemene deler, een subsidiekanon? Of nog iets anders? Bijvoorbeeld, een poging de burger te laten denken dat hij belangrijk is, of een actie om het imago van wetenschappers op te poetsen na de fraude van hoogleraren aan de Universiteit Tilburg en Vrije Universiteit Amsterdam?

*Voorspelling:* Wat zullen de concrete 11.000 vragen opleveren voor de Nederlandse wetenschapsagenda? Zullen er nieuwe onderzoeksdomeinen worden aangeboord? Dat is onwaarschijnlijk. Zullen er problemen door opgelost worden? Waarschijnlijk niet. Er zal nog steeds werkloosheid zijn, het vluchtelingenprobleem zal niet in goede banen geleid worden, het gaat op zijn best 'over' en er zullen Nederlanders zijn en blijven die zich hier niet thuis voelen. Wat er waarschijnlijk zal gebeuren is, dat groepen die al sterk staan in onderzoek nog meer claimen en krijgen dan ze al deden en hebben. Die groepen vormen immers een afvaardiging die de jury gaat vormen en die beslist over welke vragen beantwoord worden.

## 2. Baanbrekend onderzoek

Weinig wetenschappers publiceren in tijdschriften als *Nature* en *Science*. Het tijdschrift *Public Library of Science* (PloS) One van 9 juli 2014 vermeldt onderzoek naar publicaties van uitgever Elsevier - een 'grote speler' op het gebied van wetenschappelijke tijdschriften - en naar voordrachten tussen 1996 en 2011. Wat laat dat onderzoek zien?

*Eén procent maakt de dienst uit:* Er wordt geschat dat er wereldwijd ruim 15 miljoen onderzoekers zijn. Slechts 1% daarvan (ongeveer 150.000) slaagt er in jaarlijks (van 1996 tot 2011) te publiceren. Dit wordt aangeduid met UCP: *Uninterrupted, Continuous Presence*. Het betekent dat een aantal teams, departementen, discipline groepen en hele universiteiten geen of slechts een enkele auteur heeft die tot deze 1% behoort. Deze één procent maakt de dienst uit over wát onderzocht wordt en of het gepubliceerd wordt. Telt daar niet de *context of reputation*? Thema's wisselen. De tijdgeest heeft invloed op wat kans maakt gesubsidieerd te worden. Er werden 3.190 artikelen geteld die meer dan 1000 keer geciteerd waren. Ze bevatten 87% dezelfde namen, vooral van leiders van grote onderzoeksgroepen in de VS en wat minder Europa. De artikelen betroffen fysica, scheikunde, *engineering*, geologie, biologie, biotechnologie, infectieziekten, medisch- en hersenonderzoek, gezondheids- en computerwetenschappen. Dit lijkt op zangwedstrijden met begeleiding op de lier die de Romeinse keizer Nero uitschreef. Zijn moeder, de beruchte Agrippina, verbood hem filosofie te studeren, dat zou hem in de weg staan om heerser te worden. Zingen mocht wel van haar. Een zingende dictator zag ze wel zitten, een kritische filosoof minder. Nero ontpopte zich als *sing a song writer* avant la lettre. Hij nam zelf steeds de eerste prijs in ontvangst. De jury was uiteraard kritisch maar kwam toch tot de slotsom dat Nero de grootste zanger/liedjesschrijver was. Historici dichtten hem overigens wel enig muzikaal talent toe.

Of *Open Access* tijdschriften dit monopolie doorbreken is de vraag. De auteurs moeten betalen om hun publicatie gedrukt te krijgen en dat is voor PLoS nu tussen de 1350 en 2900 euro. En, wie betaalt, bepaalt. Kennis is niet alleen te koop door subsidies van bedrijven en overheid, ze moet ook verkocht worden. Het wetenschappelijk *breaking news* wordt gemaakt door een paar teams van exacte wetenschappers bij enkele universiteiten en een paar grote onderzoekinstellingen.

*Suggesties:* Vier Nederlandse bètawetenschappers deden in het NRC van 20 april, 2013 suggesties. Ik pas ze toe op de psychologie als hofleverancier van kennis en methoden voor diagnostiek:

Ten eerste moet onderzoek origineel zijn, je verricht het bijvoorbeeld op plekken waar iemand anders nooit geweest is: Noord- of Zuidpool, de binnenlanden van Afrika en Zuid-Amerika. Of, pak een onderwerp waar een ander nog niet aan gedacht heeft. De primatenonderzoeker De Waal zei dat hij coöperatief gedrag bij apen is gaan onderzoeken, omdat deze dieren meestal als agressieve overlevens worden beschouwd. Hij zei dat hij dit deed om eruit te springen. Dat is hem gelukt. Wellicht maakt zijn vermogen het gedrag van primaten te observeren meer indruk dan zijn interpretaties. De titel van zijn laatste boek luidt: *Bonobo en de tien geboden*.

Observatie en interpretatie probeer je uit elkaar te houden, maar dat is lastig. De vragen mogen diffuus en onduidelijk, niet Cartesiaans klaar en distinct zijn, zeggen onze bèta's. Dit is in strijd met nulhypothese-toetsing, logisch positivisme en de publicatiegewoonten in de psychologie.

Ten tweede moeten onderzoekers het tijdschrift en de *reviewers* kennen. Vaak is een auteur zelf ook *reviewer*. Samenwerken met Amerikaanse auteurs vergroot de kans van publicatie. Soms kun je *vooraf* vragen of het tijdschrift geïnteresseerd is in je verhaal. Het bezoek van congressen kan de kans vergroten op publicatie. Dit zijn geen intrinsiek wetenschappelijke, maar wetenschapssociologische regels.

Ten derde moet het artikel helder geschreven zijn. De titel is belangrijk: niet te aarzelend, niet te brutaal en zonder jargon. De referentiebrieven (de brieven, die met het artikel meegestuurd worden) moeten precies zijn. De Groot had moeite met deze eis van helderheid, want op grond van dit criterium zouden veel filosofische geschriften nooit gepubliceerd zijn, van Plato, Selz, tot Heidegger en Popper. De Groot<sup>6</sup> was bekend met het werk van Popper en de filosoof/denkpsycholoog Otto Selz. Beiden spelen een rol in zijn dissertatie: *Het denken van de schaker* (1946) en in zijn boek: *Methodologie* (1961). Filosofen werken een grondintuïtie uit die niet rationeel en logisch te verantwoorden is. Als het lukt, brengt die intuïtie een nieuw facet van denken/gedrag of een verrassend perspectief op de mens, zijn gedrag en de samenleving aan het licht. Deze zijn niet van meet af aan 'klaar en distinct'. Juist betrekkelijk eenvoudige intuïties zijn vruchtbaar geweest voor onderzoek en inzicht, denk aan de evolutieleer, de genetica (dominante en recessieve genen) de zwaartekracht, Freuds driedeling van de psyche in Es, Ego en Superego en het gesprek als therapie.

Ten vierde wordt aanbevolen open te staan voor onverwachte resultaten en te accepteren dat projecten niks opleveren. Veel hypothesen blijken onjuist. Deze bewering staat in contrast met wat er in psychologische tijdschriften gebeurt. Fanelli<sup>7</sup> liet zien dat rond de 90% van de onderzoekshypothesen aanvaard worden. Zou dit te maken hebben met selectief publiceren, met vermijden van riskant onderzoek, dat wil zeggen onderzoek met onzekere uitkomsten? De logica van het lineair geschreven verhaal correspondeert niet met het werkelijk verloop van onderzoek. Dat is immers niet logisch, serieel, maar grillig: een Echternach processie.

Ten vijfde wordt aanbevolen om slim met *reviewers* om te gaan en een experiment al vast te doen, waar zij om *zullen gaan* vragen. De auteurs moeten accepteren dat er vertraging is. De eerste aanbieding van een artikel en de publicatie kunnen soms meer dan een jaar uit elkaar liggen. Ook al is de verbetering klein, als er om gevraagd wordt door reviewers, moet je het doen. Als je er tegenin gaat, doe het voorzichtig en diplomatiek.

Ten zesde moet je er voor zorgen zichtbaar te zijn. Tijdschriftredacteuren zien je gemakkelijk over het hoofd en *reviewers* zitten niet op concurrenten te wachten. Je moet al gauw met enig

*breaking news* komen om je in de kijker te spelen. Dat is ook riskant: denk aan de eindhovense chemicus die het middel tegen AIDS dacht gevonden te hebben.

*Wetenschap moet je aan de man brengen:* Volgens onze vier bètahoogleraren is er naast inhoud en kwaliteit van een publicatie sprake van sociale factoren en communicatieve vaardigheden. Inhoud en communicatie, vormen een morsige tweedeling en lopen door elkaar. Dit leidt tot de vraag wat de missie van wetenschap, i.c. psychologie is. Moet die missie enkele experts boeien, maatschappelijke problemen oplossen? Moet ze de markt op: valorisatie? Moet ze kijken waar ze met haar missie van beschrijven, verklaren/ controleren en beslissen staat en af en toe de balans opmaken?

Er zijn in de psychologie geen perfecte samenhangen en geen substantiële hoeveelheden verklaarde variantie. Waarom worden veel gedragingen zo beperkt verklaard? Is een deel onverklaarbaar? En, is dat een uitvloeisel van het menselijk kennen? Is er een echte limiet, of staat de psychologie nog in de kinderschoenen?

*Een formule:* De vier hadden een formule kunnen voorstellen. Wat is de aantrekkingskracht/ zwaartekracht van een publicatie, onderwerp? Waarom niet als denkexperiment en provocatie de omgekeerde kwadraatwet van Newton gebruikt? Als de afstand tussen jouw publicatie en die van een Harvard-, M.I.T., Berkeley-medewerker verdubbelt, wordt de aantrekkingskracht van jouw publicatie vier keer zo zwak. Hoe verder verwijderd van een Harvard-, Berkeley-, Princeton- medeauteur, hoe geringer de kans op publicatie. Binnen de VS is de kans het grootst en bij het groter worden van de afstand neemt de kans kwadratisch af. Dat klopt waarschijnlijk niet. Je kunt dan proberen empirisch een predictieformule op te stellen met *cues* die, opgeteld, voorspellen hoe groot de kans is dat je artikel/onderwerp/onderzoeksaanvraag gepubliceerd wordt. Ik ben benieuwd aan welke *cues* onze bèta's denken. Psychologen willen wel langs empirische weg een predictieformule maken. Daar zijn ze goed in.

De bèta's stellen moderators van de hier verzonden kwadratische relatie voor: betrek een Amerikaan van Harvard, Princeton, Berkeley, enzovoort bij je publicatie, neem Amerikanen in dienst en hoop dat ze even blijven - in de VS hebben ze volgens mijn redenering meer kans - en schrijf helder volgens een voor die tak van wetenschap geldend *Manual*. Probeer de *reviewers* voor te zijn, doe iets afwijkends en zorg voor bijzondere resultaten maar wel zodanig, dat ze een dominante groep niet tegen de haren instrijken.

*Creativiteit:* Komen nieuwe inzichten, baanbrekend onderzoek voort uit creativiteit en/of uit lang en hard werken? Roskes et al.<sup>8</sup> bestrijden deze tweedeling. Ze stellen een *dual pathway* voor: intuïtief, flexibel en moeiteloos naast volhardend, rationeel, stap voor stap. Beide kunnen tot een creatief, nieuw resultaat leiden. Ze concluderen dat rationele vermijdingsgemotiveerde deelnemers evengoed tot een creatieve oplossing kunnen komen als creatieve, associatieve op hun doel afgaande deelnemers. De strikte scheiding tussen willekeurig, rationeel en intuïtief is niet houdbaar. Maar creativiteit en hard werken zijn niet genoeg. Er wordt door onze bètawetenschappers immers ook op gewezen, dat je moet weten hoe het wetenschapsbedrijf werkt.

Er staat geen aanduiding van interessante thema's of van de *context of discovery* in. Een antropologische studie van het wetenschapsbedrijf, zoals Luyendijk<sup>9</sup> heeft gedaan met de banken in Londen, zou hier niet misstaan. Ook nog een mooie predictieformule erbij en je weet weer iets meer van de predictoren van de productie van het gemiddelde wetenschapsbedrijf. De aanvulling van cultureel antropologische *case studies* is er al, maar er zijn andere aanvullingen te bedenken. Welke *cases* zouden interessant zijn om onder de loep te nemen? Neem, om te beginnen succesvolle én mislukte, bijvoorbeeld het CERN in Genève, de chipbouwer ASML uit Veldhoven, De Spijker van Muller, enzovoort.

### 3. Zo kom je vooruit

Het belangrijkste product in de wetenschap is momenteel een *artikel* in een internationaal tijdschrift. Men gaat er van uit dat daar de nieuwe inzichten en baanbrekende onderzoeksresultaten verschijnen. Wat daarbuiten geschreven wordt telt minder, al zijn er mooie en lezenswaardige essays buiten zulke tijdschriften verschenen. De metafysicus Peters<sup>10</sup> schreef in kerken streekblaadjes boeiende en leesbare verhalen over de persoon, de jeugd, de ouderdom, vrijheid en de relatie tussen poëzie, kunst en wetenschap, enzovoort. Ze zouden een aanvulling zijn in een psychologieboek.

Het gaat om tijdschriften met hoge impact. Hoe kom je daar in? De Amerikaanse sociaalpsycholoog Kruglanski onderwijst AIOs, PhD studenten en universitaire medewerkers artikelen zo op te schrijven dat ze gepubliceerd worden in Engelstalige tijdschriften. Artikelen bevatten vaak onderzoek van PhD studenten. Zo halen ze hun graad binnen. Kruglanski<sup>11</sup> is een praktisch man: '*... don't think big, just do your five experiments, clip them together, and hand them in.*' De APA geeft een *Publication Manual* uit, waarin in grote lijnen wordt uitgelegd hoe een artikel opgeschreven moet worden.

*Wat gepubliceerd wordt is ook niet alles:* Kruglanski en Higgins merken tegelijkertijd op dat sociale psychologie en persoonsleer meer *data-* en minder *theory-driven* zijn geworden. Ze beweren dat deze stand van zaken de vooruitgang tegenhoudt en verhindert om bruggen te slaan tussen verwante wetenschappen. Ze wijzen als oorzaak ondermeer het BAMA onderwijs aan. Dat leert studenten methoden hanteren en data analyseren, maar schiet tekort om interesse te wekken voor theorieontwikkeling. De *publish or perish attitude* zou eenvoudige, snel te begrijpen artikelen uitlokken. En, de medewerkers van universiteiten halen de vereiste hoeveelheid artikelen slechts als hun PhD studenten ze uitvoeren. Tillie<sup>12</sup>, een Amsterdamse hoogleraar Electorale Politiek, zei provocerend, dat wetenschappelijke kwaliteit het publiceren in een Engelstalig tijdschrift is dat bijna niemand leest. Het gemiddeld aantal lezers per artikel zou volgens hem drie zijn. Dat lijkt me overdreven. Het zullen geen massa's zijn, maar er zijn altijd experts en geïnteresseerden. Bij sommige onderwerpen is dat inderdaad weinig. Volgens hem bieden de meeste artikelen geen inzicht, maar informatie. Het verschil tussen journalist en onderzoeker valt zo weg.

*Cursussen om te leren publiceren:* Kruglanski geeft zijn cursus ook in Nederland, onder meer voor het *Utrecht Centre for Child and Adolescent Studies*<sup>13</sup>. Dit centrum biedt een PhD training aan met *Specialized courses offered by CAS staff*: # 1 -Publishing in Social Science en # 6 -Tacit Academic Knowledge: Hidden rules for Academic Success in Times of the Replicability Crisis. Daarin worden antwoorden gegeven op vragen, zoals: *How do I write up my research findings in a way that interests and convinces others. How can I decide on the best potential outlets for my manuscripts. How do I wisely create opportunities for myself on the job market.* In zes twee-uur sessies gaat het om hoe onderzoek gepubliceerd te krijgen, subsidies binnen te halen, productie te verhogen, om te gaan met afwijzing en hoe de indruk van *sloppy science*

te vermijden. De laatste samenstelling is afkomstig van Levelt, de voorzitter van de commissie, gevormd naar aanleiding van de fraude van Stapel. De opvatting van Kruglanski en aanhangers is die van de sofisten in Plato's tijd. Zij hadden niet veel op met *Theorein* en het *Soidzein* van de verschijnselen. Het ging hen om invloed, reputatie, aanzien, overtuigen en brood op de plank.

*Wat zeggen ervaren publicisten?* Gigerenzer<sup>14</sup> vertelde over zijn ervaring op een congresdiner. Hij zat met drie collega's en vier studenten aan een tafel met Chinees eten. De studenten wilden uit de eerste hand horen hoe hun PhD thesis af te ronden en daarna goede onderzoekers te worden. De heren - het zijn meestal mannen - waren bereid advies te geven. Met autoriteit zei een (zie hierboven): 'Doe je vijf experimenten, sla er een nietje door en leg het op het bureau van de promotor. De PhD studenten knikten tevreden.

Gigerenzer gooidde roet in het Chinese eten: *Don't follow this advice unless you are mediocre or unimaginative; try to think in a deep, bold and precise way; take risks and be courageous.* De PhD studenten bleven knikken. Ze hadden als de beroemde rabbi kunnen zeggen bij het vaststellen van een tegenstrijdigheid: U hebt beiden gelijk en u hebt ook gelijk, dat het tegenstrijdig is. Zo voorkom je de klap die soldaat Tersides kreeg toen hij voorstelde aan koning Odysseus om de belegering van Troje te stoppen na tien jaar uitzichtloos voor de muren gelegen te hebben.

Gigerenzer wees vervolgens op analyses die lieten zien dat de helft van de studies in twee bekende sociaalpsychologische tijdschriften bijna tautologisch waren. Denk aan Smedslunds<sup>15</sup> opmerkingen over pseudo-empirisch, tautologisch, semantisch onderzoek. Hij vertelde dat hij enkele artikelen in het *Journal of Experimental Psychology* van de jaren 20 tot 30 van de vorige eeuw had gelezen. Hij zei dat er zijns inziens veel verloren is gegaan: verschillende (statistische) methoden, precieze rapportage van individuele gevallen, zorgvuldige selectie van proefpersonen, nagaan of de sekse van de proefpersoon of proefleider van invloed was op het resultaat van de studie en aandacht voor *unobtrusive* metingen. Hij zag nu een dominantie van het verzamelen van data en ze publiceren in de *smallest publishable units without substantive theory*. Een citaat: '*...data without theory are like babies without parents: their life expectancy is low*' (p. 296). Maar, hij is dubbelhartig (Goethes *zwei Seelen in meiner Brust*). Op het Max Planck instituut in Berlijn wordt een publicatie in een Amerikaans hoog impact tijdschrift steeds gevierd.

*Volg je droom:* Een toespraak van Mr. Steve Jobs van Apple voor Stanford studenten (2005), zes jaar voor zijn dood staat dicht bij Gigerenzer (zijn ene ziel van de twee in zijn borst) dan bij Kruglanski. Ik heb dit op de TV gezien en aantekeningen gemaakt van de (ondertitelde) toespraak van deze frêle man. U kijkt het maar terug. Het filmpje staat vast op *You tube*. Jobs vertelde uit eigen ervaring. Het was geen lezing over Apple. Het is best mogelijk dat de toespraak voor hem geschreven is, maar zijn boodschap is duidelijk. Hieronder een parafraze van zijn toespraak. Zijn eerste ervaring was zijn adoptie. Zijn academische biologische moeder wilde een gezin voor hem vinden dat uit universitair opgeleiden bestond. Zijn adoptieouders waren geen

academici. Men kan dit zo lezen: aanleg en omgeving zijn van belang, maar geen is doorslaggevend. De tweede ervaring had te maken met zijn opleiding. Hij behaalde geen graad want de studie verveelde hem. Hij voelde zich schuldig over het geld dat hij verspilde. Hij stopte niettemin met de studie en deed wat hij interessant vond: kalligrafie. De Apple is gekenmerkt door een mooie vormgeving. Zijn derde ervaring: hij werd door zijn eigen bedrijf op straat gezet toen het niet goed ging. Hij begon een nieuw bedrijf dat zo vernieuwend was dat Apple het kocht. Hij keerde terug naar huis. Men kan het lezen als: maak dingen die je zelf mooi en de moeite waard vindt, let niet alleen op verkoopcijfers, aandeelhouderswinst. Laat je niet uit het veld slaan door anderen die je dromen onzin vinden, bewandel zijpaden en serendipiteit zal je overkomen en vergezellen. De vierde ervaring was de mededeling dat hij kanker had. Na de boodschap dat hij niet lang meer zou leven, kon hij geopereerd worden en kreeg weer perspectief. Hij toonde letterlijk het *Wunderkind* verhaal (Thompson<sup>16</sup>) ook wel het *Golden Boy Narrative* genoemd, terwijl een tragisch verhaal meer voor de hand lag. Hij eindigde met ... *follow your heart, be foolish, never give up, create beautiful things: Stay hungry, stay foolish!*

*Er tegen aan, er tegen in gaan:* In James' analyse<sup>17</sup> van de religieuze ervaring is een tegenstelling beschreven tussen de eerstgeborenen: de gezonden van geest en de tweemaal geboren: de zieken van ziel. De eersten hebben het gevoel dat het goed gaat met de wereld en dat ze aan de rechterhand van God zitten. Iets dergelijks propageert de positieve psychologie. De twee maal geboren zien kwaad, pijn en verlies in de wereld. Ze zijn melancholisch - volgens Aristoteles is dat een vruchtbare gemoedstoestand voor onderzoekers - met als dieptepunt dat ze het gevoel verliezen voor wat verloren is gegaan en het kwaad ontdekken. Deze tegenstelling bestaat ook in de Islam, waar de Soefi beweging de zwaarmoedige, bevindelijke kant benadrukt, terwijl de shar'ia het wettische en zelfingenomene benadrukt. De Griekse goden Apollo en Dionysos gaan over een georganiseerde versus de anarchistische wereld. In de Chinese cultuur gaat het over het Confucianisme versus het Taoïsme, waarbij de huidige regering terugvalt op Confucius, nu het Maoïsme verlaten is. Onder Mao waren beide verboden.

Welke houding is vruchtbaar, aan welke lessen heeft de geïnteresseerde student het meest? Die van Kruglanski, die van de gezonden van geest (James), de shar'ia, het Apollinische, het Confucianisme? Of de andere pool? Hebben de gezonden van geest niet greep op de besturen, de politiek, het bedrijfsleven, de banken en de media? In de Nederlandse media is vrijwel iedere spreker vertegenwoordiger van een belangengroep. Hij vertelt je waarschijnlijk niet wat interessant of goed is voor jou, maar voor hem en zijn groep. De nadruk op de *context of justification* van de wetenschap is er gelukkig ook nog; niet alleen de *context of selling*. Nu nog hernieuwde aandacht voor de *context of discovery* (Reichenbach<sup>18</sup>) en ontdekken welke kanten van gedrag nog niet of te weinig beschreven zijn.

## Referenties

1. Piketty's boek is vertaald in het Nederlands. Ik verwijs hier naar de grafieken die in 'Waarom Piketty Lezen? 49 reflecties op 'Kapitaal in de 21<sup>ste</sup> eeuw' beschreven zijn (2014) door R. Went (redactie). Amsterdam: Amsterdam University Press B.V. Hoofdstuk 12 van John Cassidy vermeldt zes leesbare grafieken en Paul Krugman zet het in een grafiek neer (hoofdstuk 22, p. 153).
2. Kahneman, D. (2011). *Ons feilbare denken*. Amsterdam: Business Contact.
3. Gotlieb, A. (2014). *Operatie 'werk Arthur de deur uit'*. Onder redactie van Joep Dohmen & Jeroen Wester. Amsterdam: Bertram+deLeeuw Uitgevers.
4. Piaget, J. (1970). Piaget's Theory. In: *Carmichael's Manual of Child Psychology*, 3<sup>rd</sup> Vol. (pp. 703-733). New York: John Wiley & Sons Inc.
5. De Waal, F. (2015). *Bonobo en de tien geboden*. Amsterdam: Uitgeverij Atlas Contact.
6. De Groot, A.D. (1961). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen* (12<sup>de</sup> druk, 1994). Den Haag: Mouton.  
De Groot, A.D. (1946). *Het denken van de schaker*. Den Haag: Mouton.
7. Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States data. *PLoS (Public Library of Science) One*, 5, 4, e 10271. DOI: 10.1371/journal.pone.0010271.
8. Fanelli, D. (2010). 'Positive' results increase down the hierarchy of the sciences. *PLoS (Public Library of Science) One*, 5, 3, e10068. DOI: 10.1371/journal.pone.0010068.
9. Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics*. DOI: 10.1007/s11192-011-0494-7. (published online: 11 September, 2011).
10. Roskes, M., De Dreu, C.K.W. & Nijstad, B.A. (2012). Necessity is the mother of invention: Avoidance motivation stimulates creativity through cognitive effort. *Journal of Personality and Social Psychology*, 103, 2, 242-256.
11. Luyendijk, J. (2014) *Dit kan niet waar zijn. Onder bankiers*. Amsterdam: Atlas Contact.
12. Peters, J.A.J. (1984). *De Vraag naar het Zijn*. Keur uit zijn verspreide geschriften bijeen gebracht en ingeleid door H.H. Berger, C.E.M. Struyker Boudier en Th. van Veldhoven. Kampen: Kok Uitgeverij.
13. Kruglanski, A.W. & Higgins, E.T. (2004). Theory construction in social personality psychology: Personal experiences and lessons learned. *Personality and Social Psychology Review*, 8, 2, 96-97.
14. Tillie, J. *Wetenschap* in: NRC van zaterdag 5 april: Opinie en Debat (p. 5).
15. Programma van het Utrecht Center for Child and Adolescent Studies, z.j. verzonden November, 2014.
16. Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford: Oxford University Press.
17. Gigerenzer, G. (2011). Personal reflections on theory and psychology. *Theory and Psychology*, 20, 6, 733-743.

15. Thompson, M. (1982). A three dimensional model. In M. Douglas (Ed.). *Essays in the sociology of perception* (pp. 31-63). London: Routledge & Kegan Paul Ltd.
16. James, W. (1902, 1982). *Varieties of religious experience*. Hammondsworth. In 1982 printed as Penguin Book. Nederlandse vertaling beschikbaar (2010).
17. Reichenbach, H. (1938). *Experience and Prediction: An analysis of the foundation and structure of knowledge*. Chicago Illinois.: The University of Chicago Press.

## VI Slot

### Terugdenkend aan Diagnostiek

*Terugblikkend constateer, beweer, vind en geloof ik het een en ander. Onder meer dat bedeksels niet objectief, neutraal of omvattend zijn. Er is altijd een ander gezichtspunt mogelijk. Ons perspectief verschuift steeds. Elk doel is multipel realiseerbaar. Precies daarom moet je bedeksels verdedigen; met de hakken in het zand om je zelf en anderen te overtuigen. Anders rest vergetelheid als overigens rustig en plezierig vooruitzicht, en zoals Nijhoff het zegt: 'zingend en zonder herinnering het volgende (derde) land ingaan'.*

*Ik ga in op de positie van de diagnostiek in de psychologie en de wetenschap, kies er een en licht die toe. Het is een persoonlijke overdenking. 'Nog een keer door die hoge gang' is de titel van 100 schoolgedichten (H.Werkman, samensteller; 2001, Amsterdam: Ooievaar). Nog één keer wat ik zeggen wil over diagnostiek: ze gaat over het gedrag van een individu, de cliënt en dat is niet waar psychologisch onderzoek hoofdzakelijk over gaat. Diagnostiek is geen pilaarheilige die het in haar eentje redt. Ze moet een positie in de (sociale) wetenschap vinden. Ze heeft haar kennisbestand nodig, al is dat overmatig en onoverzichtelijk. Diagnostiek moet bovendien methoden die geldigheid van die kennis borgen op waarde schatten voor de vraag, het probleem van de cliënt. Met dit alles moet de diagnosticus dealen. Hoe houdt hij zich staande? Ten slotte zeg ik iets over lezen en schrijven.*

*Diagnostiek is geen onderzoek bij steekproeven in de logisch positivistische, empirisch-analytische zin van het woord. Ze gaat over gedrag van één cliënt.*

Psychologische diagnostiek is geen biologisch, medisch, natuurwetenschappelijk of sociaal wetenschappelijk onderzoek met steekproeven. Het is eerder zoek naar het waarom en het hoe van het gedrag, het doel dat een cliënt nastreeft, gedrag dat hij prefereert of dat hem dwarszit. Van chemie, biologie, neurologie, wiskunde en van nog veel meer heeft een diagnosticus geen verstand. Dat onderzoek kan hem slechts metaforen bieden om gedrag te beschrijven. Gedrag is de ziel van de psychologie. Op de ziel geeft de psychologie soms af. Ze definieert haar in ieder geval gemakzuchtig. Wijnberg<sup>1</sup> is dichter, schrijver en hoogleraar Cultureel Ondernemerschap. Een bijzondere combinatie, gedoopt door hemzelf en de UVA. Na een verhaal over de geldprofessor citeert hij Emily Dickinson, een zichzelf in de weg zittende diepgraafster. Ze heeft ongeveer 1800 gedichten geschreven die na haar dood in een lade gevonden werden: *I gave Myself to Him/And took Himself for Pay/ The solemn contract of a Life Was ratified, this way*. Wijnberg vult aan: 'Maar afgezien van Emily Dickinson en twee of drie anderen is de economie van de ziel een verwaarloosd veld, ook omdat de ziel door de meeste onderzoekers zo armzalig gedefinieerd wordt, als ze die al definiëren'.

Diagnostiek sluit geen gedrag uit en zit niet vast aan een specifieke vorm. Een dergelijke onbepaaldheid lokt het aanbrengen van orde uit. De wetenschapper houdt het in eerste instantie bij een top-down-bureaustoel-aanpak en maakt een abstracte formele of semantisch conceptuele theorie. Een praktisch georiënteerde psycholoog kijkt om zich heen en kiest een bottom-up-aanpak en beschrijft gedrag op allerlei manieren. Wetenschappers hopen dat het bestuderen van deeltjes en het stapelen van feiten tot inzicht leiden. Het zijn alle geoorloofde en zinvolle manieren om er iets van te maken: *poiein*.

Om zicht op gedrag als 'bijna alles' te krijgen zijn drie niet-willekeurige bronnen van theorievorming onderscheiden: impliciete alledaagse theorieën, expliciete theorievorming uit tekstboeken (de heersende paradigma's van Kuhn<sup>2</sup>) en alternatieven voor expliciete theorievorming. Deze kun je ordenen in drie oriëntaties: Individuele Verschillen, Ontwikkeling en Fysische & Sociale Context. Deze bevatten de theorieën uit onze tekstboeken. Combinaties van telkens twee bronnen leveren de bekende tweedelingen van Persoon-Situatie, Organismisch-Mechanistisch en Aanleg-Omgeving op. Deze ordening verlost je van het opsommen van al die theorieën.

Simon<sup>3</sup> inspireert om contrafactisch te streven naar een rijke en voldoende beschrijving van de cliënt. Ik ga niet mee in zijn kijk op intelligentie als het manipuleren van abstracte symbolen. Hij zet, zo doende en denkende, de relatie met de biologische *wetware* en psychologische *software* tussen haakjes. Hij denkt die niet nodig te hebben. Hij bedoelt de symbolen en structuren niet metaforisch, maar letterlijk. Dit mag dan populair zijn, maar de diagnosticus heeft niet veel aan zulke abstracties.

In de diagnostiek is de aandacht eenzijdig gericht op individuele verschillen. Ontwikkeling en Sociale context zijn niet te ontkennen gedragskenmerken. Elk gedrag heeft een verloop en is gesitueerd. Een (huis)arts kijkt, als het goed is, naar de context: het gezin, de buurt, de culturele achtergrond. En, hij (m/v) laat je een paar keer terugkomen om te kijken hoe je vraag / probleem zich ontwikkelt. De dokter heeft het beter voor elkaar dan de diagnosticus.

Diagnostiek gaat over gedrag. Gedrag is onderwerp van elke wetenschap, want zonder menselijk gedrag is er geen wetenschap. De mens en zijn gedrag zijn bijgevolg intrinsiek een *go-between* tussen alle wetenschappen. Het risico van dit uitgangspunt is dat je buiten je boekje gaat en uitspraken doet over zaken waar je geen verstand van hebt. Moet je maar voor lief nemen. Diagnostiek richt zich op psychologische mechanismen en constructen/ begrippen/ hypothetische constructen die het gedrag van de cliënt beschrijven, verklaren en voorspellen. Daarbij zijn formele modellen welkom, maar hebben geen prioriteit. Eerst de empirie en de daarin vermoede orde, structuur of het ontbreken daar van.

Op de vlakte blijven is zo gek nog niet voor diagnostiek. Deftiger: je kunt een *horizontal world view* verdedigen (Van Dijk & Wilthagen<sup>4</sup>). Deze steunt op Wittgensteins en James's werk waarin het hogerop zoeken door middel van reductie, het zoeken naar de kleinste eenheid en naar verborgen onder- en bovenliggende lagen secundair is aan de concrete beschrijving met behulp van voorbeelden. Wittgensteins<sup>5</sup> opmerking (1953, sec. 2 xiv): *'The existence of experimental method makes us think we have the means of solving problems which trouble us; though problem and method pass one another by'* heeft niets van zijn kracht verloren.

Diagnostiek heeft robuuste bevindingen nodig om het gedrag van de cliënt te kunnen voorspellen en controleren. Meta-studies beloven zulke bevindingen te bieden. Ze laten echter bescheiden tot gemiddelde d- en r- waarden met grote variabiliteit zien. Voorbeelden: r capaciteiten - werkprestaties tussen .15 en .65, predictieve validiteit van tests: r tussen de .02 en .75 en seksverschillen in schoolprestatie d-waarden tussen -0,30 tot 0,17. Dat duidt op steekproef- en mogelijk ook lokaal gebonden zijn van resultaten. Er zijn geen universele gedragswetten.

Diagnostiseren vraagt een open, eclectische houding. Ze is een activiteit *sui generis*. Daar bedoel ik mee, dat een diagnosticus zijn werk voor een deel zelf moet maken. Dat wil niet zeggen dat hij autonoom is. Hij is schatplichtig aan het kennisbestand van de psychologie, de sociale context en het instituut waar hij werkt, de cultuur en de samenleving. Dat betekent het kiezen van een positie ten opzichte van de stand van zaken in wetenschap en ten opzichte van de samenleving.

#### *Wetenschappelijke verantwoording van diagnostiek*

Wetenschappelijke activiteit is bedoeld om de werkelijkheid binnen en buiten de mens - het Ik en het Niet-Ik - te beschrijven, verklaren, controleren en voorspellen. Aan de Ik - niet-Ik tweedeling kunnen we ons bijna niet onttrekken. Het Ik (n = 1 onderzoek) is de *core business*

van de diagnostiek. Wat tot het Ik gerekend wordt wijzigt zo nu en dan. Leer-, ontwikkelings- en informatieverwerkende mechanismen worden bijvoorbeeld vervangen door biologische en neurologische processen. Een diagnosticus beschrijft en verklaart gedrag voornamelijk met psychologische mechanismen en gedragsconstructen. Daar heeft hij een beetje verstand van. We zijn van ons gedrag enigszins eigenaar (laat Freud het niet horen) en dat zijn we niet van neurologische en biologische mechanismen en processen.

Naast voorspellen, *verklaren*/controleren en beslissen is *beschrijven* een doel van diagnostiek. Het bereiken van die doelen hangt af van conceptuele analyses en empirische bevindingen, onder meer in de gedaante van resultaten van meta-analyses en van gezond verstand. Soms vraagt onderzoek naar de bekende weg en is het pseudo-empirisch.

*Verklaren* gaat terug op de Griekse Pre-Socratici en Plato en Aristoteles. Het ging hen om het vinden van een verklarend principe (oorsprong, archè), zodat alles opgenomen kon worden in een overkoepelend geheel. Zo worden uiteenlopende zaken verklaard: Aristoteles wilde bijvoorbeeld met zijn principe van de gulden middenweg zowel evenwichtstoestanden bij natuurverschijnselen (harmonie der sferen) als politiek (zijn vorm van stadsstaat-democratie) en ethiek (hoe het goede leven te bereiken) verklaren.

*Beschrijven* is onder meer categoriseren, denk aan Linnaeus en de DSM. Dat kan een doel op zich zijn. Deze voorkeur is terug te vinden bij psychometrici, maar ook bij structuralisten als Piaget. De laatste is sympathiek omdat hij het voor elkaar kreeg om Kantiaanse wijsgerige begrippen te operationaliseren in taken en experimentjes. Van vertegenwoordigers van beide werk- en denkwijzen wordt verwacht dat ze een brug naar observeerbaar gedrag slaan. Als dat niet gebeurt, is het weliswaar knap en elegant, maar zonder betekenis voor diagnostiek.

In de diagnostiek komen beide, *beschrijven* en *verklaren*, als *hybride* aan de orde. Diagnostische kennis bestaat voor een deel uit categoriesystemen, structuren en formules én berust op empirische feiten én op ervaringskennis én op diagnostische praktijken als observatie, experimentjes om gedrag te ontlocken aan de cliënt en het gebruik van een valide instrumentarium. Die twee moeten elkaar kunnen 'vinden', formeel gezegd op elkaar af te beelden zijn.

Elke activiteit - ook de wetenschappelijke - is gebaseerd op alledaagse, geleefde waarneming. Dat is de basis van onze kennis van de buitenwereld en onszelf: Merleau-Ponty<sup>6</sup> en zijn opvolger aan de Sorbonne: Piaget.

Het willen kennen en begrijpen van gedrag en de fysische werkelijkheid vindt zijn oorsprong in onze rationele en emotionele uitrusting. Rationaliteit leidt tot theoretische constructen, analyse, het scheppen van de orde in verschijnselen en gedrag en tot de eis van empirische toetsing. Emotionaliteit leidt tot acties om het verlangen naar inzicht te stillen en schiept de behoefte ons lot te verbeteren. Emotie brengt ongerustheid en onzekerheid, maar ook de bereidheid te werken, te handelen om ons inzicht te vergroten en ons lot te verbeteren. Zonder emotie geen actiebereide, je zou bijna zeggen, geen ondernemende wetenschapper.

Er is een pikorde in de wetenschap. Deze zet de psychologie achter de fysica, scheikunde, neurologie en biologie. Simonton<sup>7</sup> heeft aan de hand van zeven criteria een hiërarchie van wetenschappen beschreven. Als gammawetenschap staat psychologie/diagnostiek op de middelste trap van de wetenschappelijke prestigeschaal. De *science markers* zijn volgens hem: (1) concentratie van geciteerd worden (2) vroege, snelle *impact rate* (3) peer evaluatie consensus (4) snelheid van vergeten worden (5) prominente aanwezigheid van grafieken: de mate waarin grafieken verschijnen in artikelen in belangrijke tijdschriften (6) hoe vaak geraadpleegd en (7) de theorie-wet verhouding. Dit levert de volgende rangorde op: 1. Fysica 2. Scheikunde 3. Biologie 4. Psychologie en 5. Sociologie. Psychologie staat in het midden met sociologie onder zich en de fysica en de biologie boven zich. *Binnen* de psychologie is er ook een pikorde met functioneleer aan de top en ontwikkelings- en klinische psychologie onderaan. Dat zal zo blijven en ze zullen niet van elkaar leren.

De identiteit van psychologen ligt minder vast dan die van fysici en medici. Meedeinend met de tijdgeest en zo nu en dan geconfronteerd met een identiteitscrisis, richt de psycholoog zich op neurologie en hersenonderzoek. Het geloof in het 'DNA-maals' wordt substitueert van het geloof in het 'hierna-maals'. Evenals bij religie gaat het soms het om belangen, al wordt het geloof als liefde voor waarheid en mens 'gecommuniceerd'. Onzekerheid over psychologie als wetenschap uit zich ook in het omarmen van statistische modellen. De modellen zijn 'waar, geldig' op de wijze van regels van het schaakspel. Bij confrontatie met psychologische onderzoeksvragen lopen ze door hun complexiteit soms voor op de realiteit van concrete gedragingen.

Sociologie, antropologie, taalkunde en wijsbegeerte worden onder psychologie geplaatst in Simontons hiërarchie. Deze disciplines gaan over groepsgedrag en culturele praktijken. Deze liggen niet zover van de psychologie af. Ze formuleren mechanismen die er toe doen om het gedrag van een cliënt te beschrijven, te begrijpen en te verklaren.

De hang naar hogere wetenschap ('de geur van hoger honing', zegt Nijhoff in 'Het lied der dwaze bijen') en naar exactheid is een mogelijk symptoom van twijfel of psychologie een wetenschap is. Misschien bevat ze slechts informatie over samenhangen tussen gedragingen onderling en met de sociale en fysische context. Wetenschappelijke tijdschriften zijn in dat geval bladen, *journals/magazines* en onderzoekers *journalists*. Had Roskam toch gelijk? Deze twijfel aanvaard ik. Misschien is dat wel het geval. De vorm van publiceren volgens het *APA Manual* is overgenomen van de exacte wetenschappen, maar dit leidt er niet toe dat de inhoud ook wetenschap is. Misschien is het slechts informatie, nuttige overigens. Immers, hoe vaak laat empirisch onderzoek niet zien, dat we over de kennis van samenhangen en effecten van interventies niet zo hoog van de toren hoeven te blazen.

Dispuut op grond van argumenten is onvermijdelijk in wetenschappelijk en elk ander werk. Dat is iets anders als twijfel. Wetenschappers kunnen echter ontspreken door zich in te stellen op het overwinnen van anderen en op het behalen van succes door bijzaken-barok. Dat wil

zeggen dat ze zich met overmaat aan geleerdheid en sluwheid richten op kleine gebieden of populaire maatschappelijke thema's. Ze vertonen daarbij een vaardigheid die weinigen hen zullen verbeteren. 'Geldig' en 'ongeldig' en 'waar' en 'vals' zijn *tactische* begrippen in dienst van het eigenbelang (vrij naar Menno ter Braak<sup>8</sup>).

Wetenschap heeft concurrenten in systemen die snel en afdoende vragen over onze gedragingen beantwoorden. Dat zijn ideologieën, religies, heersende opvattingen, reclame, lobbyisten, goeroes, enzovoort. De psychologie is vooralsnog geen geslaagd systeem. De resultaten van empirische studies en succes van behandelingen geven daar onvoldoende reden toe. Misschien moet ze nooit een geslaagd systeem worden.

#### *De wetenschapper*

Is de diagnosticus een wetenschapper? Misschien, een beetje, *sort of*. Hij maakt gebruik van het kennisbestand en moet voldoende kennis van zaken hebben om dat bestand op waarde te schatten. Maar hij heeft niet de pretentie de wetenschap vooruit te helpen. Dat is toevertrouwd aan wie, ja, aan de wetenschappers. Het 19<sup>de</sup> eeuwse romantisch beeld van de wetenschapper als dat van een uitzonderlijk begaafde is achterhaald. Het cliché van transpiratie (95%) en inspiratie (5%) komt dichterbij wat er gebeurt in onderzoek en praktijk: respectievelijk routine bij het uitvoeren van een project plus het opschrijven volgens richtlijnen.

Een wetenschapper is conservatief. Wat hij te weten is gekomen is kostbaar en moeizaam verworven. Een succesparadigma wordt niet gauw opgegeven. De VU wiskundige Meester<sup>9</sup> voegt er aan toe dat sommigen arrogant zijn en hun kennis en inzichten overschatten. Wat van hun opvattingen afwijkt, wordt als onwetenschappelijk gekwalificeerd. Het ging in dit concrete geval over onzekerheden in de evolutieleer, geuit door een promovendus van Meester. Deze werden opgevat als het afwijzen van de evolutieleer ten faveure van een Bijbelse duiding van het ontstaan van de aarde en het leven. Dat ging Meester te ver.

Wetenschappers en iedereen hebben nooit volledig toegang tot de fysische en sociale werkelijkheid. Kennen verschuift de horizon, verlegt de grens, maar het einde komt nooit in zicht. Kennen neemt niet of nauwelijks toe door integratie binnen en tussen wetenschappen. Daar moeten zij en wij het mee doen; meer zit er vooralsnog niet in.

Wetenschappers vertellen soms wat het publiek en tijdschriftredacties willen horen en vermijden te zeggen wat deze niet willen horen. Gemakkelijk bij het publiek en redacties liggende thema's worden uitgewerkt, bijvoorbeeld man-vrouw verschillen, agressie, positieve psychologie, persoonlijke groei en ontwikkeling, genetische basis van gedrag en de rol van hersenen voor begaafdheid en pathologie. Wetenschappers zijn de laatste 50 jaar meer van het publiek afhankelijk geworden dan daarvoor. Dit wijzigt ook weer na een zekere tijd. Ze werken waarschijnlijk het best in de luwte.

Wetenschappers kunnen het publiek en bestuurlijke elite een spiegel voorhouden en aan de hand van gegevens of informatie tonen dat ze niet het centrum van de wereld zijn, geen onschuldige *tabula rasa*, niet rationeel of irrationeel, niet vrij of onvrij, niet gelijk of ongelijk, niet onafhankelijk of afhankelijk, niet slecht of goed, niet transparant of enigmatisch, enzovoort. De boodschap komt er op neer dat het publiek, de cliënt en de elite zichzelf voor een deel moeten uitvinden. Een wetenschapper kan dat niet voor hen 'regelen'.

#### *Scientists go, where the money goes*

En wie niet in onze neoliberale utilitaristische samenleving? Dit uit zich in aansluiting bij modieuze thematieken, goed in het gehoor liggende theorie en methodologie. Een ter plekke verzonden typologie van psychologen met mijn voorkeur voor driedelingen luidt: De Opportunist: steeds weer een nieuw thema kiezen dat het goed doet: van geleerde hulpeloosheid naar positieve psychologie. De Monopolist: één helder populair thema kiezen en dat monopoliseren, bijvoorbeeld de Big Five, *Attachment*, *Adoptie*, *Echtscheiding*, enzovoort. De Bekeerling: van persoonlijkheidspsycholoog naar psychometricus of vice versa; van de fenomenologie naar het logisch positivisme en omgekeerd. Op je schreden terugkeren heeft iets moedigs.

Wetenschappers mogen, moeten en kunnen meestal of af en toe in vrijheid analyseren, denken en theoretiseren. Kennis uit alle bronnen, van logica, wiskunde, biologie, neurologie, filosofie, sociologie, antropologie tot de verhalen van je buurvrouw mag daarbij betrokken worden. De vrijheid is er ook aan de methodologische zijde: er rust niet bij voorbaat een taboe op intuïtieve, subjectieve of kwalitatieve benaderingen naast de experimentele en correlatieve methoden. Het is wel oppassen om niet te verdwalen in vertalingen in de vorm van reducties, verborgen mechanismen, eigenbelang, producten van neurale activiteiten, hersenscans, kleinste eenheden en veronderstelde hogere mentale bewustzijnstoestanden. Dit verdwalen lijkt op het *Lost in translation*, de film, waar de twee congresgangers, Bill Murray en Scarlett Johansson, niets verstaan van de lezingen en maar met elkaar gaan praten, zoals studenten bij hoorcolleges. De verticale metafoer van het dieper en hoger zoeken, bevat als iedere metafoer - ook de horizontale - het risico ... *of holding us captive* (Wittgenstein; 1953, Sectie 115).

#### *Wetenschap en samenleving*

Wetenschap en/of kennis/inzicht/informatie worden door elites gebruikt om hun positie te behouden en privileges te beschermen. Het publiek wordt op afstand gezet, terwijl het ziet dat wetenschappers het oneens zijn en er zich onder hen fraudeurs bevinden<sup>10</sup>. Er was eens de tendens fraude binnenshuis te houden. De cultureel antropoloog Köbben werd afgehouden van het kond doen van zijn resultaten over de gang van zaken en mogelijke ontsporingen in de wetenschap. Nu is er veel aandacht voor en als gewoonlijk slaat het even door. Van de weeromstuit wordt in 2015 aan burgers, het publiek gevraagd thema's en vragen te noemen, waar onderzoekers zich over moeten buigen. Van het ene extreem naar het andere. Waar of

geldig is niet wat de gemiddelde burger zegt en denkt. Dat kan niet, want ze hebben verschillende en niet-gearticuleerde geldigheidscriteria en vooral verschillende belangen.

### *Het kennisbestand*

Vooruitgang en stagnatie in diagnostiek hangen samen met ontwikkelingen in de psychologie. Is er na Watson, William James, Skinner, Freud, Piaget, Newell, Simon, Tversky en Kahneman en in Nederland na mijn inspiratiebronnen Strasser, Linschoten, Kouwer en De Groot iets nieuws te verwachten? Een doorbraak of een andere kijk, die ons dichter brengt bij begrijpen/verklaren van het gedrag van de cliënt? Komt er een nieuw en belangrijk facet van gedrag aan het licht? Kan er originaliteit verwacht worden op een gebied dat zo breed is en waar alle mogelijke standpunten, voorkeuren en temperamenten al lang en vaak beschreven en vertoond zijn en iedere zogenoemde nieuwe gedachte meteen in een vertrouwde categorie ondergebracht kan worden? Of doen er zich subtiele, bijna onopgemerkte verschuivingen voor, zoals met de PC en het internet? Er zijn vragen genoeg over voor de onderzoeker en diagnosticus.

De inhoud van het kennisbestand is niet te overzien. Meta-studies vatten samen, maar leiden zelden tot nieuwe conclusies. De gesofisticeerdheid van analyses neemt toe, maar ook dat is bijna nooit aanleiding om eerdere conclusies te herzien. Er is te veel en te verspreide inhoud. Je kunt maar een vingerhoed lezen uit het *mer à boire* dat de psychologische literatuur is. Of meer ook tot beter inzicht leidt, betwijfel ik. Het lijkt er veeleer op dat onderzoek de oude, al bestaande kennisvoorraad op peil houdt: de winkel blijft op orde; er is geen revolutie of omslag. Kwam die er maar.

Pogingen om verschillende theorieën en methoden te verzoenen of te integreren werken niet. Geef ze afzonderlijk plaats en aandacht, want combinaties leiden tot onwerkbare hybriden. Maak wel een orde en laat zien dat ze niet tot logisch tegenstrijdige uitspraken over gedrag kunnen leiden zijn, en dat ze niet tegen elkaar ingeruild kunnen worden.

Meta-analyses en meta-analyses van meta-analyses brengen geen nieuws en maken geen eind aan eenzelfde type onderzoek of een specifieke onderzoeksvraag. Heeft onderzoek ooit een vraag volledig beantwoord, zodat we er nooit meer op terug hoeven te komen? Lost meta-analyse een probleem definitief op? Is meta-analyse niet slechts boekhouden? Is deze analyse niet de uiting van het westers empiristisch geloof in de Zelfopenbaring van het Zijn? Meta-analyses, Big Data en *data mining* zijn niettemin zinvol, al steunen ze op de empiristische veronderstelling: hoe meer onderzoek, hoe beter. Empirisch psychologisch onderzoek ontucht immers steeds weer en matigt verwachtingen. Onderzoek kan niet zonder een gedisciplineerd 'geloof'. Ze heeft buiten de empirie om bedachte structuren en modellen nodig en ook hermeneutisch gewonnen betekenissen. Onderzoek kan ook niet buiten de emotie iets te willen en iets te ondernemen om te kennen, leren, begrijpen, verklaren, duiden, enzovoort.

Methoden, onderzoeksontwerpen, statistiek en protocollen functioneren als ze helpen de vraag, het probleem van de cliënt te analyseren. Je mag vragen wat KTT en IRT, studies van het klinisch oordeel, statistische modellen en regels opleveren voor het diagnostisch handwerk. Wat leren heuristieken en *biases, de multi-attribute-utility-theory*, het hypothese toetsend model (HTM) en de Bayesiaanse regel je als diagnosticus? Hun intrinsieke kwaliteit wordt niet betwist, wel de pretentie om het diagnosticeren van het gedrag van de cliënt afdoende te regelen.

Protocollen, regels en statistische modellen aanvaarden we alle als nuttige, maar beperkte hulpen bij het diagnosticeren van de cliënt. Ze kunnen hun doel voorbijschieten, gegeven de beperkte informatieverwerkende capaciteit en de natuurlijke werkwijze van de diagnosticus. Terughoudendheid ten opzichte van DBCs is gerechtvaardigd. De D (diagnose) stelt beperkt in staat tot beschrijven, voorspellen en controleren van het gedrag van de cliënt. De practicus kan dan ook met recht zeggen dat de meerwaarde van diagnostiek gering is. De B (behandeling) kent echter ook beperkingen: effectgroottes van *treatments* zijn bescheiden. Voor kinderen en adolescenten met psychische klachten is er bijvoorbeeld in Nederland een honderdtal DBCs. Het is onwaarschijnlijk dat ze passen bij 100 verschillende problemen of bij 100 te onderscheiden groepen cliënten of bij 100 stoornissen, laat staan dat van elk uitgezocht is of ze *evidence-based* is.

#### *De diagnosticus*

Hij krijgt te veel en te uiteenlopende informatie op zijn bord. Je kunt je voorstellen dat hij houvast zoekt bij de DBCs, het HTM, bij goedgekeurde tests en bij wetenschappers die zeggen te weten 'hoe diagnostiek op een goede manier' te bedrijven. Toch moet hij het steeds weer zelf uitzoeken bij iedere cliënt. Informatie uit het kennisbestand is niet eenduidig, de feiten berusten op steekproeven en de voorschriften zijn niet steeds gebruiksvriendelijk en vaak complex. Hij bijt niet in de hand die hem voedt, maar hij moet uitkijken niet overvoerd te raken en er zelf voor zorgen dat hij de goede ingrediënten binnenkrijgt.

Er is interactie met de cliënt. Spontane interacties kennen doorgaans een grillig verloop. In de diagnose is er een zekere stilering, maar het wordt nooit een chemisch proces of een vaste architectuur. Het gedrag van de cliënt is ten slotte niet volledig te vatten in één laag (bijv. responsen, of driften, motieven, enz.) en in één theorie of epistemologische oriëntatie.

Hij moet zijn activiteit verdedigen. Zijn handelen wordt soms geridiculiseerd door onderzoekers (werk van doctorandi) en door hulpverleners (weinig meerwaarde; je levert een mager halffabricaat). Beider houding belemmert het zicht op wat diagnostiek kan en is en ze kunnen van de practicus leren (Smedslund<sup>11</sup>).

Hij ziet dat de grenzen van voorspellen, verklaren/controleren en beslissen bereikt zijn als onderzoek voortgaat op de wijze waarop het nu vormgegeven en gepubliceerd wordt. Hij kan dit illustreren aan de hand van resultaten uit meta-studies.

Hij is zich bewust dat gedragingen en sociale contexten tijdgebonden, dynamisch en variabel zijn. Hij heeft daarom graag beschikking over betrouwbaarheidsintervallen, effectgroottes en *power* van toetsen. Hij kijkt zelf (*eye-balling*) en gebruikt zijn gezond verstand.

Als aanhanger van Searle<sup>12</sup> vraagt hij om een semantische interpretatie van elke formule, latente trek, patroon van latente variabelen, enzovoort. Er moet een relatie gelegd worden tussen model of formule en gedrag /gedragsconstruct /psychologisch mechanisme. De relatie wordt nooit definitief gelegd, maar je moet er naar blijven streven. Het is niet anders. Als het wel anders was, zouden wetenschap en onderzoek al voltooid zijn.

Het kennisbestand is van dien aard, dat de diagnosticus in zijn streven aan de cliënt recht te doen eclecticus mag zijn en praktisch gezien ook is. Hij streeft naar een rijke en voldoende beschrijving van de vraag van de cliënt.

Hij ziet de proliferatie van constructen/mechanismen niet als teken van vooruitgang. Er is semantische gelijkenis in persoons-, intelligentie- en context constructen, zoals Smedslund beweert. Hij zoekt geen steun en inspiratie bij meer sexy disciplines. Hij blijft bij zijn leest.

Hij beschouwt de bijdrage van neurologische en biologische parameters aan het voorspellen en verklaren als een beperkte aanvulling en kent de empirische steun voor zo'n opmerking: meta-studies genoeg.

Hij weet dat de meeste instrumenten en constructen voor (afwijkende) gedragskenmerken vooral individuele verschillen in kaart brengen.

Hij stemt in met Ellemers<sup>13</sup>, die zegt: '...in ons vakgebied bestaan geen definitieve antwoorden op eenduidige vragen' (p. 12). De zekerheid de *evidence-based* factoren c.q. behandelingen bieden is beperkt en bestaat nog minder op individueel niveau.

Hij is zich bewust dat veel weten over individuele verschillen variabelen niet gelijk is aan veel weten over een individuele cliënt. De psychologie levert voor een deel iets anders op als dat wat de diagnostiek nodig heeft.

Ondertussen blijft hij een manasje van alles dat van alles wat af moet weten, gestileerd moet kunnen interacteren met zijn cliënt en een analyse moet geven van zijn vraag en/of probleem. Daarna geeft hij het diagnostisch proces vorm. Dat is niet volledig te protocolleren, ook al zijn er het HTM, de Multi Attribuut Nut Theorie, Bayesiaanse regels, reeksen richtlijnen en strenge, doch rechtvaardige supervisors. De onderzoeker kan van de practicus leren. De kritiek van de onderzoeker is vaak methodologisch van aard. Het gaat hem om dwalingen in de leer. Dat is niet het hele verhaal. Diagnostici aller landen verenigt U; blij onafhankelijk en focus op de unieke cliënt.

### *Lezen en schrijven*

Elke tekst is een farmacon volgens Plato. Dat wil zeggen zowel een medicijn tegen het vergeten als een gif voor het geheugen en het kennen. Elke lezer loopt het risico ontgift te moeten worden na lezing van wat dan ook en na het aanhoren van een verhaal van wie dan ook. Elke tekst verhuult en onthult. Door iets te zeggen/schrijven, zeg je/schrijf je iets anders niet.

Een tekst kan men op verschillende manieren lezen. Voor de meeste wetenschappers van nu komt dat volgens Fijnaut (NRC mei, 2015) er op neer, dat ze werk van een ander zelden lezen. Als het doen, wordt een enkele keer - zoals ik ooit bij een promotie meemaakte - gezegd dat de schrijver het over iets anders had moeten hebben. Vaker wordt naar eigen thema's gezocht en vastgesteld of eigen werk er wel of niet of (on)juist in verwerkt is: 'Ik word (niet) geciteerd'; 'Ik word verkeerd geciteerd', 'Wordt mij rechtgedaan'? Ook wordt de nieuws waarde van het werk betwist: 'Waar heb ik dat eerder gehoord'.

Een voordehandliggende, maar lastige opdracht van lezen is nagaan of het waar of juist wat er staat. Er zijn daarvoor geen 0 - 1 waarheidscriteria in de psychologie. Het gaat eerder om hermeneutiek: ontdekken van aannemelijke betekenissen en het verschuiven van een perspectief. Nog een manier is retorisch lezen: de argumentatiestructuur ontrafelen. Elke lezer toont wel iets van al deze manieren.

Leesstof is er niet om door studenten gerepliceerd te worden, maar om zich eigen te maken. Dat wil ondermeer zeggen kritisch evalueren en valoriseren. Elke lezer heeft zijn eigen unieke referentiekader bijgevolg valt zijn kader nooit samen met dan van anderen. Als dat het geval zou zijn, zouden lezers en schrijvers al uitgepraat zijn voor ze aan iets beginnen. De lezer die het kader met de schrijver helemaal deelt, vindt alleen bevestiging. Of je daar nu veel van leert? Lezer en schrijver moeten 'schuren', niet te veel, maar wel iets, anders levert het weinig op. Elke tekst, protocol, model, empirisch resultaat wordt rammelende ballast, dood gewicht als het niet wakker gelezen en getoetst wordt door een geïnteresseerde, gemotiveerde en onafhankelijke lezer.

Schrijven is pretentief. Zeker, als het geen nieuwe inzichten biedt en niet oorspronkelijk is. Dat is hier het geval, en dat weet de diagnosticus ook. Hij kan dus geschriften over diagnostiek onafhankelijk bekijken. Hij moet het toch zelf doen.

### *Ten slotte, een ter plekke verzonnen diagnostische persona*

Wat de diagnosticus vermag, is beperkt. Dat weet hij zelf ook. Dit stel ik voor als de *persona* van 'mijn' diagnosticus. Het is een ter plekke verzonnen archetypisch beeld. *Persona* verwijst oorspronkelijk naar het masker dat Romeinse toneelspelers droegen om iemand, 'a character' degelijk en overtuigend uit te beelden. Wie is hij, gegeven zijn kennis, karakter en drijfveren? Welk masker draagt hij?

Hij kent geen wetten, waarheden, dogma's. Hij verwacht niet dat nieuwe apparaten of technieken beslissend zijn voor psychologische diagnostiek. Hij toont relativiseringsvermogen en een door het empirisch kennisbestand van de psychologie getemde verbeelding over het doorgronden van de gesitueerde persoon/denker/kenner die zijn cliënt is. Hij is probleem- en *niet* oplossingsgericht. Hij kan ertegen dat hulpverleners, als pedagogen en psychiaters hem dat verwijten. Hij is een evenwichtskunstenaar die het midden houdt tussen het Hegeliaanse *Dass Wahre ist das Ganze* en het reductionisme van het behaviorisme, neurologisme, empirisch-analytisch methodologisme, staticisme, monotheïsme of de beperking tot één psychologische thema of één methodische invalshoek.

Het midden houden, maat houden is één van de vier Aristotelische deugden. Ze kwamen in de Middeleeuwen vaak aan de orde bij begrafenistoelafspraken. Daar hebben we nu de *Big Five* voor. Ze passen wel enigszins op diagnostiek van de cliënt: matigheid, verstandigheid, voorzichtigheid, dapperheid, sterkte en rechtvaardigheid. Het is overigens uitkijken met ijveren voor deugden, want de deugd is door niets zo veel schade toegebracht dan door de saaiheid van haar pleitbezorgers.

Vrij naar Kant: *Man kann denken wass man will*. Bij deze. Met zeggen wat je wil moet je wat voorzichtiger zijn.

## Referenties

1. Wijnberg, N.M. (2015). *Alle collega's dood*. Amsterdam: Van Genneep.
2. Kuhn, T.S. (1962). *The structure of scientific revolutions*. International Encyclopaedia of Unified Science. Vol. 2. Chicago: The University of Chicago Press.
3. Simon, H.A. (1991). *Bounded rationality and organizational learning*. Online beschikbaar: Permalink: <http://dx.doi.org/10.1287/orsc.2.1.125>.
- Larkin, J.H., McDermott, J., Simon, D.P. & Simon, H.A. (1980). Models of competence in solving physics problems. *Science*, 208, 1335-1342.
4. Van Dijk, L. & Wilthagen, J. (2014). The horizontal world view. A Wittgensteinian attitude towards scientific psychology. *Theory and Psychology*, 24, 1, 3-18.
5. Wittgenstein, L. (1953; translated 1999). *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford, UK: Blackwell Publishing.
6. Merleau-Ponty, M. (1945). *Phénoménologie de la Perception*. Paris Gallimard: translated by Colin Smith 'Phenomenology of Perception' (1958). London: Routledge Classics. Nederlandse vertaling beschikbaar: Fenomenologie van de Waarneming.
7. Simonton, D.K. (2004). Psychology's status as a scientific discipline: Its empirical placement within a hierarchy of the sciences. *Review of General Psychology*, 8, 59-67.
- Simonton, D.K. (2009). Varieties of (scientific) creativity. A hierarchical model of domain-specific disposition, development, and achievement. *Perspectives on Psychological Science*, 4, 5, 441-453.
8. Menno ter Braak, geciteerd in L. Hanssen (2003): Biografie Deel II: *Sterven als polemist*. Amsterdam: Meulenhoff.
9. Meester, R. (2014). *Arrogant: waarom wetenschappers vaak minder weten dan ze denken*. Utrecht: Ten Have.
10. Over fraude, zie Kolfshoten, F. van (2012). *Ontspoorde wetenschap: Over fraude, plagiaat en academische mores*. Amsterdam: Uitgeverij De Kring. Al eerder heeft hij samen met Köbben geschreven over fraude in het wetenschapsbedrijf. Daarvoor heeft Köbben zich ingehouden om zijn cultureel antropologisch onderzoek naar het wetenschapsbedrijf te publiceren. Kennelijk zaten sommige instituten daar niet op te wachten.
- Het Slot* van Kafka (1999) heb ik gelezen als het ontbreken van welke inhoud dan ook (Amsterdam: Athenaeum-Polak-Van Genneep).
11. Smedslund, J. (2009). The mismatch between current research methods and the nature of psychological phenomena: What researchers must learn from practitioners. *Theory & Psychology*, 19, 778-794.
12. Searle, J. (1984). *Minds, Brains and Science: Can computers think*. Harvard University Press: Cambridge Massachusetts (pp. 28-41).
13. Ellemers, N. (2013). Terug naar de grote lijn. *De Psycholoog*, 48, 12, December, 2013, pp. 12-20.

Voor resultaten uit meta-studies die uiteenlopende criteria prediceren, zie Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszin.T.W. & Reed, G.M. (2002). Amplifying issues related to psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist*, February, 2002, pp. 140-143. Ze verdedigen zich in deze bijdrage tegen de kritiek dat voorspelling op basis van medische en biologische variabelen tot betere resultaten leidt dan psychologische. Naar de 2001 studie met resultaten over predictieve validiteit van tests is eerder in de tekst verwezen.

