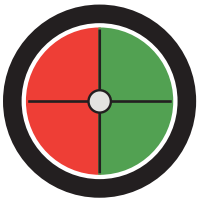


leadership.be

(On)Betrouwbaarheid en validiteit van theorieën,

modellen en psychometrische testen.



INHOUDSTAFEL

Doel van dit rapport	2
Voorafgaandelijke informatie over de gebruikte terminologie	3
Betrouwbaarheid (reliability) van tests	3
Validiteit (validity) van tests	4
Conclusie	5
Analyse van (therapeutische) methoden	6
Cognitieve gedragstherapie (CGT)	7
NLP	8
REBT	10
Analyse van meetinstrumenten	12
Belbin teamroltest	12
Enneagram	15
HBDI (Herrmann Brain Dominance Instrument)	17
LIFO (Life Orientations)	21
MBTI – Myers-Briggs Type Indicator®	23
NEO-PI-R (big five persoonlijkheidstest)	31
N-IAS	37
Thomas-Killmann Instrument (TKI)	39
Eindconclusie	40
Kan het inzetten van tests nuttig zijn?	40
Het type van tests en hun wenselijkheid	40
Wat is “wetenschappelijk”?	41
De gevaren van pseudo-wetenschappelijke tests of methodes	41
Eén of meerdere tests?	44
Aanbevelingen	44

Doel van dit rapport

De bedoeling van dit rapport is onze klanten op een zo onafhankelijk mogelijke manier te informeren en tevens onze belofte waar te maken om een betrouwbare en vooruitstrevende partner voor onze klant te zijn.

Wij willen een partner zijn die betrouwbaar en vooruitstrevend is.

- ▶ Betrouwbaar door samen te werken met onderzoekers en professoren uit de universitaire wereld en door voor modellen te kiezen die een degelijke **wetenschappelijke onderbouwing** hebben. De modellen die wij hanteren zijn liefst altijd gebaseerd op peer-reviewed wetenschappelijk onderzoek, gepubliceerd in degelijke tijdschriften. Daarbij houden wij ons ver van pseudo-wetenschap.
- ▶ Betrouwbaar ook door enkel te werken met **senior consultants** die een bewezen track-record hebben in het bedrijfsleven en die reeds verschillende jaren actief zijn in **training, coaching en begeleiding** van leidinggevenden.
- ▶ Vooruitstrevend door open te staan voor **nieuwe ontwikkelingen**, zonder evenwel te snel mee te gaan in niet bewezen modetrends. Om de combinatie betrouwbaar en vooruitstrevend te kunnen realiseren, volgen wij de wetenschap op de voet dankzij eigen inspanningen en onze contacten met onderzoekers en professoren.

Dit betekent dat wij ook afstappen van tests, modellen en theorieën, waarvan gerepliceerd wetenschappelijk onderzoek bewijst dat ze onvoldoende betrouwbaar of zelfs foutief zijn. Modellen waarvan het onderzoek niet overwegend eenduidig is, laten wij dus vallen.

Zo werkt Performance Coaching al geruime tijd niet meer met MBTI, ondanks het feit dat dit een zeer moeilijke keuze was. MBTI is namelijk een commercieel interessant meetinstrument en ervan afzien heeft ons dus aanvankelijk omzet gekost, en wellicht nog.

Eens men echter keuzes maakt, moet men consequent zijn. Waarden zijn namelijk maar waarden en écht geloofwaardig als men deze in daden omzet.

Tenslotte nog dit: wij zijn ons ervan bewust dat sommige bedrijven en klanten nog met een aantal van de in dit rapport “gewraakte” modellen en tools werken, en dat dit soms op onbegrip of kritiek kan stuiten. Dit nemen we er bij, want wie kleur bekent, moet ook de gevolgen dragen. Wij gaan ervan uit dat **kwaliteit en eerlijkheid** op langere termijn echter bovendrijven en daarom deze keuze.

Voorafgaandelijke informatie over de gebruikte terminologie

Betrouwbaarheid (reliability) van tests

Als over betrouwbaarheid wordt gesproken in psychometrische testverslagen, dan betekent dit in 'mentaal' meestal dat men nagaat of de test opnieuw dezelfde resultaten (of alleszins dezelfde interpretatie van de verschillende meetresultaten) oplevert bij een latere meting (te vergelijken met de betrouwbaarheid van een meetlat: meet de lat dezelfde afstanden telkens opnieuw?).

Er zijn nochtans minstens twee manieren om betrouwbaarheid te meten:

Vaak worden testpersonen gevraagd om na enkele weken opnieuw de test af te leggen. Als de test vaak dezelfde resultaten meet, dan spreekt men van een goede **test/retest betrouwbaarheid**. Nochtans dient hierover ook gezegd dat dit geen garantie voor juistheid is: de test is niet nieuw meer voor de testpersoon, de testpersoon kan een andere situatie voor ogen hebben, enzovoort.

Interne consistentie meet eerder of alle items (vb vragen) in de test op dezelfde manier meten wat ze moeten meten. De interne consistentie bekijkt ook de correlatie tussen de items die allen dezelfde variabele zouden moeten meten. Vaak wordt dit gemeten door de test in twee te splitsen: beide delen van de test zouden dan hetzelfde resultaat moeten geven (split-half betrouwbaarheid).

De betrouwbaarheidsscores worden uitgedrukt in punten na de komma. In de academische wereld wordt nu 0,70 als 'voldoende betrouwbaar' geacht. Scores van 0,80 worden als zeer hoog beschouwd. Een score van 1.0 zou betekenen dat de test 100% de variantie verklaart, terwijl een score van 0,0 betekent dat dit puur toeval is. Men kan scores echter niet omzetten in procenten (zie verder bij opmerkingen).

Bij ernstige testontwikkeling dient de samenstelling van de steekproef goed te gebeuren en dienen er ook "replicaties" te volgen op verschillende steekproeven

Opmerkingen:

- Betrouwbaarheid is de bovengrens van validiteit. Een validiteitscijfer kan, theoretisch gezien, nooit boven de betrouwbaarheidscijfers uitkomen.
- Betrouwbaarheid zegt niets over het feit of de theorie of het model al dan niet juist is en op wetenschappelijk onderzoek berust. **Het dus is een louter statistisch gegeven!**
- Om een betrouwbaarheidsindex te kunnen interpreteren moet je natuurlijk ook weten in welke ratio deze is uitgedrukt; meestal is dat een r-correlatie, maar het kan ook een tau-index of nog een aantal andere zijn. De ene ratio geeft al gauw een veel hoger getal dan de andere. R. is de meest gebruikte.

Om een concreet idee te geven wat het effect van betrouwbaarheid is maken we de vergelijking met een IQ meting. Stel, iemand heeft een IQ score van 100 op een test. De test heeft een betrouwbaarheid van 0.90. Dan ligt de ware score (met 95% zekerheid) ergens tussen 90 en 109 IQ. Bij een betrouwbaarheid van 0.60 zou de echte score ergens liggen tussen 81 en 119 IQ. Als je weet dat 80 de grens van mentale retardatie is en dat sommigen vanaf 125 IQ over hoogbegaafdheid durven spreken, is dergelijk laag cijfer van betrouwbaarheid erg bedenkelijk. Bij een betrouwbaarheid van 0.40 zou je tegen de man of vrouw in kwestie moeten vertellen dat het IQ ergens tussen 77 en 123 ligt, dus de persoon in kwestie kan mentaal geretardeerd zijn of ook wel hoogbegaafd.

Validiteit (validity) van tests

Dit zegt iets over het feit of er “standaards” zijn opgenomen. Meten de vragen wel wat er gemeten moet worden (bijvoorbeeld de stellingen naar voren gebracht in de theorie of hypothese)?

Face validity: Zegt iets over de geschiktheid van een instrument om iets bepaalds te meten; hierbij wordt het instrument zelf beoordeeld. Het is de minst betrouwbare vorm van validiteit en wij raden aan hier nooit rekening mee te houden.¹

Content validity: Gaat na of de vragen de inhoud, vermeld in de theorie, meten. De meeste tests beantwoorden hieraan omdat de vragen meestal zijn afgeleid van de theorie.

Criterion validity: (ook soms predictive validity genoemd). Gaat na of de test een voorspellende waarde heeft. Meestal wordt hiervoor een vergelijking gemaakt met andere tests die al bewezen hebben een hoge validiteit te hebben. Over de voorspellende waarde:

- een IQ-test meet dat sommige kinderen een hoger IQ hebben dan anderen, en de test voorspelt dat ze dus betere schoolresultaten zouden moeten halen
- als dit zo blijkt te zijn in werkelijkheid, is er sprake van een hoge *criterion validity*.

Construct validity: deze meet de aparte delen van de samenstelling van het model. Meestal zijn modellen complexe constructies van eenvoudiger bestanddelen. Die eenvoudiger bestanddelen (of variabelen) test men dan met diverse andere testen. Vervolgens giet men de bekomen resultaten in een soort matrix en leidt men zo af of er een correlatie tussen variabelen en de test bestaat (is er een hoge overeenkomst tussen variabele en testresultaten waar nodig, en een lage overeenkomst wanneer er geen connectie zou moeten zijn - dit wordt ook “convergent” en “discriminant” genoemd).

¹ Zie onder andere Trochim, W. (2002). "Measurement Validity Types." en Lacity, M., & Jansen, M. A. (1994). Understanding qualitative data: A framework of text analysis methods. *Journal of Management Information Systems*, 11, 137-166. Ook in België bestaat hierover bij wetenschappers consensus over: face validity is “waardeloos”.

Opmerkingen:

- Of een test waardevol is of niet kan je statistisch alleen maar afleiden als de test zowel valide als betrouwbaar is. Naast de statistische validiteit en betrouwbaarheid moet een test verder kritisch bekeken worden: zo zijn vele tests zo vaag opgevat dat ze een hoge betrouwbaarheid geven; nochtans kan de test waardeloos zijn. Een ander euvel bij zelftests is dat de resultaten niet noodzakelijk de werkelijkheid weergeven: diverse onderzoeken geven immers aan dat ongeveer 25% van de bevolking geen accuraat zelfbeeld heeft.
- Bovendien moet wetenschappelijk onderzoek ondermeer repliceerbaar zijn. Dit betekent dat het onderzoek door andere onderzoekers (in andere labo's indien het onderzoek betrekking heeft op laboratoriumproeven) herhaalbaar moet zijn en daarbij de onderzoeksresultaten van de eerdere studies bevestigen.

Conclusie

Hoewel betrouwbaarheid en validiteit veel kunnen zeggen over de test op zichzelf – en dus “gepruts” onderscheidt van “professioneel werk”, **zeggen deze begrippen niets over de juistheid van de theorie die er achter schuil gaat. Het zijn louter statistische gegevens.**

Zo is het best denkbaar dat bepaalde testen zo ruime en/of vage beschrijvingen geven van wat ze meten, dat iedereen zich herkent. In het geval de test alleen maar een “zelfbeoordeling” inhoudt (en geen input van derden) zou statistisch gezien de betrouwbaarheid en validiteit ongetwijfeld hoog kunnen liggen, maar moet men toch nog alert blijven of daardoor de waarde van de test overeind blijft. Bij een aantal tests is dit zo (ze blijken voldoende waarde te hebben als zelfbeoordeling), bij een aantal andere tests zien we duidelijk dat een 360° gebruik meer waardevolle resultaten oplevert.

Daarom een aantal aanbevelingen:

- a) **Ga nooit af op de eigen vermeldingen van onderzoeksresultaten** bij de verdelers of ontwerpers zelf, zeker niet als het om commerciële tests gaat.
- b) **Ga na of er onafhankelijke onderzoeken zijn uitgevoerd**, en welke resultaten deze gaven; zo kan je bijvoorbeeld via de betalende sites **academisch onderzoek** achterhalen bij “psychlit”, bij “medline” en bij “Buros”. De afwezigheid van onderzoek is veelbetekenend: dit betekent vrijwel steeds dat de test zelfs niet de moeite waard werd gevonden om te onderzoeken.
- c) Het is steeds zinvol na te gaan **hoeveel testpersonen** deelnamen aan de testen (100 lijkt wel een erg absoluut minimum – 1000 is al heel wat beter) en of er een voldoende sociografische (demografische) spreiding was (in het geval u zich over een grote populatie wil uitspreken).
- d) **Vraag desnoods de mening aan een prof** die toegang heeft tot onderzoeksresultaten.

Analyse van (therapeutische) methoden

Cognitieve gedragstherapie (CGT)

Wat is CGT?

Deze moderne vorm van therapie is in feite het samengaan van twee strekkingen: de cognitieve therapie en de gedragstherapie. De gedragstherapie baseert zich op resultaten uit de experimentele psychologie. Principes als klassieke en operante conditionering, leerprincipes enz. zijn afgeleid uit deze onderzoeken.

Deze aanpak gaat er grotendeels van uit dat onze leergeschiedenis bepaalt hoe we denken, voelen en handelen. Bepaalde interne of externe stimuli (vb een welbepaalde situatie) roept gevoelens, gedachten (inclusief verwachtingen) en gedragingen op.

In een klinische context (met patiënten) wordt meestal gewerkt in multidisciplinaire teams (psychiaters, neurobiologen, psychologen, sociaal assistenten).

Bijna altijd gaat de therapeut eerst een gedragsanalyse en een functie- en een betekenisanalyse opstellen, om daarna met de patiënt/coachee doelen te stellen.

Wat is de waarde van CGT?

Het onderzoek naar werkbare psychotherapie zit de laatste jaren in een stroomversnelling. Uit experimenteel onderzoek en uit onderzoek waarbij de superioriteit van de behandeling dient aangetoond te worden ten opzichte van placebo controlegroepen of andere bona fide behandelingen blijkt cognitieve gedragstherapie zeer werkzaam bij het oplossen van bepaalde emotionele problemen:

- ▶ Angststoornissen gaande van angst voor spreken in publiek, gebrek aan assertiviteit tot regelrechte fobieën zoals smetvrees, dwangneurosen. Hierbij wordt meestal beroep gedaan op een graduele exposure in vivo (= geleidelijke blootstelling aan het voorwerp van de angst)
- ▶ Depressie
- ▶ Boosheid en agressie waarbij vooral de cognities aangepakt worden en waarbij graduele exposure in vivo secundair zijn.

Ook met de strenge criteria die in medische studies worden gehanteerd blijkt cognitieve gedragstherapie en nog specifiek de exposurebehandeling, heel succesvol te zijn voor de behandeling van angst. Dit leidde ondermeer in Nederland en Frankrijk tot een erkenning van dit soort therapie (en in Frankrijk besloot de Inserm in 2003 na een diepgaande studie dat de cognitieve gedragstherapie steeds superieur was aan ondermeer psychoanalyse en familiale of groepstherapie).

Als **professionele coach** is het onontbeerlijk een aantal van deze methodes onder de knie te hebben omdat het verschil tussen een pathologie (ziektebeeld) en normaal gedrag vaak erg klein is. In de professionele context krijgt een coach vaak te maken met ondermeer volgende problemen waar CGT nodig is:

- ▶ Angst om te spreken voor een groep
- ▶ Algemeen gebrek aan assertiviteit
- ▶ Te autoritair gedrag van leidinggevenden

In België is één van de leidende onderzoekers Dirk Hermans. U vindt een artikel van zijn hand op: <http://ppw.kuleuven.be/leerpsy/dirk/artikelGT.htm>

NLP

NLP of Neuro-Linguistic Programming™ is een beschermde benaming voor opleidingen tot NLP-practitioner. NLP is een zogenaamde therapeutische methode.

Het is een vrij populaire methode in sommige HR en opleidingsmiddelen. Vanuit wetenschappelijke kant² geniet NLP echter minder respect.

Een kort overzicht waarom NLP niet zou doen wat het beweert te doen en een zuiver commercieel model zou zijn:

- Neurowetenschappelijk onderzoek ondersteunt de stellingen van NLP niet:
 - De beschrijving “neuro-linguïstisch” heeft niets met neurolinguïstiek te maken; dit vak onderzoekt hoe spreken en taal verstaan door de hersenen wordt gestuurd. De huidige onderzoeken ter zake maken gebruik van klinische studies (afasiepatiënten) en van moderne beeldtechnieken zoals positronemissie tomografie en magneto-encefalografie. Deze technieken bestonden destijds niet.
 - De huidige onderzoeken hebben aangetoond dat de typische NLP-bewering “hoe je het zegt denk je het” (PRS: Primary Representational System) volkomen onjuist is: de relatie tussen taal en denken, tussen taal en cognitie is aanzienlijk minder direct. Er is integendeel vaak een verschil tussen de intentie en de manier van gedragen of praten. Empirisch onderzoek toonde ondermeer het volgende aan:
 - De keuze van woorden heeft niets van doen met de “denkmode” (visueel, auditief, kinestetisch...)
 - De richting van de oogbewegingen zegt evenmin iets over het zogenaamde PRS
 - “Matching” door hetzelfde woordgebruik werkt niet: researcher Cody stelde zelfs vast dat NLP-therapeuten die bewust probeerden aan taal-matching te doen, door de behandelde als “minder betrouwbaar” en “minder effectief” werden aanzien!
 - NLP beweert dat het interne processen heeft blootgelegd die we gewoonlijk het ‘denken’ noemen. Dat is de academische wereld ook nu (2003!) nog niet gelukt.
 - NLP heeft nooit (niet in het verleden, noch in het heden) de bevindingen van ook maar één neurowetenschapper vermeld of gebruikt om zijn concept uit te leggen.

² Critici:

- Hoogleraar Willem J.M. Levelt, directeur van het Max-Planck-Instituut voor Psycholinguïstiek te Nijmegen in “Hoedt u voor NLP”, Skepter 9(3), september 1996)
- Dylan Morgan: “A scientific Assessment of NLP” in Journal of the National Council for Psychotherapy & Hypnotherapy Register, Spring 1993
- Dr Heap (British Society of Experimental and Clinical Hypnosis)
- Robert Todd Carrol: “The skeptic’s Dictionary: Neuro-Linguistic Programming”
- John Mickelwait en Adrian Woolridge, in “De Toverdokters” (Business Bibliotheek, Contact)
- Alan Sokal en Jean Bricmont: “intellectueel bedrog” (EPO/De Geus)
- Gustaaf Cornelis en Geerd Magiels : “onzin verkoopt: pseudowetenschappen en de bedrijfswereld (Wonder en is gheen wonder, 2° jaargang nr 4, 12/2002)

- Wat betreft hun beweringen over het nut van hypnose, is de wetenschap er tot op de dag van vandaag niet in geslaagd aan te tonen dat onze geheugenprestaties onder hypnose toenemen (Spanos, Wilson...)
- De beoefenaars van NLP slagen er zelf niet in exact te beschrijven wat NLP juist doet. De eigen modellen van NLP zijn wetenschappelijk niet te achterhalen. Sommige modellen (denken, voelen, doen) en technieken (bijvoorbeeld CBT) die NLP gebruikt zijn bekende therapeutische inzichten of technieken die echter niets (origineel) met NLP vandoen hebben, maar die NLP wel een aura van betrouwbaarheid kunnen geven, omdat een aantal van deze technieken wel degelijk werkzaam zijn (bijvoorbeeld de meer dan 1.000 studies over de REBT-methode van Albert Ellis of de CBT van Beck).
- De resultaten die NLP zou opleveren als therapeutisch middel zijn niet onderzoekbaar en zijn nog nooit onderzocht: de beweringen berusten op “testimonials” en “anekdotische” vertellingen. Vermits er geen tests bestaan, is er ook geen statistisch onderzoek naar validiteit en betrouwbaarheid mogelijk. Er wordt geen medewerking verleend aan bijvoorbeeld categorie I onderzoeken die empirisch onderbouwde resultaten geven: Behandelingen in categorie I zijn behandelingen die bewezen zijn in ten minste twee willekeurig uitgevoerde gecontroleerde “trials” (RCT’s), waarbij de superioriteit van de behandeling dient aangetoond te worden ten opzichte van placebo controlegroepen of andere bona fide behandelingen.
- Richard Bandler en John Grinder bedachten NLP als “een nieuwe vorm van psychotherapie” in de jaren '70, toen hersenonderzoek nog niet de doorbraken kende zoals die de laatste 30 jaar, en vooral de laatste 5 jaar, hebben plaatsgevonden.
- Richard Bandler was een wiskundige en heeft nooit psychologie gestudeerd (!)
- Grinder was een taalkundige en heeft nooit psychologie gestudeerd (!)
- De grondleggers schijnen zich bewust te zijn geweest van de onjuistheid van hun beweringen: in hun boek “Reframing” schrijven ze: *‘Dus, als we toevallig iets noemen waarvan je uit wetenschappelijk onderzoek, of uit statistieken, weet dat het onjuist is, realiseer je dan dat we je een ander niveau van ervaring aanreiken’.*
- Bepaalde uitspraken van Grinder of Bandler zelf zijn hoogst ongewoon: *“...Bijvoorbeeld geloof ik dat het erg goed was dat noch ik (Grinder) noch Bandler gekwalificeerd waren in de domeinen waarin we ons begeven, met name in de psychologie en de therapeutische toepassing ervan; dit is trouwens één van de stellingen die Kuhn heeft geïdentificeerd in zijn historisch onderzoek naar paradigmawisselingen.”* (Grinder) Dit terwijl Thomas Kuhn in zijn boek “The Structure of Scientific Revolutions” op geen enkel ogenblik aangaf dat het beter is om niets af te weten van een bepaald domein om daar doorbraken in het denken te bereiken. Kuhn beschreef dat wetenschappelijke theorieën evolueerden omdat *“wetenschappelijke theorieën afbrokkelen wanneer nieuwe data niet meer kunnen uitgelegd worden aan de hand van de oude theorieën, maar wel aan de hand van nieuwe theorieën”.* Hij beperkte zijn uitspraken trouwens tot de harde wetenschap, en niet de “zachte”...
- En als leuke uitsmijter: het eigen ‘mentale model’ van Bandler en Grinder heeft Bandler niet belet Grinder juridisch te vervolgen voor miljoenen Dollars!

Conclusie: NLP is meer een (commerciële) geloofsovertuiging dan wetenschap en de beweringen die typisch eigen zijn aan NLP (en die niet gekopieerd zijn uit bewezen therapeutische methodes) zijn ofwel bewezen onjuist, ofwel niet onderzoekbaar.

REBT

REBT staat voor **rational emotive behavioral therapy**. REBT is een therapeutische methode die valt onder de zogenaamde CBT (cognitive behavioural therapy – of cognitieve gedragstherapie)

REBT is een vorm van gedragstherapie en wordt vaak aangevuld met diverse technieken en methodes uit andere CBT-vormen.

Geschiedenis

Na 6 jaar psychoanalyse te hebben beoefend, verlaat Albert Ellis in 1953 deze methode om zijn patiënten te helpen. Na heel wat klinisch onderzoek ontwikkelde hij stapsgewijs de RET-methode. Eerst begon hij zijn patiënten 'huiswerk' te geven en vaardigheidstraining. Vooral met fobieën maakten zijn patiënten ongelooflijke vooruitgang – zeker in vergelijking met psychoanalyse. Daar ontdekte hij ook reeds dat het traditionele concept van inzicht in het verleden absoluut niet genoeg is voor diepe en langdurige verandering. Zelfs als patiënten inzicht krijgen in de redenen waarom ze zich nu nog 'fout' gedragen, is dit niet voldoende, maar wel een noodzakelijke eerste stap. Deze stelling van destijds kwam de laatste jaren terug in de belangstelling nadat onderzoeker-psycholoog-journalist Dr. Daniel Goleman al had aangetoond dat gedragingen pas veranderen over een periode van 6 tot 9 maand en dat de ingrediënten motivatie, feedback en uitgebreid oefenen zijn.

In het begin ondervond Albert Ellis veel weerstand omdat hij psychoanalyse bekritiseerde en pleitte voor een "rationele analyse". Hij stelde in 1958 dat mensen twee soorten beliefs hebben: rationele (RB's) en irrationele (IB's), en dat wanneer mensen frequente en intense IB's hebben, ze dan meer verstoord zijn dan wanneer ze zwakke IB's hebben. Later (nu nog steeds!) werd hier veel onderzoek naar gedaan en deze stelling werd in meer dan 1.000 artikels (!!!) bevestigd.³ Over de REBT-methode en andere CBT's bestaan heel wat publicaties: één van de belangrijkste vindt men terug in "**empirically supported psychological interventions: controversies and evidence**", Annual Review Psychology, 2001. 52:685-716.

Hierin vindt men terug dat CBT's erg werkzaam zijn in tal van behandelingen, en vaak in zogenaamde "categorie I" behandelingen: dit zijn behandelingen die bewezen zijn in ten minste twee willekeurig uitgevoerde gecontroleerde "trials" (RCT's), waarbij de superioriteit van de behandeling dient aangetoond te worden ten opzichte van placebo controlegroepen of andere bona fide behandelingen.

Het is pas in de jaren 60 dat Albert Ellis erkenning kreeg, en gevolgd werd door andere bekende therapeuten die hun eigen versie van een CBT maakten.⁴ Nu behoort REBT zeker tot de "mainstream" therapie in de Verenigde Staten.

Kritische noten

Toch kan er ook kritiek geleverd worden op de REBT, met name op het erg lineaire ABC-schema: Activating event → Beliefs → Consequences (feelings & behaviour). Recent onderzoek heeft immers uitgewezen dat (irrationele) overtuigingen niet altijd

³ Onder meer: Clark, 1997; Glass & Arnkoff, 1997, Schwarz, 1997

⁴ Onder meer Glasser (1965), Beck (1976), Bandura (1997) en Meichenbaum (1977).

vooraf gaan aan de zogenaamde gevolgen zoals emoties. Bij angstgevoelige mensen is het zelfs zo dat proefondervindelijk werd vastgesteld dat hun limbisch systeem (een gebied in de hersenen dat bekend staat als het “productiecentrum voor emoties”) al heeft gereageerd en al een fysische reactie heeft uitgelokt (verstijving, hartritme wijziging, adrenaline en cortisol...) nog vooraleer de persoon zich bewust is van het gebeurde. Voor vele vormen van angst beschouwen moderne therapeuten de gedachten (met de typische negatieve verwachtingen) als een “bijproduct” van de angst.

REBT toepassen zoals het “protocol” het voorschrijft is dus minder gepast voor angst, maar bijvoorbeeld erg werkzaam voor problemen van zelfbeheersing van boosheid (irritatie, frustratie, ongeduld...), aangevuld met andere methoden uit de cognitieve gedragstherapie.

De huidige REBT-methode is een geactualiseerde versie van de oude RET (waarover de meeste boeken nog gaan), waarin ook inzichten en methodes worden gebruikt zoals ontwikkeld door Beck, Seligman en anderen.

In een **professionele coachingcontext** is de REBT methode uitermate geschikt voor het coachen van de autoritaire mensen, ook als dit het gevolg is van gebrek aan beheersing van gevoelens als boosheid, irritatie, ongeduld... Bovendien leert onze eigen ervaring dat dit op een zeer hoog niveau (directie!) aanvaard wordt omdat het opstellen van schema's (ABC) en het Disputen (uitdagen van irrationele gedachten) een erg rationele aanpak is.

Ook perfectionisten en mensen met een laag relativiseringsvermogen zijn gebaat bij deze aanpak.

Analyse van meetinstrumenten

Belbin teamroltest

Deze test kan op twee manieren afgenomen worden: ofwel via de Interplace Software (in België mogelijk in Nederlands, Frans en Engels), ofwel via een combinatie van psychometrische tests waarbij nog een accurater beeld kan worden neergezet (o.m. de 16PF van Catell, de DCT of Distorted Concepts Test en de PPQ ofte Personal Preferences Questionnaire)

Interplace is een testinstrument (ook op het internet beschikbaar: www.belbin.com) dat de "natuurlijke rollen" van mensen in **groepswork** meet (dus niet het individueel functioneren).

Over de validiteit en betrouwbaarheid bestaat lichte controverse. Een aantal gebruikte instrumenten scoren hier minder goed - zoals de 16PF van Catell die veel minder goed is om de BIG 5 accuraat in kaart te brengen dan de NEO-PI-R van Costa & McCrae. Maar het concept en het instrument werd gedurende meer dan 9 jaar onderzoek (betaald door de Britse staat) geteste en vertoont een zeer sterk punt: de ongelooflijke hoge voorspellende waarde van welke teams hoge prestaties gaan leveren. In totaal werden 120 teams getest met een business game. In de laatste oefening in Henley voorspelden de onderzoekers de rangorde van winnaars en verliezers voor de acht deelnemende teams: hun voorspelling was bijna helemaal juist.

Validiteit en betrouwbaarheid:

- Eigen onderzoek gepubliceerd in "Management Teams, Why they succeed or fail" uit 1981
- **Furnham, A., Steele, H. & Pendleton, D.** (1993,a). A psychometric assessment of the Belbin Team Role SPI. Journal of Occupational and Organizational Psychology, 66, 245-257 (kritiek)
- **Furnham, A., Steele, H. & Pendleton, D.** (1993,b). A response to dr. Belbin's reply. Journal of Occupational and Organizational Psychology, 66, 261. (kritiek)
- **Dulewicz V.** (1995) "A validation of Belbin's team roles from 16PF and OPQ using bosses' ratings of competence", Journal of Occupational and Organizational Psychology, vol. 68, pp. 81-99.
- **Broucek, W.G. & Randell, G.** (1996). An assessment of the construct validity of the Belbin SPI and OA from the perspective of the five-factor model. Journal of Occupational and Organizational Psychology, 69, 389-405
- **Fisher, S.G., Hunter, T.A. and Macrosson, W.D.K.** (1998), "The structure of Belbin's team roles", Journal of Occupational and organizational Psychology, Vol. 71, pp. 283-8. De validiteit (interne consistentie coëfficiënt) zou .85 bedragen.
- **Senior, B.** (1998). An empirically-based assessment of Belbin's team roles. Human Resource Management Journal, 8,3,54-60
- **Senior, B. & Swailes, S.** (1998). A comparison of the Belbin SPI and OAS as measures of an individual's team roles. International Journal of Selection and Assessment, 1, 1-8

- **Fisher, S.G., Hunter, T.A. & Macrosson, W.D.K.** (2001). A validation study of Belbin's team roles. *European Journal of Work and Organizational Psychology*, 10,2,121-144
- **Aritzeta, A., Swailes, S. & McIntyre-Bhaty, T.** (2004). Further evidence on the validity of the Belbin Team Role SPI and the OAS.
- **Onafhankelijk onderzoek** in Nederland:
 - Van Dongen, Marieke, Selectie van teamleden, Gids voor personeelsmanagement, juli 1994
 - Schenning, Janneke & Van der Sluis, Marianne, 1997, pre-post-then. Onderzoek naar de effecten van een teamrol training. Ongepubliceerd doctoraalonderzoek. RU Utrecht, Psychologie van Arbeid.
 - Buwalda S. & van den Hout E. (maart 2003) in "Psychometrisch onderzoek naar de Belbin Teamrolinstrumenten", werkstuk voor de afstudeerrichting arbeids- en organisatiepsychologie onder begeleiding van dr. Dirk van Dierendonck en in samenwerking met drs. Rob Groen.

Deze test is geen typologie zoals bijvoorbeeld MBTI, want de test plaatst mensen NIET in één vak maar geeft aan in welke (meerdere!) teamrollen je jezelf het beste voelt. De meeste mensen hebben twee à drie teamrollen die hem/haar van nature goed liggen (sterktes, talenten...) en nog een aantal teamrollen die zij vrij gemakkelijk kunnen opnemen.

Een bijkomend voordeel is dat de kandidaat niet alleen naar zichzelf kijkt, maar dat de kandidaat dankzij het **360° gebruik** te weten komt hoe anderen zijn gedrag percipiëren en interpreteren. Een kandidaat kan namelijk best een test invullen vanuit de "innerlijke beleving" (hoe de kandidaat "is") maar ander gedrag stellen (hoe hij zich "gedraagt" en "overkomt").

Persoonlijkheidstrekken verhogen wel de kans op een bepaalde "range" aan gedragingen maar voorspellen geen rechtstreeks specifiek gedrag.

Toch dienen we te eindigen met enkele kritische noten

- ▶ Niet alle onderzoekartikels zijn positief over de gebruikte tests en hun betrouwbaarheid of validiteit. Dit in tegenstelling tot bijvoorbeeld instrumenten als de N-IAS en de NEO-PI-R waarvoor de peer-reviewed artikels in belangrijke publicaties quasi uitsluitend unaniem zijn. Furnham bijvoorbeeld levert kritiek. Ook al omdat de 16PF van Catell minder accuraat blijkt om de Big Five te meten dan de NEO-PI-R.
- ▶ Er bestaat niet zo veel onderzoek naar deze methode.
- ▶ Belbin UK doet geen moeite om de instrumenten te verbeteren of bijvoorbeeld de 16PF van Catell te vervangen door de NEO-PI-R. De commerciële uitbating blijkt goed te lopen en de zoon van professor Belbin ziet het nut er niet van in (in Nederland worden er wel pogingen ter verbetering gedaan).
- ▶ Het meetinstrument gebruikt ipsatieve vragenlijsten die kort gezegd een gedwongen keuze doen maken, zelfs al is men het met geen enkele stelling eens. Op die manier bekomt men al snel een meer extreem type⁵. Gelukkig betekenen de ipsatieve

⁵ ondermeer:

- Tenopyr, M.L., "Artifactual reliability of forced-choice scales", *Journal of applied psychology* (1988), 73, 749-751.
- Baron, H, "Strengths and Limitations of Ipsative Measurement, *Journal of Occupational and Organizational psychology* (1996), 69, 49-56

vragenlijsten bij Belbin uit een scoring van 10 zinnen waardoor het geforceerde effect van ipsatieve vragenlijsten sterk wordt verminderd. De meeste ipsatieve vragenlijsten werken namelijk met gedwongen keuzes tussen twee vragen. Bovendien vervalt de Belbin methode gelukkig ook niet in “één typologie” (in tegenstelling tot bijvoorbeeld type-instrumenten zoals MBTI – zie verder) maar maakt het een combinatie van teamrol-voorkeuren.

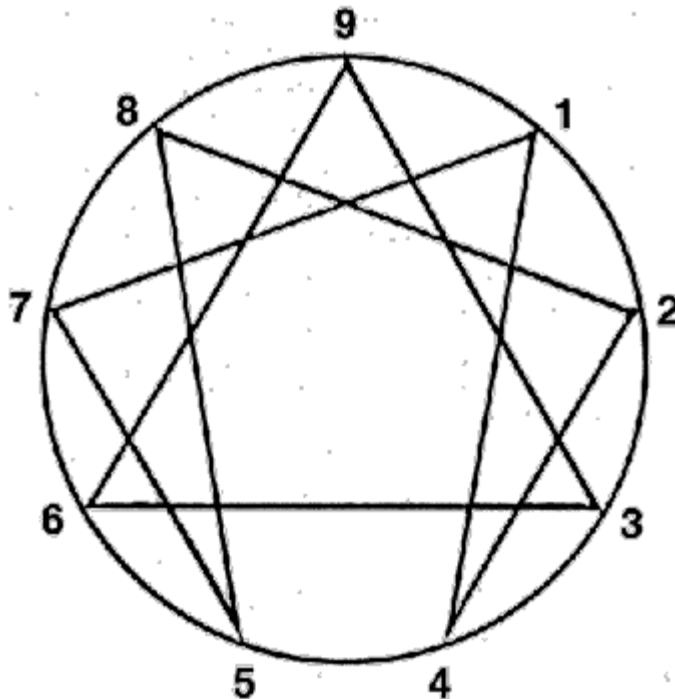
Conclusie:

De juistheid van de meting kan dus beter, hoewel de meting minder controversieel is dan bijvoorbeeld MBTI, waar enkel onderzoek door aanhangers positieve resultaten oplevert. Het instrument is nuttig als men de persoonlijkheid niet anders kan in kaart brengen (bijvoorbeeld door een combinatie van IQ-test met NEO-PI-R en N-IAS test) of wanneer het opzet specifiek naar teams gericht is.

Performance Coaching besloot in de context van trainingen tot een beperkt gebruik van het instrument (als bewustmaking van teamrollen) en zeker ook alleen waarvoor het bedoeld is: zijn voorkeuren in een team leren kennen.

-
- Meade, A.W; “Psychometric problems and issues involved with Creating and using ipsative measures for selection?” *Journal of Occupational and Organizational Psychology* (2004), 77, 531–552 *The British Psychological Society*

Enneagram



“Het enneagram is een geometrische figuur in de vorm van een cirkel waarin lijnen zijn getrokken tussen negen punten die zich op gelijke afstanden van elkaar op de cirkelomtrek bevinden (het Griekse woord ennea betekent negen).”

Het enneagramsymbool is afkomstig van George Gurdjieff (±1866 -1949) uit Armenië. Als we een blik werpen op enkele van zijn ideeën, stelt dit het enneagram eens in zijn oorspronkelijk “esoterisch” daglicht: Zijn visie op de mensheid was nogal pessimistisch. Hij beschouwde mensen als machines of automaten die louter reageren op externe prikkels. Ze beschikken niet over een bewustzijn en hebben geen controle over hun gevoelens, gedachten en gedrag.

Een letterlijk citaat: *“Vooral de maan heeft de mensheid in haar greep. Zij is een wezen dat zich voedt met de levensenergie die bij de dood van een organisme vrijkomt. Zielen van overledenen worden door de maan aangetrokken en stellen haar in staat zich te ontwikkelen tot een soort aarde, waarbij de temperatuur geleidelijk zal stijgen.”*

Er zijn al meer dan 10 boeken over het enneagram in de Nederlandse taal verschenen, en allen hebben ze zonder uitzondering een zeer wollig taalgebruik dat soms zelfs onbegrijpelijk of minstens voor meerdere interpretaties vatbaar is. Dit maakt het voor de lezers gemakkelijker om iets van waarheid te herkennen in de woordenbrij. Het barnum-effect dus. (In de wetenschappelijke psychologie wordt de neiging van mensen om zichzelf te herkennen in een brede algemene persoonsomschrijving ook wel aangeduid als het ‘Barnum-effect’ – genoemd naar de “grootste circusman aller tijden P.T. Barnum”, die ook wel bekend staat als een meesterlijke psychologische manipulator.)

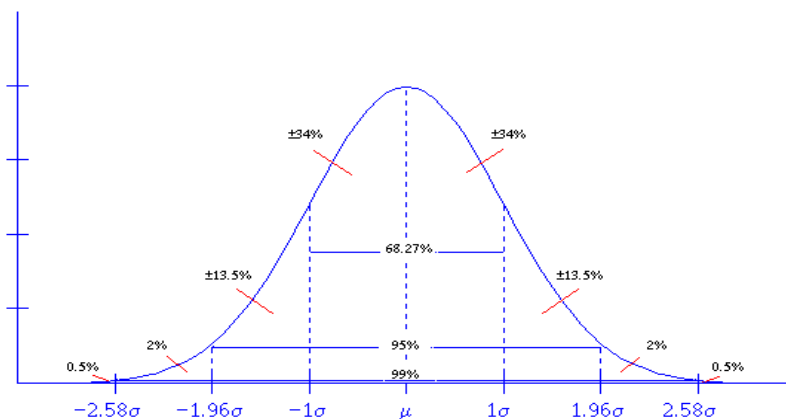
Nederlands hoogleraar in psychologie professor Emeritus Dr. Wim Hofstee was in 2005 zo moedig te pogen de alarmklok te luiden. *“Het enneagram is geen serieus*

wetenschappelijk model. Het behoort tot het onoverzichtelijke en snel wisselende domein van de lekenpsychologie: een bloeiende handel. Een psycholoog die is aangesloten bij het Nederlands Instituut van Psychologen (de beroepsvereniging, red.) en die met het enneagram werkt, loopt het risico een klacht aan broek of rok te krijgen ex artikel III.3 van de Beroepscode, dat over deskundigheid gaat," waarschuwt de expert in de persoonlijkheidsleer. "Voor niet-psychologen gelden zulke beperkingen natuurlijk niet. Maar aangezien de enneagram-industrie zich beweegt op het terrein van de persoonlijkheid - waar wij een wetenschap en een professie voor hebben - moeten we het beestje bij zijn naam noemen: beunhazerij."

Vaak verdedigen verdedigers van het enneagram zich met Machiavellaanse redeneringen op een vraag naar de wenselijkheid van dergelijke pseudo-wetenschap: kort gezegd vinden zij dat het niet belangrijk of het enneagram juist is of juiste uitspraken doet, als je mensen maar aan het denken zet en of je er interessante werkvormen kan mee opzetten.

Het lijkt ons meer aangewezen om voor dergelijke doelen degelijke en wetenschappelijke en vooral accurate instrumenten te gebruiken. Mensen worden al genoeg misleid in deze tijden. En als je over mensen uitspraken doet, dan zou dit mogen (en waarom niet "moeten") zo accuraat mogelijk zijn.

Als we kijken naar meer wetenschappelijke methoden dan toont de huidige stand van de wetenschap aan dat de Big Five momenteel het beste de persoonlijkheid in kaart brengen. De meest erkende test, namelijk de NEO-PI-R, telt voor elk domein van de Big Five 6 facetten. Vijf domeinen x 6 geeft in totaal 30 facetten. 30 facetten die op hun beurt onderlinge combinaties vormen, geven dus veel meer dan 4, 8, 9 of 16 persoonlijkheidstypes zoals in meer populistische benaderingen het geval is. Of zoals wij stellen: een type bestaat niet! Hoogstens kunt u op een bepaald facet een duidelijke tendens op het continuüm (vb introversie – extraversie) hebben...



HBDI (Herrmann Brain Dominance Instrument)

Dit instrument claimt na te gaan wat iemands geprefereerde denkstijl is:

- Blauw: analyticus: logisch denken, analyse van feiten, verwerken van getallen
- Groen: organisator: denkt in planningen, organiseert de gegevens, maakt gedetailleerde overzichten
- Rood: communicator: benadert de omgeving interpersoonlijk, intuïtief, expressief
- Geel: creator: gebruikt verbeeldingskracht, denkt in grote lijnen, conceptualiseert

Een eerste opvallende vaststelling (december 2004) is de afwezigheid van kritiek van gerenommeerde groeperingen van wetenschappers; dit kan erop duiden dat ofwel de theorie onjuist werd bevonden, ofwel dat de gegevens van de testresultaten niet ter beschikking worden gesteld van onafhankelijke onderzoekers. In de wereldwijde databases van medici en psychologen vonden we dit:

- Psychlit: één artikel waarin de tests HBDI en MBTI met elkaar vergeleken worden (zie verder)
- Medline: geen enkel artikel

Wij vroegen dus aan een aantal wetenschappers om één en ander na te gaan en deze kwamen tot **onthutsende vaststellingen**:

- **De hersenwerking** zoals uitgelegd op de website is **onjuist**:
 - HBDI stelt dat niet alleen de neocortex bestaat uit een linker- en rechterdeel, maar ook het limbische systeem. Hieruit leidt HBDI af dat deze 4 delen andere functies hebben en zelfs een denkvoorkeur impliceren (dit wordt verder nergens bewezen, ook niet door de zogenaamde neuro-psychologen waarmee hij samenwerkte). Over de dichotomie van het brein (links- rechts specialisaties) bestaan (ook nog bij dokters) veel vooroordelen en misverstanden: het is echter een feit dat de **huidige stand van de wetenschap** aantoonde dat het brein in hoge mate plastisch is (functies en specialisaties kunnen zich verplaatsen, hersenen veranderen voortdurend tot bij seniliteit) en dat er wel hoog gespecialiseerde **gebieden** bestaan.

Bij het verwerken van nummerwoorden (zoals “één”, “twee”...) werd wel aangetoond dat dit vooral in de linkerhersen helft plaatsvond, maar dat zowel rechts als links actief waren bij het verwerken van Arabische cijfers (“1”, “2”...). Bij het analyseren van processen zoals lezen werd vastgesteld dat subregioenen in beide hemisferen actief waren. Zelfs ruimtelijk inzicht activeerde beide hemisferen, hoewel de activering op een verschillende manier verliep. Links is beter in ruimtelijke categorisering (boven, onder, links, rechts...) en rechts is beter in afstanden verwerken (meters); maar zelfs hier heeft neuroimaging aangetoond dat de twee hemisferen samenwerken. Over de specialisaties: de frontale lobben dienen voor planning en actie. De temporale lob (zijkant) staat in voor geluidswaarneming, geheugen en herkenning van objecten. De pariëtale lob (bovenaan) dient voor gewaarwording en ruimtelijke ordening. De occipitale lob (achterhoofd) dient voor het zicht. Dit is een ruwe indeling, want de lobben zijn verder onderverdeeld in neurale netwerken die voor specifieke informatie-uitwisselingsprocessen dienen. **Elke complexe vaardigheid**, zoals optellen, het herkennen van woorden... **is het gevolg van gecoördineerde actie van verschillende van deze gespecialiseerde neurale netwerken in de verschillende delen van ons brein.** PET en fMRI kunnen wel de regioenen in het brein tonen die aangesproken worden tijdens lezen of rekenen, maar ze

kunnen niet de dynamische interacties van de mentale processen in kaart brengen.


- HBDI stelt dat de cerebrale hemisferen verbonden zijn door het corpus callosum en dat de twee helften van het limbische systeem bijeengehouden worden door de “Hippocampal Commissure”. Het limbische systeem is echter niet enkel een groepering van diverse organen (de olfactorische cortex, de amygdala, de hippocampus, de fornix...) die samen de mediale temporale lob (meer populair: “het limbische systeem”) genoemd worden, de organen worden in tegenstelling tot wat hbdI beweert ook “bijeengehouden” door de anterior commissure, the posterior commissure, Ganser’s commissures enz... Van “twee helften” van het limbische systeem is dus geen sprake (amygdala bevindt zich bijvoorbeeld rechts, hypothalamus links).
- De neurobiologen, psychiaters en neuropsychologen die de hersenwerking onderzoeken werken niet met “metaforen” maar baseren hun bevindingen op anatomisch onderzoek en met medische apparatuur (waarnaar ook HBDI verwijst: zie elders)
- U kunt zelf de bouw van de hersenen zien op “the whole brain atlas” <http://www.med.harvard.edu/AANLIB/cases/caseNA/pb9.htm>

■ De **testvragen zelf zijn niet goed opgevat**:

De meeste onderdelen uit de vragenlijst worden zeer beperkt bevraagd, onvoldoende om een juist beeld te kunnen vormen;

- Echte vragenlijsten over “handedness” of handvoorkeur (links- of rechtshandig)nemen veel meer items op dan de 2 vragen die HBDI opneemt (ondermeer potlood versus vork, gooien, slaan, oogdominantie...)
- Voor het onderscheid te maken tussen introversie en extraversie wordt slechts met één vraag gewerkt!
- Enzovoort

■ De beweringen over **nature/nurture verhoudingen** zijn fout:

- HBDI stelt “we feel that on average nurture represents 70% of who we are compared to 30% for nature”. Dit is totaal onjuist; recent onderzoek toont juist aan dat voor de meeste gedragingen biologie (genetica om juist te zijn) minstens 50% van de variatie in persoonlijkheid kan verklaren, voor sommige persoonlijkheidstrekken loopt dit zelfs op tot 80%. Enkele voorbeelden:
 - IQ is erfelijk in relatie tot de leeftijd van de onderzochte personen (vb op de leeftijd van 18 jaar speelt erfelijkheid 82% rol, op 28 jaar is dit reeds 85%, op 50 jaar 80%.ⁱ Naarmate kinderen opgroeien, wordt de eigen erfelijke aanleg steeds belangrijker: iemands huidig IQ is een goede voorspeller van zijn latere 
 - Ziektes zoals schizofrenie en depressie hebben een duidelijke erfelijke component (bepaalde vormen meer dan andere).
 - Op vrijwel elke leeftijd blijft erfelijke aanleg de belangrijkste determinant van gedragsproblemen:
 - Externaliserende gedragsproblemen (agressief, tegendraads, hyperactief): op jonge leeftijd (3 jaar) is de invloed van erfelijke aanleg 76%, daarna neemt de erfelijke invloed af tot ongeveer 50% op volwassen leeftijd (48% op 26-jarige leeftijd)ⁱⁱ

- Internaliserende gedragsproblemen (angstig, depressief, teruggetrokken) was op 3-jarige leeftijd 70% erfelijk bepaald, op de leeftijd van 26 jaar nog 60%.

■ Er staan **leugens, onjuistheden en oncontroleerbare gegevens** op de website die erg misleidend (dus zelfs strafbaar) zijn:

- Roger Sperry, Robert Ornstein, Henry Mintzberg... worden aangehaald als “Key Contributors to the whole brain technology”. Dit terwijl HBDI elders naar deze mensen verwijst als hun inspiratiebron en zegt dat de HBDI een “metaforisch model” is. Inderdaad: een metafoor, en geen wetenschappelijk instrument. E bestaat geen enkele publicatie van de eerder aangehaalde wetenschappers over de HBDI. U kunt bijvoorbeeld surfen naar <http://www.rogersperry.info/> waarop u geen enkele link met Herrmann of HBDI zal vinden!

Hierna volgt een uittreksel uit de mail die de zaakwaarnemer van Dr. Sperry mij toezond:

“Dear Mr. Vermeren:

Thank you for your e-mail. .../... I am not familiar with Ned Hermann or the HBDI. I do not believe that Dr. Roger Sperry was aware of either the individual nor of the instrument.

My impression, having reviewed the website, is that this system is a business consulting one based very loosely on recent theories of brain functioning, personality typology, and industrial-organizational psychological principles. I will forward this to Mrs. Sperry to determine whether my answers are valid.

Thank you,

Antonio E. Puente”

En op 30/12/2004:

“Have now heard from Mrs. Sperry and Dr. Bogen;

Mrs. Sperry- Roger nor her had/have ever heard of the concept nor of the person

Dr. Bogen- “Ho, Ho, Ho” (he thought it was a practical joke on my part)

Oh, well...Happy New Year.

Tony Puente”

En hier het kort maar krachtige antwoord van Henry Mintzberg:

Dear Mr. Vermeren

I certainly never worked with Herrmann nor do I recall citing his work, and I did no research as such on the brain, just an article using Sperry and Ornstein's work called "Planning on the Right Side and Managing on the Right."

Thank you for this.

Henry Mintzberg

- Eerst is er sprake van “more than 60 doctoral dissertations”, later is er sprake van “over 50 doctoral dissertations”. Omdat deze traceerbaar zijn werden deze opgezocht en er is slecht één doctorale dissertatie⁶ terug te vinden. Zelfs deze ene haalde de publicatie in een gereputeerd vakblad niet. Om te kunnen verschijnen moet immers bewezen worden dat er geen fouten in het onderzoek zitten en dergelijke meer: wetenschappelijke publicaties zijn altijd onderhevig aan peer review. Bovendien blijkt deze éne doctorale dissertatie niets meer te zijn dan een vergelijking tussen HBDI en MBTI die bovendien verscheen in het “Journal of psychological type”, een puur promotieblad voor ... MBTI.

⁶ Power, Sally J. and Lundsten, Lorman, “Studies that compare type theory and left-brain-right-brain theory”, *Journal of Psychological Type*, Vol. 43, 1997. Power, SallyJ, Kummerow, Jean M. and Lundsten, Lorman, “A Herrmann Brain Dominance profile analysis of the sixteen MBTI types in a sample of MBA Students, *Journal of Psychological Type*, Vol. 79, 1999.

- Er zijn meerdere indicaties dat het gaat om **pseudo-wetenschap**:
 - Er wordt verwezen naar een onderzoek waarin men probeert de link te leggen met astrologie (geen astronomie): de pseudo-wetenschappelijke leer waarbij de namen die vroegere volkeren gaven aan sterrenconstellaties nu allerlei verklaringen opleveren gaande van de invloed op uw persoonlijkheid tot verklaringen over uw levensloop en dood. Wat betreft de astrologie: de namen van de sterrenbeelden zijn eerder toevallig gegeven en verschillen bovendien van cultuur tot cultuur. De astronomen (de wetenschappers die de ruimte bestuderen) zouden het zich geen moer aantrekken als de “Waterman” de “Koalabeer” zou heten maar de astrologen zouden totaal andere uitspraken doen. Hun “wetenschap” hangt immers af van die namen en niet van natuurverschijnselen. Maar ondanks duizenden experimenten die wél deugdelijk werden uitgevoerd en die aantonen dat astrologie klinkklare onzin is, blijft dit een enorme aantrekkingskracht hebben op mensen.
 - “Meer dan 100 artikels” (= bewering op website) zijn **ENKEL verschenen in “the international brain dominance review”**, een eigen tijdschrift...
 - Er is **geen enkel artikel** verschenen in peer-reviewed tijdschriften.
 - Er wordt verwezen naar “bioritme” als een verklaring voor persoonlijkheid.

De conclusie van de geraadpleegde mensen is ongemeen hard: “dit is pure kwakzalverij”. Dat mensen zichzelf herkennen zal dan wel te maken hebben met het Barnum-effect (zie achteraan voor uitleg) en het feit dat ze de test over zichzelf invulden, maar dat geldt evenzeer voor de horoscoop in uw favoriete tijdschrift...

Tenslotte wijzen wetenschappers erop dat het in diverse landen strafbaar is medische uitspraken te doen (ook in ons land) of psychologische tests af te nemen als niet-psycholoog.

Wetenschappers die reageerden op hbd:

- Rosemary Lyndall, Clinical neuro-psychologist, Perth, Western Australia
- John M. Price, Ph.D, Psychology Department, Davis, University of California
- Danny Rouckhout, licentiaat psychologische wetenschappen en docent psychometrie (en ook inductieve statistiek) aan de Lessius Hogeschool te Antwerpen, maar daarnaast ook verbonden aan de dienst psychiatrie van het UZA Wilrijk

LIFO (Life Orientations)

Lifo is zowel een methode als een test (vragenlijst volgens de auteur) die wel degelijk pretendeert dingen te meten.

Opvallend is de afwezigheid van onafhankelijk onderzoek en literatuur over LIFO, hoewel de vertegenwoordigers beweren over een database met miljoenen mensen te beschikken, wat meteen ook zeer twijfelachtig kan worden gevonden. In PsycInfo, Ebsco Host of Science Direct (academische databases) was niets te vinden.

Er bestaan (op de website en op eenvoudig verzoek te bekomen cijfergegevens) van de vertegenwoordigers zelf en deze legden we voor aan specialisten in psychometrie en statistiek. Hierna de antwoorden:

- **Argumentatie van LIFO voor betrouwbaarheid en validiteit:** enerzijds wordt veel moeite gedaan om de terminologie en methodiek van betrouwbaarheid toe te lichten (en de methode op die manier een wetenschappelijk imago te geven??), maar anderzijds moet je dan lezen dat het eigenlijk niet zoveel uitmaakt omdat de LIFO survey geen test is...en dat je de cijfers van betrouwbaarheid niet te strikt moet interpreteren. Van het ogenblik dat je een meting doet, moet die meting beantwoorden aan elementaire eisen van betrouwbaarheid en validiteit. Dus ook de LIFO.
 - De **interne consistentie** (alpha) is gewoon zwak te noemen: met uitzondering van een aanvaardbaar cijfer van 0.70 liggen alle waarden laag. Over de reden daarvan wordt niets gezegd. Alleen probeert men de pil te verzachten door op te merken dat de betrouwbaarheid niet te strikt mag worden genomen omdat de waarde afhangt van o.a. de opzet van het onderzoek enz. Dat is juist, interne consistentie cijfers worden niet enkel bepaald door de inhoud van de vragen, maar ook door de samenstelling van de proefgroep waarop de berekening gebeurde. Daarom ziet men dat bij ernstige testontwikkeling de samenstelling van de steekproef ten eerste goed wordt omschreven en dat er replicaties volgen op verschillende steekproeven. Niets te merken van dergelijke replicaties bij de LIFO, verwonderlijk omdat men er nogal trots op is dat de methode "reeds bij twee miljoen mensen werd aangewend en voortdurend wordt verbeterd". Een tweede betrouwbaarheidsonderzoek is dan toch niet veel gevraagd.
 - De **stabiliteit** over een periode van vijf weken (test-hertest) wordt ook onderzocht. Op een steekproef van amper 63 studenten. Terug zeer lage cijfers. Daarover wordt gezegd dat dit mogelijks een gevolg kan zijn van het feit dat de respondenten ondertussen een andere visie (voorkeur) ontwikkelden. Dat is één mogelijkheid. Men vergeet er wel bij te zeggen dat het ook een gevolg kan zijn van een slechte stabiliteit. Hoe dan ook, de lage cijfers zijn niet te interpreteren: ofwel is het de weergave van een stabiele test afgenomen bij personen wiens voorkeur veranderde, ofwel is het de weergave van een instabiele test bij niet-veranderde personen, ofwel is het de weergave van een combinatie van de twee. Dus, zinloze cijfers. Elke verandering kan spontaan zijn en hoeft geen gevolg van je training te zijn. Meer technisch: dergelijke lage

stabiliteiten geven ook zeer grote standaardmeetfouten van verschil (standard error of difference). Dat wil zeggen dat er een zeer grote score toename of afname zal nodig zijn om te spreken van werkelijke verandering. Ondanks alles meent men toch te mogen zeggen...the LIFO survey measures pretty much the same thing in people over time...(p.4). Je moet maar durven.

- **Betrouwbaarheid:** er staan een reeks van pamfletachtige opmerkingen in de brochure/op de website die niets te maken hebben met betrouwbaarheid, maar in het beste geval met validiteit. De kwaliteit van de vertaling, het brede gebruik (25 landen), de massa respondenten (twee miljoen) zijn irrelevante gegevens en dragen zeker niet bij tot de betrouwbaarheid. In de wetenschap gaat men van psychometrische tests uit van een betrouwbaarheid van 0.70 die moet behaald worden. Bij LIFO is men reeds tevreden met een maat van 0.4 tot 0.60. Gesteld dat we dit aannemen wil dit ook zeggen dat we te maken hebben met een grote meetfout en dus met een zeer breed betrouwbaarheidsinterval. Kortom, de concrete score van een persoon is zeer relatief en kan dus evengoed stukken lager of hoger liggen. Om een concreet idee te geven wat dit kan betekenen passen we dit even toe op een IQ-meting met deze verschillende betrouwbaarheden.
 - Stel, iemand heeft een IQ score van 100 op een test. De test heeft een betrouwbaarheid van 0.90. Dan ligt de ware score (met 95% zekerheid) ergens tussen 90 en 109 IQ.
 - Nemen we nu de betrouwbaarheidcijfers die LIFO zelf publiceert: Bij een betrouwbaarheid van 0.60 zou de echte score ergens liggen tussen 81 en 119 IQ. Als je weet dat 80 de grens van mentale retardatie is en dat sommigen vanaf 125 IQ over hoogbegaafdheid durven spreken, klinkt dit nogal idioot. Bij een betrouwbaarheid van 0.40 zou je tegen de man of vrouw in kwestie moeten vertellen dat het IQ ergens tussen 77 en 123 ligt, dus de persoon in kwestie kan mentaal geretardeerd zijn of ook wel hoogbegaafd.

- Over **de validiteit:** Gezien de lage betrouwbaarheidcijfers zal de validiteit van de LIFO Survey noodzakelijkerwijze ook laag zijn. Betrouwbaarheid is immers de bovengrens van validiteit. Een validiteitscijfer kan, theoretisch gezien, nooit boven de betrouwbaarheidscijfers uitkomen. Men mag zich dus verwachten aan validiteitscijfers in de buurt van 0.20 en 0.30.

- **Over de theorie** waarop LIFO gebaseerd is: men haalt Erich Fromm en Carl Rogers aan en zegt ervan dat zij zorgden voor “soundly based psychological theories”; deze theorieën zijn jammer genoeg wel degelijk achterhaald. Het is duidelijk dat zij worden gevolgd door “gelovigen” of “adepten”, serieuze wetenschappers doen dit niet.

Niet alleen blijkt de **test** niet meer of niet minder dan **waardeloos** te zijn, ook de theorie en het model zijn zeer mager, en geven blijk van weinig inzicht in leidinggeven, om maar één voorbeeldomein aan te halen.

MBTI – Myers-Briggs Type Indicator®

Dit is zowat de meest verspreide commerciële test om het “type” van mensen te meten.

MBTI® werd reeds in het begin van de vorige eeuw ontwikkeld door Isabel Briggs Myers (1897 – 1979), waarbij ze het “werk” van haar moeder, Katharine Cook Briggs verderzette. Noch Katharine noch Isabel waren opgeleid in de psychologie. Isabel was ‘bachelor’ in politieke wetenschappen. Katherine maakte Isabel attent op het boek “*Psychological Types*” van Carl Gustav Jung. Ze werden samen wat ze zelf noemen “fervente type observeerders”. Isabel leerde naar eigen zeggen “*test construction*” door de “*personeelstesten van een plaatselijke bank te bestuderen*”.

In tegenstelling tot Interplace (Belbin – enkel functioneren in groep) beweert MBTI zowel nuttig te zijn voor individueel functioneren als voor functioneren in groep. MBTI is verder “een test om te kijken **hoe mensen naar zichzelf kijken**”.

Er zijn een aantal merkwaardige vaststellingen te doen die aanleiding geven tot enkele kritische bedenkingen:

1. PROBLEMEN MET DE TEST ZELF

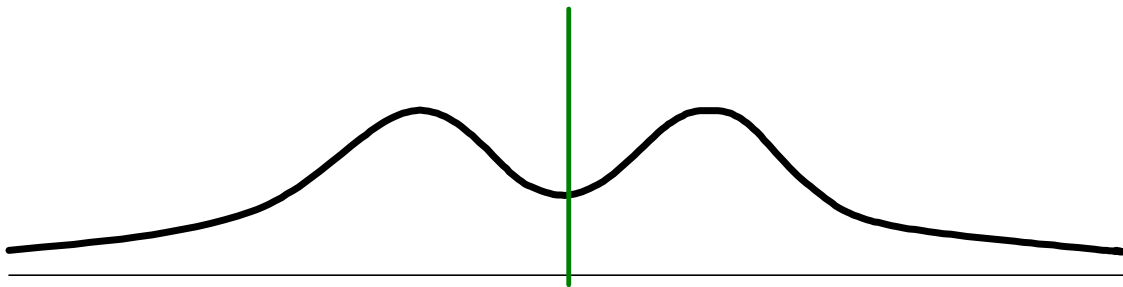
- Alert Management Consultants - officiële distributeurs van de MBTI voor de Benelux - prijzen het instrument aan als 'uiterst betrouwbaar' en 'door honderden onderzoeken tijdens de laatste 50 jaar gevalideerd', maar hierop valt het een en ander af te dingen. Het is waar dat er in Amerika veel onderzoek rond de MBTI is verricht. De resultaten zijn echter vaak gemengd of onduidelijk, en bovendien raden onafhankelijke onderzoekers het gebruik van het instrument af! (zie bijv. het overzicht van Fitzgerald en Kirby, 1997). Onafhankelijk en peer-reviewed onderzoek laat namelijk absoluut niet dergelijke fraaie validiteit en betrouwbaarheid zien als de vele artikels die gepubliceerd worden in ... the Journal of Psychological Type. Dit is namelijk een tijdschrift geheel en al gewijd aan MBTI en helemaal niet onderworpen aan peer-review.
- Reeds in 1957 werd de test geweigerd door the Educational Testing Service (ETS) bij hun uitgebreid en duur onderzoek naar de wenselijkheid MBTI op te nemen in hun gamma. Na een negatieve interne studie, kozen ze om MBTI niet verder op te nemen in hun gamma, en verbraken ze hun relatie met Myers. Consulting Psychologists Press verworven in 1975 de rechten voor de distributie van de test.
- In een onderzoek bij 1.291 mensen⁷ werden de vragen en schalen zelf onderzocht. Factoranalyse om uit te maken of de 4 bipolaire dimensies van de MBTI werkelijk bestaan, toonde aan dat er op basis van de vragen niet 4 maar 6 dimensies waren. Bovendien vonden de onderzoekers een grote meetfout. 83% van de verschillen tussen de deelnemers konden niet worden verklaard door de MBTI.
- In andere studies stelde men bovendien vast dat de JP en de SN schalen sterk met elkaar correleerden, wat dus niet mag (ze meten quasi hetzelfde).⁸ Ook andere

⁷ Sippes, G.J. en Alexander R.A. en Friedt, L.: “Ieme analysis of the Myers-Briggs Type Indicator”, Educational and psychological Measurement, vol. 45, n° 4 (1985) blz 789-796

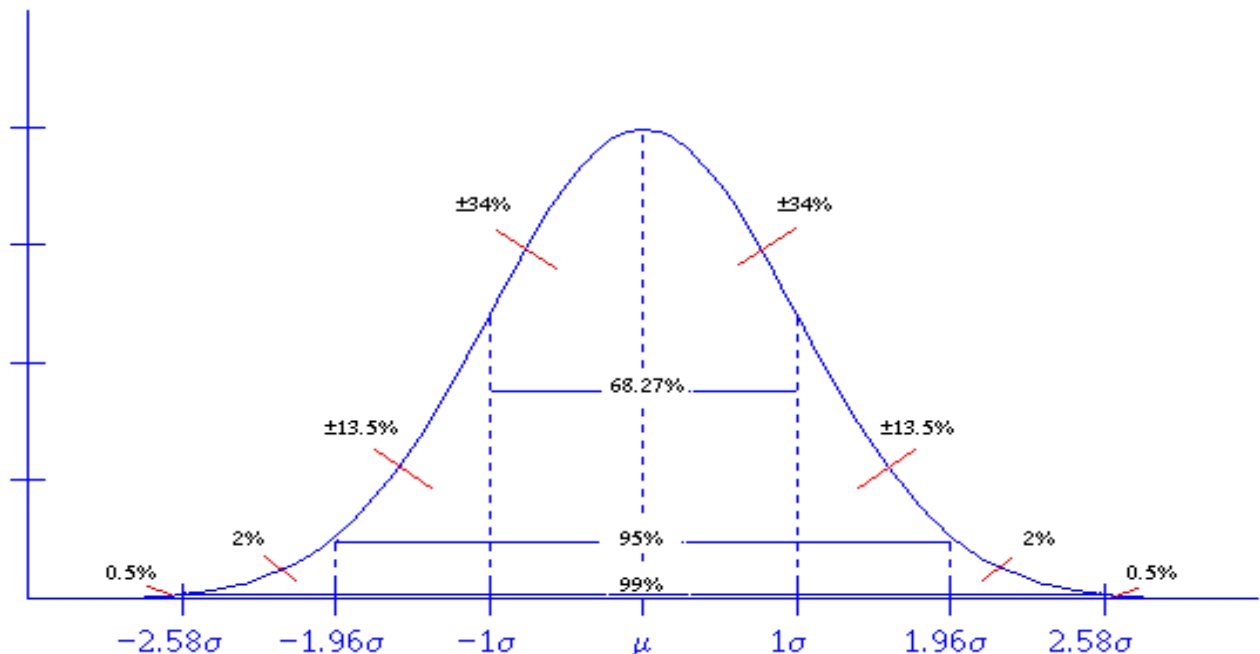
⁸ McCrae, R.R en Costa P.T.: “Reinterpreting the Myers-Briggs Type Indicator from the Perspective of the Five-Factor Model of Personality.”, Journal of Personality, vol. 57, N° 1 (1989), blz 12-40

schalen correleren sterk (Berr, Chruuch & Waclawski in 2000, Isabel Myers (!) et al. in 1998; Sipps & DiCaudo in 1988).

- De test bevat een kapitale fout, namelijk dat men er van uitgaat dat de schalen “discontinu” of “bimodaal” zijn (figuur 1) terwijl alle wetenschappelijk onderzoek⁹ uitwijst dat trekken verdeeld zijn over een continue schaal (de Gausscurve in figuur 2). Het gros van de bevolking zit dus in de middenzone. Nemen wij bijvoorbeeld de uitgesproken extraverten, dan zitten die langs de ene kant van de extremititeit van de Gausscurve (figuur 2) terwijl de uitgesproken introverten langs de andere kant liggen. Nu wordt door een berekening iedereen in een vak “geduwd”. Je bent ofwel extravert ofwel introvert, terwijl het gros van de bevolking noch extravert, noch introvert is, maar in balans.



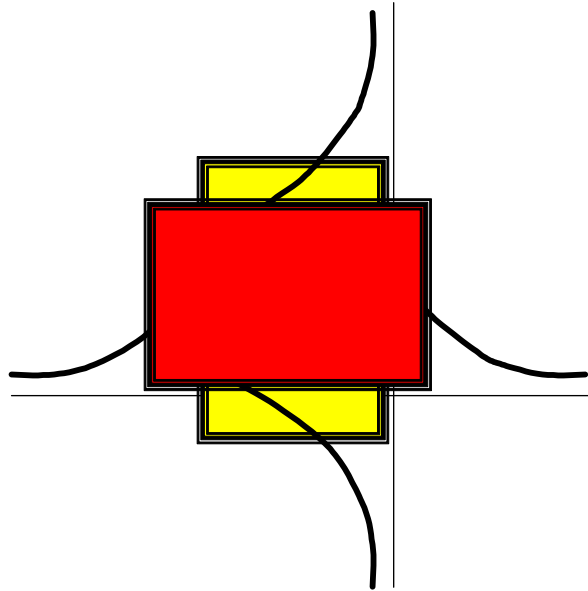
Figuur 1: het foutieve uitgangspunt van MBTI: elke dimensie (I/E, S/N, T/F, J/P) is weergegeven volgens een discontinue schaal met een “dal” in het midden.



Figuur 2: de werkelijke verdeling van de bevolking bij dimensies zoals Extraversie/Introversie, denken/voelen (Bron: D. Rouckhout)

⁹ Bijvoorbeeld Stricker and Ross (1962, 1964) in “An assessment of Some Sturctural Properties of the Jungian personality Typology.” *Journal of abnormal and social psychology*, Vol. 68 (1964), blz 62-71

Het probleem met MBTI is bovendien dat de fouten worden uitvergroot doordat verschillende schalen op elkaar telkens met dezelfde Gausskurve zitten. Voor alle schalen geldt dus dat het gros van de populatie in ELK van deze schalen in het midden zit. Het gros van de bevolking kan dus NIET in type worden onderverdeeld.



Figuur 3: indien men twee schalen op elkaar plaatst, blijft het gros van de bevolking nog steeds in het midden zitten. Dit blijkt ondermeer ook uit een groot longitudinaal onderzoek uitgevoerd door the National Institute on Aging (NIA – USA) waarin mensen om de twee jaar worden opnieuw getest op medisch, psychologisch en ander vlak.

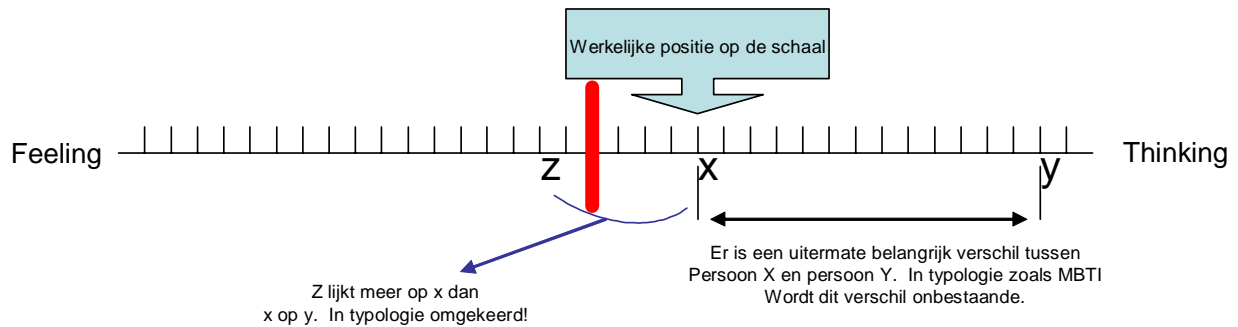
Cohen (1983) toonde aan dat door een schaal dichotoom te maken je de “gedeelde variantie” met een andere variabele met .64 (r^2) vermindert. Bij twee dichotome schalen die men met elkaar combineert loopt dit al op tot .40 (r^2). MBTI combineert zelfs 4 dichotome schalen, waardoor de voorspellende kracht haast tot nul wordt gereduceerd.

Diverse andere onderzoekers¹⁰ waarschuwen voor het type schalen dat MBTI gebruikt (dichotome en ipsatieve – gedwongen keuze).

Een andere manier om naar de schalen te kijken is de volgende: bij een type wordt de scheiding in het midden gemaakt (rode lijn in figuur 4): je bent ofwel introvert ofwel extravert, ofwel thinking ofwel feeling... Dit terwijl je in werkelijkheid misschien lichtjes extravert bent maar niet erg extravert. Figuur 4 geeft dit het beste weer:

¹⁰ ondermeer:

- Tenopyr, M.L., “Artifactual reliability of forced-choice scales”, *Journal of applied psychology* (1988), 73, 749-751.
- Baron, H, “Strenghts and Limiations of Ipsative Measurement, *Journal of Occupational and Organizational psychology* (1996), 69, 49-56
- Meade, A.W; “Psychometric problems and issues involved with Creating and using ipsative measures for selection?” *Journal of Occupational and Organizational Psychology* (2004), 77, 531–552 *The British Psychological Society*



Figuur 4: het probleem met een gedwongen keuze tussen het ene uiterste en het andere uiterste. Zie ook het onderzoek van Vermeulen.

Dit is ook één van de redenen waarom je je zo **gemakkelijk herkent (90 tot 100% herkenning) in vier (!!!) van de 16 types** wanneer je de lettercombinaties samenstelt vooraleer je de gedwongen verdeling doet: je maakt combinaties waarbij je de lettertegenpolen waarop je score niet ver uiteenligt ook telkens combineert.

- In het **eigen officiële verslag** over de betrouwbaarheid en validiteit van het MBTI-instrument¹¹ staat te lezen: “Omdat persoonlijkheid “glad” is om te meten, kunnen psychologische instrumenten niet dezelfde consistentie vertonen die we kunnen verwachten van bijvoorbeeld een meetlat.” Maar er zijn algemeen aanvaarde standaards voor psychologische instrumenten... Het is belangrijk om te weten dat het MBTI® instrument tegemoetkomt aan die standaards en ze zelfs overschrijdt op vlak van test-hertest betrouwbaarheid.” .../... “Bij retest komen mensen uit op **drie tot vier voorkeuren** voor type in 75 tot 90% van de gevallen.” .../... “Er zijn een paar groepen waarbij de betrouwbaarheid erg laag is, en dus moet er met voorzichtigheid gehandeld worden bij het gebruik van het MBTI® instrument bij deze groepen (bijvoorbeeld bij kinderen; +sommige practitioners melden ons ook dat er soms “noise” is bij sommige groepen van respondenten).”

Bedenkingen van diverse auteurs:

- Met hoeveel worden de standaards voor test-hertest betrouwbaarheid overschreden? Dit is onduidelijk.
- **Waarom zijn 16 schalen nodig als iedereen zich kan terugvinden in 4 schalen?** Waarom wordt er meestal maar één 4-lettertype aan mensen voorgesteld als andere types zo nauw aansluiten? Het toekennen van slechts één 4-lettertype houdt serieuze **risico's op stereotypering** in.
- Een duidelijker aanwijzing (ook naar de buitenwereld zoals bedrijven die willen investeren in psychometrische instrumenten) over welke groepen best niet getest worden omwille van lage betrouwbaarheid, zou van pas kunnen komen.
- Robert I Winer, M.D., beschrijft dat MBTI een foutmarge van “**tenminste 25%**” heeft. Het instrument blijkt resultaten op te leveren die niet overeenstemmen met het bevestigde en/of geobserveerde type van de testpersonen. Dit wordt toegeschreven aan “**ongeschikte contexten**”. Zo zou, nog steeds volgens Winer, de betrouwbaarheid “**onder andere**” afnemen wanneer iemand naar een baan op zoek is: om te overleven gaan mensen liegen over zichzelf. Zelfbeoordelingen blijken

¹¹ Drs Gordon Lawrence and Charles Martin, hoofdstuk 7 in het boek “Building People, Building Programs”

accurater te worden naarmate de persoon daadwerkelijk juiste informatie over zichzelf zoekt.

- Vermeulen (1996) analyseerde de MBTI scores van in totaal 362 Nederlandstalige cliënten van een psychologisch adviesbureau. Uit zijn berekeningen blijkt dat de betrouwbaarheid van geen van de acht schalen volgens de normen van de COTAN (Commissie Testaangelegenheden van het Nederlands Instituut van Psychologen) 'goed' genoemd zou kunnen worden. De betrouwbaarheid komt nergens hoger dan .82 (bij de schalen Beheersing en Waarneming). **Twee schalen (Intuïtie en Gevoel) hebben een betrouwbaarheid lager dan .70 wat als onvoldoende moet worden gekenschetst.** De gegevens over de validiteit zijn evenmin om over naar huis te schrijven. De verwachte correlaties met andere tests komen in het overgrote gedeelte (20 van de 27) van de gevallen niet uit. "Deze cijfers geven geen aanleiding om de MBTI tot een valide instrument te bestempelen", concludeert Vermeulen. Mogelijk één van de belangrijkste psychometrische problemen bij de MBTI ligt in de dichotomisering en de discontinuïteit (zie hiervoor) bij de schaalduo's: een cliënt die score 11 heeft bij Extraversie en 10 bij Introversie behoort bij de Extraverte persoonlijkheidstypen, net als iemand bij wie de scoreverhouding 20-1 is. Bij de omzetting van de schaalscores tot de persoonlijkheidstypen treedt aldus een aanzienlijk informatieverlies op. Deze gang van zaken is eigenlijk alleen te verdedigen als de schaalscores bimodaal oftewel tweetoppig verdeeld zouden zijn. Hiervan is echter geen sprake. Ook de MBTI schalen zijn 'normaal' verdeeld (veel scores in het midden, weinig aan de extremen). Toch "dwingt men iedereen in een type".
- Howes, R.J. en T.G. Carskadon stelden reeds in een studie in 1979 (!) vast dat de test-hertest betrouwbaarheden ronduit zwak waren: na een test-hertest interval van slechts 5 weken hadden 50% van de mensen een ander type als uitkomst...In 1983 repliceerden Mc Carley en Carskadon deze resultaten. **Zelfs Isabel Myers¹²** et Al. zelf rapporteerden in 1998 dat 35% van de testpersonen hadden een andere type-score na een test-hertest interval van slechts 4 weken!
- Zulke zwakke test-hertest stabiliteit werd meermaals ontdekt, ondermeer door Feenor (Centre for Creative Leadership Greensboro, 2001), Mastrangelo (University of Baltimore, 2001), Bess, T.L. & Harvey R.J. (Journal of Personality Assessment, 2002)
- De bewering in de MBTI manual dat de "4 schalen samen een geheel vormen dat verschillend is dan de som van de delen" wordt nergens bewezen, en onafhankelijk onderzoek kon dit al evenmin bevestigen.

2. OVER DE THEORIE WAAROP DE TEST GEBASEERD IS

- **MBTI werd "uitsluitend" ontwikkeld op basis van het werk van Carl Jung. Heel wat (empirische) studies hebben ondertussen uitgewezen dat de persoonlijkheidstheorieën van Jung fout waren.¹³** Volgens het onderzoek is zijn

¹² Myers, I.b., Mc Caulley, M.H., Quenk, N.L., & Hammer, A.L.: "Manual: a guide to the development and use of the Myers-Briggs Type Indicator (1998), Palo Alto, Consulting Psychologist Press.

¹³ Kritiek valt te lezen bij:

- Don McGowan: "What is wrong with Jung?" (Amherst, N.Y. Prometheus Books, 1994)
- Richard Noll: "The Jung Cult: Origins of a Charismatic Movement" (Free Press, 1997)

beroemde karakterindeling (*Psychologische Typen*, 1921) niet alleen “een volstrekt onhandelbaar systeem door het kwartet denken-voelen-ervaren-intuïtie te combineren met de karakterindeling extravert-introvert die teruggaat op de beroemde psycholoog William James (1842-1910)”. Hij deed bovendien geen enkel statistisch onderzoek naar zijn “observaties”. Jung deed ook geen gecontroleerd onderzoek en hij baseerde zich teveel op zijn eigen “inzichten” en percepties.

- Over de component “intuïtie” zijn er veel bedenkingen te maken; tenslotte beweerde Jung dat er mensen zijn die “intuïtief” dingen aanvoelen, bijvoorbeeld dat er “een rode wagen van achter de straathoek zal komen”...Intuïtie wordt nu wetenschappelijk geduid als “het geheel aan ervaringen dat vorm geeft aan je impliciete kennis”. Onze voorhoofdskwabben “scannen” voortdurend de omgeving en zo kan het bijvoorbeeld zijn dat een brandweerman “intuïtief” de juiste keuzes maakt omdat zijn geheel aan ervaringen hem vertellen dat er iets fout is. De intuïtie zoals beschreven door Jung lijkt eerder op het paranormale.
- Op verschillende punten blijken de onderzoeksgegevens van de MBTI trouwens zelfs niet in overeenstemming met Jungs theorieën. Zo zouden mensen volgens Jung bij het ouder worden hun hulpfuncties (d.w.z. de niet dominante functies) in een 'individuatieproces' moeten ontwikkelen. De MBTI echter beschouwt de voorkeuren als bij de natuur van de persoon behorend, net als links- of rechtshandigheid. En inderdaad blijkt uit de onderzoeksgegevens met de MBTI niets van individuatieprocessen, althans volgens McCrae en Costa (p. 32), die concluderen: "...those who embrace Jung's theory should probably avoid the MBTI".

-
- Richard Noll: “The Aryan Christ: The Secret Life of Carl Jung” (Random House, 1997)
 - Anthony Storr: “Feet of Clay – saints, sinners, and madmen: a study of gurus.” (New York; The Free Press, 1996)
 - Marcel Hulspas: “Psychiater, Praatjesmaker, profeet – leven en werk van Carl Gustav Jung” (Skepter 10(3), september 1997)
 - Robert Todd Carroll: “The skeptic’s Dictionary: Myers-Briggs Type Indicator®” (www.SkepDic.com)
 - Jon Noring: “Personality Type Summary”: “personality models do not put people in boxes, people put people in boxes...”
 - Robyn M. Dawes: “Psychology and Psychotherapy Built on Myth” (New York: The Free Press, 1994)
 - McCrae, R., & Costa, P.T., Jr. (1986). Clinical assessment can benefit from recent advances in personality psychology. *American Psychologist*, 41, 1002.
 - McCrae, R. & Costa, P.T., Jr. (1989). Reinterpreting the Meyers-Briggs Type Indicator from the perspective of the Five-Factor Model of personality. *Journal of Personality*, 57, 17-40.
 - John Mickelwait en Adrian Woolridge, in “De Toverdokters” (Business Bibliotheek, Contact)
 - Alan Sokal en Jean Bricmont: “intellectueel bedrog” (EPO/De Geus)
 - Fitzgerald, C. en Kirby, L.K. *Developing Leaders: Research and Applications in Psychological Type and Leadership Development*. Davies-Black, Palo Alto, California, 1997.
 - Vermeulen, J. Een validatieonderzoek van de Nederlandse versie van de MBTI: Werkstuk onder supervisie van Dr. A. Evers. Universiteit van Amsterdam (faculteit psychologie), maart 1996.
 - Gustaaf Cornelis en Geerd Magiels : “onzin verkoopt: pseudowetenschappen en de bedrijfswereld (Wonder en is gheen wonder, 2° jaargang nr 4, 12/2002); “Pseudowetenschappelijke persoonlijkheidstypologieën zoals die van Myers-Briggs geven medewerkers een illusie van inzicht in en aandacht.”

- Briggs en Myers hebben zelf de dichotome schaal “Judgment vs Perception” uitgevonden. Dit was niet gebaseerd op Jung.
- Volgens verschillende onderzoekers, psychiaters en psychologen (zie voetnoot) zijn **de beschrijvingen van de types zo vaag en ruim dat je niet anders kan dan je er in grote mate in herkennen**. *“De byzantijnse complexiteit maakt dat elk soort van gedrag ultiem past in elk persoonlijkheidstype”* stelt Todd Caroll. Deze **neiging van mensen om zichzelf te herkennen in een brede algemene persoonsomschrijving ook wel aangeduid als het ‘Barnum-effect’ (zie achteraan in onze conclusies voor meer details hierover)**. De vaagheid van de beschrijvingen maakt dat in wetenschappelijke kringen over MBTI smalend wordt gepraat als **“een horoscoopmodel”**.
- Een andere reden voor de “lichtgelovigheid” in de resultaten is dat de test **in erg flatterende taal types beschrijft**. Wie leest over zichzelf dat hij “een innovatieve denker en goede probleemoplosser is, die goed mensen kan begrijpen en motiveren, maar misschien minder goed is in het gedetailleerd tot een goed einde brengen van projecten” zal dit graag geloven, of de uitspraak nu echt is of niet. Nog een gevolg van het Barnum-effect;
- In 1975 nam CPP (Consulting Psychologists Press, Inc) de test over van Isabel Briggs Meyers en begon dit commercieel uit te baten. Ook zij zijn erg “ruimdenkend”; *“als een type niet past probeer dan eens een letter te veranderen”* en *“als je jezelf niet herkent, dan is het belangrijk een goede feedbacksessie met de practitioner te organiseren”*.

CONCLUSIES:

Hoewel onderzoek door “aanhangers” aangeeft dat de validiteit en betrouwbaarheid aanvaardbaar zijn, zijn de conclusies in onderzoek door onafhankelijke wetenschappers zéér overwegend negatief.

- ▶ De psychometrische kwaliteiten zijn ronduit te zwak en uitspraken doen over mensen geeft een te grote kans op fouten (tot 50% - zie voorgaande uiteenzetting).

Over het feit dat de onderliggende theorie niet klopt, bestaat in de academische wereld wereldwijd grote consensus en dat is dan ook de reden waarom er geen budgetten worden vrijgemaakt om veel onafhankelijke validiteits- en betrouwbaarheidsstudies op te zetten. Sommige mensen zien hierin “een complot”, wat natuurlijk klinkklare onzin is; hoe zouden onafhankelijke wetenschappers wereldwijd ooit kunnen verleid worden tot het opzetten van een complot?

Hoewel sommige mensen schermen met het idee dat indien men de tests in trainingen gebruikt, het doel is “mensen aan het praten en denken te zetten over verscheidenheid” en dat het daarom “niet belangrijk is of een test al of niet wetenschappelijk is”.

- ▶ Deze redenering is te eenvoudig want het indelen van mensen in typologieën geeft wel eens aanleiding tot een duidelijke stereotypering of tot “typearrogantie” (*“mijn stijl is beter dan die andere, ik ben extravert en dat is beter dan introvert...”*).
- ▶ Het effect dat een aantal tests daarmee beoogt, namelijk respect krijgen voor elkaars anderszijn, wordt daarmee dus niet bereikt, integendeel.

- ▶ Mensen leggen zich ook vaak “fatalistisch” neer bij de resultaten, terwijl de test helemaal niet juist hoeft te zijn. Mochten ze de test binnen de 4 tot 5 weken opnieuw doen, zouden ze overigens grote kans hebben dit zelf te ontdekken, want de kans op een ander type is groot!

Bovendien gaan tests zoals deze MBTI – zoals reeds eerder gesteld - voorbij aan andere vaststellingen uit wetenschappelijk onderzoek:

- Ongeveer 65% van de bevolking heeft geen uitgesproken “tendens” qua interactiestijl noch persoonlijkheid
- Uit onderzoek met 360° instrumenten blijkt dat ongeveer 25% van de bevolking een verkeerd zelfbeeld heeft op vlak van hun gedrag ten opzichte van anderen.

Deze vaststellingen gaan lijnrecht in tegen de praktijken van sommige tests en de MBTI in het bijzonder om iemand “een vakje” toe te wijzen, vaak op basis van kunstmatige ingrepen zoals een gedwongen verdeling en een geforceerde berekening.

ALTERNATIEF: De Big Five of het “Five Factor Model for personality” is een wetenschappelijk onderbouwde theorie die vele malen tegen het licht werd gehouden met de meest rigide methodes en waarover de grootst mogelijke consensus bestaat. De Big Five test die de meeste erkenning geniet is **de NEO-PI-R** van Costa & McCrae. Deze test mag alleen worden afgenomen en teruggekoppeld door psychologen of speciaal daartoe opgeleide mensen. (uitgeverij: Swets & Zeitlinger; Harcourt Test Publishers Verbindingsdok Oostkaai 13, bus 1, 2000 Antwerpen, Telefoon: +32 (0)3 202.05.09, Fax: +32 (0)3 202.05.10, website: <http://www.swetest.be/>). Zie ook de beschrijving elders in deze brochure.

NEO-PI-R (big five persoonlijkheidstest)

Dit is een persoonlijkheidsvragenlijst die de Big Five persoonlijkheidstrekken meet:

- Emotionele stabiliteit (Engels: Neuroticism)
- Extraversie
- Openheid voor ervaringen
- Altruïsme (vriendelijkheid)
- Consciëntieusheid

De test werd oorspronkelijk in 1992 ontworpen door Paul Costa en Robert McCrae voor het National Institute of Aging (NIH, Baltimore) in het kader van een grootschalig longitudinaal onderzoek waarin een grote groep mensen gedurende hun levensloop wordt gevolgd. De bestaande persoonlijkheidstests (MBTI, 16PF Catell...) werden niet geschikt bevonden, en uit noodzaak ontwikkelden zij een nieuwe persoonlijkheidsvragenlijst

De NEO-PI-R is in vele andere talen genormeerd en gevalideerd. De Nederlands-Vlaamse versie werd genormeerd door de professoren Hoekstra, Ormel, (beiden Universiteit Groningen) en De Fruyt (Universiteit Gent). De Franstalige versie werd gevalideerd en genormeerd door Rolland en Petot van de Université de Paris X-Nanterre.

De NEO-PI-R werd ontwikkeld voor het beschrijven van de normale variabiliteit van persoonlijkheidsverschillen, en niet primair om psychologische afwijkingen te meten. Dit maakt het instrument breed inzetbaar voor diverse doeleinden, ondermeer het bedrijfsleven.

De NEO-PI-R geeft een uitgebreide beschrijving van de vijf dimensies van persoonlijkheid: neuroticisme (emotionele stabiliteit versus emotionele labiliteit), extraversie, openheid, vriendelijkheid en conscientieusheid. Iedere dimensie is verder onderverdeeld naar zes facetten, zodat in totaal 30 facetten worden gemeten van de persoonlijkheid.

Diverse geautoriseerde anderstalige bewerkingen van de NEO-PI-R zijn beschikbaar, ondermeer in het Frans en het Nederlands.

- *Nederlandse versies:* NEO PI-R en NEO-FFI door H. A. Hoekstra, J. Ormel, & F. de Fruyt en uitgegeven door Swets & Zeitlinger, Lisse, The Netherlands.
- *Franse versie van de NEO PI-R* door J. P. Rolland & J. M. Petot en uitgegeven door Les Editions du Centre de Psychologie Appliquee, Paris.

De NEO is in Swets Test Manager af te nemen met behulp van expertrapportages. Door gebruik te maken van deze expertrapportages kunnen niet-psychologen snel toegang krijgen tot het instrument

Er bestaat zowel een “zelfrapport als een “observatorrapport”

Validiteit en betrouwbaarheid:

- **Betrouwbaarheid:** De interne consistentie coëfficiënten variëren van .86 tot .95 voor de “dimensieschalen” (big five) en van .56 tot .90 voor de “facetschalen” (6 per dimensie).
- **Validiteit:** Er is een grote validiteit tussen zelfscore, “peer”score en partnerscore. Er is een sterke validiteit gevonden voor construct validity en voor convergente en divergente validiteit.

Onafhankelijk onderzoek (buiten het onderzoek van Costa & McCrae)

- Botwin, M. D. (1996). Review of the Revised NEO Personality Inventory. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 862-863). Lincoln: University of Nebraska Press.
- Briggs, S. R. (1992). Assessing the Five-Factor Model of personality description. *Journal of Personality*, 60, 253-293.
- Butcher, J. N., & Rouse, S. V. (1997). Personality: Individual differences and clinical assessment. *Annual Review of Psychology*, 47, 87-111.
- Dolliver, R. H. (1987). A review of the NEO Personality Inventory. *Journal of Counseling and Development*, 66, 107-108.
- Hess, A. K. (1992). [Review of The NEO Personality Inventory]. In J. C. Conoley & J. J. Kramer (Eds.), *The eleventh mental measurements yearbook* (pp. 603-605). Lincoln, NE: Buros Institute of Mental Measurements.
- Hogan, R. (1989). [Review of The NEO Personality Inventory]. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 546-547). Lincoln, NE: Buros Institute of Mental Measurements.
- Juni, S. (1996). Review of the Revised NEO Personality Inventory. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 863-868). Lincoln: University of Nebraska Press.
- Leong, F. T. L., & Dollinger, S. J. (1990). NEO Personality Inventory. In D. J. Keyser & R. C. Sweetland (Eds.), *Test Critiques* (Vol. VIII, pp. 527-539). Austin, TX: PRO-ED.
- McLachlan, J. (1992, July). [Product review: Revised NEO Personality Inventory]. *QuadraPsych: Section on Computers in Psychology*, 6-7.
- Morey et al: Leslie C. Morey, Brian D. Quigley, Charles A. Sanislow, Andrew E. Skodol, Thomas H. McGlashan, M. Tracie Shea, Robert L. Stout, Mary C. Zanarini, John G. Gunderson. Substance or Style? An Investigation of the NEO-PI-R Validity Scales *Journal of Personality Assessment*. Volume 79, Issue 3, Page 583-599, 2002
- Piedmont, R. L. (1997). Test review: The NEO PI-R. *Newsnotes*, 32(3), 3-4.
- Piedmont, R. L. (1998). *The Revised NEO Personality Inventory: Clinical and research applications* [Test review]. New York: Plenum.
- Schinka, J. A., & Borum, R. (1994). Readability of normal personality inventories. *Journal of Personality Assessment*, 62, 95-101.
- Tinsley, H. E. A. (1994). NEO Personality Inventory-Revised. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques* (Vol. X, pp. 443-456). Austin, TX: PRO-ED.
- Young M.S. & Schinka J.A. (2001): Research Validity Scales for the NEO-PI-R: Additional Evidence for Reliability and Validity. *Journal of Personality Assessment*. Volume 76, Issue 3, Page 412-420, 2001
- Widiger, T. A. (1992). [Review of The NEO Personality Inventory]. In J. C. Conoley & J. J. Kramer (Eds.), *The eleventh mental measurements yearbook* (pp. 605-606). Lincoln, NE: Buros Institute of Mental Measurements.
- Widiger, T. A., & Trull, T. J. (1997). Assessment of the Five-Factor Model of personality. *Journal of Personality Assessment*, 68, 228-250.

Een neerslag van de meer dan 1500 onderzoeksartikelen gepubliceerd in een waaier van “blind peer-reviewed” tijdschriften (“journals”) is terug te vinden op: http://www3.parinc.com/uploads/pdfs/NEO_bib.pdf. De Web of Science biedt het meest recente overzicht.

Deze test blijkt ook een duidelijke link te kunnen leggen met leiderschap:

- Anderson, S. (1993). Psychological type and leadership. *Dissertation Abstracts International*, 54, 378.
- Asherian, A. (1993). Managerial leadership and effectiveness, and their relationship to personality in the former USSR. *Dissertation Abstracts International*, 54, 1709.
- Ashton, M. C., Lee, K., & Paunonen, S. V. (1999). What is the central feature of extraversion?: Social attention versus reward sensitivity. *Journal of Personality and Social Psychology*, 83, 245–251.
- Avolio, B. J. (1999). *Full leadership development*. Thousand Oaks, CA: Sage.
- Avolio, B. J., Bass, B. M., & Jung, D. I. (1999). Re-examining the components of transformational and transactional leadership using the Multifactor Leadership Questionnaire. *Journal of Occupational and Organizational Psychology*, 72, 441–462.
- Avolio, B. J., Dionne, S., Atwater, L., Lau, A., Camobreco, J., Whitmore, N., et al. (1996). *Antecedant predictors of a "full range" of leadership and management styles* (Tech. Rep. No. 1040). Alexandria, VA: United States Army Research Institute for the Behavioral and Social Sciences.
- Barling, J., Weber, T., & Kelloway, E. K. (1996). Effects of transformational leadership training on attitudinal and financial outcomes: A field experiment. *Journal of Applied Psychology*, 81, 827–832.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9-30.
- Barrick, M.R. & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Barrick, M.R. & Mount, M.K. (1993). Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. *Journal of Applied Psychology*, 78, 111-118.
- Barrick, M.R., Mount, M.K. & Judge, T. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9-30.
- Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York: Free Press.
- Bass, B.M. (1990). *From transactional to transformational leadership: Learning to share the vision*. *Organizational Dynamics*, 18, 19-31.
- Bass, B. M. (1998). *Transformational leadership: Industrial, military, and educational impact*. Mahwah, NJ: Erlbaum.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155–162.
- Block, J. (1995) A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187–215.
- Bono, J. E., (2001). Self-determination at work: The motivational effects of transformational leaders. Unpublished doctoral dissertation, University of Iowa.
- Bono, J. E., & Ilies, R. (2002). The transference of positive emotions: From leader to follower? In J. E. Bono, *Leadership and emotions: A multipleperspective approach*. Symposium conducted at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.
- Bono, J. E., & Ilies, R. (2003) A multilevel examination of links between leaders and followers. In N. Ashkanasy (Chair), *Multi-level perspectives on emotions in organizations*. Symposium conducted at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Bono, J. E., & Judge T. A. (2004). Personality and Transformational and Transactional Leadership: A Meta-Analysis. *Journal of Applied Psychology*, Vol. 89, No. 5, 901-910.
- Bragason H (1998): *Relationship between leadership styles and personality traits*. Amsterdam Vrije Universiteit (doctorale thesis)
- Burns, J. M. (1978). *Leadership*. New York: Harper & Row.
- Bycio, P., Hackett, R. D., & Allen, J. S. (1995). Further assessments of Bass's (1985) conceptualization of transactional and transformational leadership. *Journal of Applied Psychology*, 80, 468–478.
- Carless, S. A. (1998). Assessing the discriminant validity of transformational leader behavior as measured by the MLQ. *Journal of Occupational and Organizational Psychology*, 71, 353–358.
- Church, A. H., & Waclawski, J. (1998). The relationship between individual personality orientation and executive leadership behavior. *Journal of Occupational and Organizational Psychology*, 71, 99–125.
- Cooley, J.C., & Impara, J.P. (Eds.). (1995). *12th mental measurements yearbook*. Lincoln, NE: Buros Institute and University of Nebraska Press.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor (NEO-FFI) Inventory Professional Manual*. Odessa, FL: PAR.
- Costa, P. T., Jr., & McCrae, R. R. (1995). *Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory*. *Journal of Personality assessment*, 64 (1), 21-50.

- Costa, P. T., Jr., & McCrae, R. R. (1997). *Stability and change in personality assessment: The revised NEO personality inventory in the year 2000*. *Journal of Personality Assessment*, 68 (1), 86-94.
- Crampton, S. M., & Wagner, J. A. (1994). Percept-percept inflation in microorganizational research: An investigation of prevalence and effect. *Journal of Applied Psychology*, 79, 67-76.
- De Fruyt, F., & Mervielde, I. (1999). RIASEC types and Big Five traits as predictors of employment status and nature of employment. *Personnel Psychology*, 52, 701-727.
- De Fruyt, F., & Salgado, J. (2003a). Personality and IWO Applications: Introducing personality at work. *European Journal of Personality*, 17, S1-S3.
- De Fruyt, F., & Salgado, J. (2003b). Applied personality psychology: lessons learned from the IWO field. *European Journal of Personality*, 17, 123-131.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral & Brain Sciences*, 22, 491-569.
- Digman, J.M. (1990). *Personality structure: The emergence of the five-factor model*. In M.R. Rosenzweig & L.W. Porter (Eds.), *Annual review of psychology*, (Vol. 41, pp 417-440). Palo Alto, CA? USA: Annual Reviews.
- Dubinsky, A. J., Yammarino, F. J., & Jolson, M. A. (1995). An examination of linkages between personal characteristics and dimensions of transformational leadership. *Journal of Business and Psychology*, 9, 315-335.
- Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, 45, 735-744.
- Eysenck, H. J., & Eysenck, S. G. B. (1975). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: Educational and Industrial Testing Services.
- Fuller, J. B., Patterson, C. E. P., Hester, K., & Stringer, D. Y. (1996). A quantitative review of research on charismatic leadership. *Psychological Reports*, 78, 271-287.
- Goleman, D. (1995). *Emotionele Intelligentie*. New York, Bantam.
- Graziano, W. G., & Eisenberg, N. H. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795-824). San Diego, CA: Academic Press.
- Graziano, W. G., Jensen-Campbell, L. A., & Hair, E. C. (1996). Perceiving interpersonal conflict and reacting to it: The case for agreeableness. *Journal of Personality and Social Psychology*, 70, 820-835.
- Hackman, M. Z., Furniss, A. H., Hills, M. J., & Paterson, T. J. (1992). Perceptions of gender-role characteristics and transformational leadership behaviors. *Perceptual and Motor Skills*, 75, 311-319.
- Hogan, J., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 849-870). San Diego, CA: Academic Press.
- Hogan, R., Curphy, G.J., & Hogan, J. (1994). *What we know about leadership effectiveness and personality*. *American Psychologist*, 49, 493-504.
- Hough, L. (1992). The "Big Five" personality variables—construct confusion: Description versus prediction. *Human Performance*, 5, 139-155.
- Hough, L. M., & Ones, D. S. (2002). The structure, measurement, validity, and use of personality variables in industrial work, and organizational psychology. In N. Anderson, D. Ones, H. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work & organizational psychology: Vol. 1* (pp. 233-277). London: Sage.
- House, R. J., Shane, S. A., & Herold, D. M. (1996). Rumors of the death of dispositional research are vastly exaggerated. *Academy of Management Review*, 21, 203-224.
- House, R. J., Spangler, W. D., & Woycke, J. (1991). Personality and charisma in the U.S. presidency: A psychological theory of leader effectiveness. *Administrative Science Quarterly*, 36, 364-396.
- House, R.J., & Howell, J.M. (1992). *Personality and charismatic leadership*. *Leadership Quarterly*, 3(2), 81-108.
- Howell, J. M., & Avolio, B. J. (1993). Transformational leadership, transactional leadership, locus of control, and support for innovation: Key predictors of consolidated-business-unit performance. *Journal of Applied Psychology*, 78, 891-902.
- Howell, J. M., & Higgins, C. A. (1990). Champions of change: Identifying, understanding, and supporting champions of technological innovations. *Organizational Dynamics*, 19, 40-55.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In E. Pervin & O. John (Eds.), *Handbook of personality* (pp. 102-138). New York: Guilford Press.

- Hunter, J.E., Schmidt, F.L., Rauschenberger, J.M. & Jayne, M.E.A. (2001). Intelligence, motivation, and job performance. In C.L. Cooper & E.A. Locke (Eds), *Industrial and organizational psychology* (pp 278-303). Oxford, UK: Blackwell.
- Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology, 85*, 751–765.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765–780.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. (2002). Discriminant and incremental validity of four personality traits: Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology, 83*, 693–710
- Koehler, C. DS. (1993). Personality traits associated with transformational leadership styles of secondary principals in Christian schools. *Dissertation Abstracts International, 53*, 4153–4154.
- Kouzes, J. M., & Posner, B. Z. (1997). *The leadership challenge*. San Francisco: Jossey-Bass.
- Lipsey, M.W., & Wilson, D.B. (1993). *The efficacy of psychological, educational, and behavioural treatment: Confirmation from meta-analysis*. *American Psychologist, 48*(12), 1181-1209.
- Locke, E., & Latham, G.P. (1984). *Goal setting: A motivational technique that works!* Englewood Cliffs, NJ: Prentice Hall.
- Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology, 71*, 402–410.
- Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformation and transactional leadership: A metaanalytic review of the MLQ literature. *Leadership Quarterly, 7*, 385–425.
- MacDonald, K. (1995). *Evolution, the five-factor model, and levels of personality*. *Journal of Personality, 63*, 3.
- Makiney, J. D., Marchioro, C. A., & Hall, R. J. (April, 1999). Relations of leader perceptions to personality, leadership style and self schema. Paper presented at 14th Annual Meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.
- Marusic, I., & Bratko, D. (1998). Relations of masculinity and femininity with personality dimensions of the five-factor model. *Sex Roles, 38*, 29–44.
- McAdams, D.P. (1994). *Can personality change? Levels of stability and growth in personality across the life span*. In T.F. Heatherton & J.I. Weinberger (Eds.), *Can personality change?* (pp. 299-314). Washington, DC: American Psychological Association.
- McCrae, R. R. (1994). Openness to experience: Expanding the boundaries of Factor V. *European Journal of Personality, 8*, 251–272.
- McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin, 120*, 323–337.
- McCrae, R. R., & Costa, P. T., Jr. (1997). Conceptions and correlates of openness to experience. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 825–847). San Diego, CA: Academic Press.
- McNeese-Smith, D. K. (1999). The relationship between managerial motivation, leadership, nurse outcomes, and patient satisfaction. *Journal of Organizational Behavior, 20*, 243–259.
- Meibaum, A. J. (1986) The antecedents and consequences of transformational leadership. *Dissertation Abstracts International: Section B: the Sciences & Engineering, 53*, 215.
- Monroe, M. J. (1997). Leadership of organizational change: Antecedents and implications. *Dissertation Abstracts International, Humanities and Social Sciences, 58*, 1364.
- Murphy, K. R., & DeShon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53*, 873–900.
- Northouse, P. G. (1997). *Leadership: Theory and practice*. Thousand Oaks, CA: Sage.
- Phillips, J. S., & Lord, R. G. (1986). Notes on the practical and theoretical consequences of implicit leadership theories for the future of leadership measurement. *Journal of Management, 12*, 31–41.
- Ployhart, R. E., Lim, B. C., & Chan, K. Y. (2001). Exploring relations between typical and maximum performance ratings and the five-factor model of personality. *Personnel Psychology, 54*, 809–843.
- Prochaska, J.O., Norcross, J.C., & DiClemente, C.C (1994). *Changing for good*. New York: William Morrow.
- Prochaska, J.O., Diclemente, C.L., & Norcross, J.C. (1992). *In search of how people change. Applications to addictive behaviours*. *American Psychologist, 47*(9), 1102-1114.
- Ross, S. M., & Offerman, L. R. (1997). Transformational leaders: Measurement of personality attributes and work group performance. *Personality and Social Psychology Bulletin, 23*, 1078–1086.

- Schmidt, F. L., Viswesvaran, C., & Ones, D. S. (2000). Reliability is not validity and validity is not reliability. *Personnel Psychology*, 53, 901–912.
- Stogdill, R.M. (1974). *Handbook of leadership*. New York: Free Press.
- Van Eron, A. M. (1992). Key components of the transformation/transactional leadership model: The relationship between individual differences, leadership disposition, behavior and climate. *Dissertation Abstracts International*, 52, 4507.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574.
- Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 767–793). San Diego, CA: Academic Press.
- Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, 75, 315–321.
- Yukl, G., & Van Fleet, D. D. (1996). Theory and research on leadership in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 147–197). Palo Alto, CA, USA: Consulting Psychologists Press.

Conclusie:

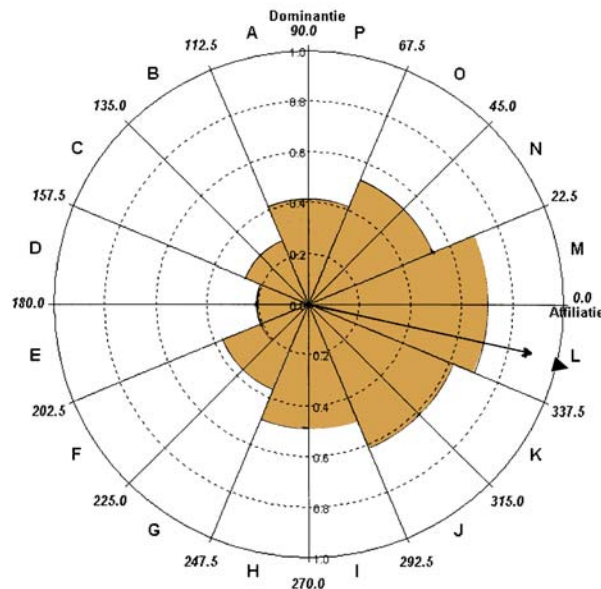
Samen met het interpersonal circumplex is dit de test waarover de grootste overeenstemming bestaat in de academische wereld. De test kan in verschillende domeinen worden toegepast.

De neutrale verwoording van vragen en rapportages maakt dat de NEO-PI-R in verschillende settings bruikbaar is, zoals verschillende toepassingen binnen de HRM-sector, als selectie, mobiliteit, ontwikkeling, en coaching. Per toepassingsterrein zijn diverse onafhankelijke en peer-reviewed studies beschikbaar in de internationale vakliteratuur die het gebruik van de NEO-PI-R illustreren en rechtvaardigen.

N-IAS

N-IAS staat voor “Nederlandse Interpersoonlijke Adjectieven Schalen” en kan worden gebruikt als 360°

Vrouwen (gemiddeld)



Over een periode van vijftig jaar is het interpersoonlijk circumplexmodel geëvolueerd van een “intuïtieve constructie” (T. Leary & The Kaiser Foundation) naar een **“wetenschappelijk gevalideerd model”** dat als referentiepunt voor vele theorieën en tests geldt. Diverse auteurs noemen het interpersoonlijk circumplex dan ook een “integratief model”; het gebruikt de kennis van de belangrijkste aanpakken en het biedt een kader om deze inzichten uit diverse richtingen te integreren en te systematiseren (Horowitz, 2004; Pincus & Ansell, 2003):

- Biologisch onderzoek: invloed van biologie en niet-gedeelde omgevingsfactoren
- Cognitieve gedragstherapie: cognities en schema’s spelen een rol, vooral bij de interpretatie van (ambigu) gedrag
- Humanistische aanpak: “zelf” versus “ander”
- Psychodynamische aanpak: motieven
- Communicatietheorie
- ...

Samen met de “Big Five” (die persoonlijkheidstrekken meet) is de IAS -R (Interpersonal Adjective Scales - Revised) één van de weinige tests waarover in de academische wereld zo een grote consensus bestaat (deze Big Five tests meten wel andere zaken).

De **onafhankelijke studies** naar validiteit en betrouwbaarheid zijn te omvangrijk om op te sommen. De interpersoonlijke circumplexschalen zijn in diverse landen naar de eigen taal vertaald en daar opnieuw (met succes) onderzocht op validiteit en betrouwbaarheid. De oorspronkelijke modellen werden ook omgezet naar andere toepassingen (vb. QTI – Questionnaire on Teacher Interaction, IIP-C – Inventory of Interpersonal Problems) en opnieuw betrouwbaar en valide gevonden.

De IAS-R van Wiggins bestaat sinds 1988 in het Engels en nog enkele andere talen, en sinds kort dus ook in het Nederlands na jaren van onderzoek en onderzoek door Danny Rouckhout en Rik Schacht. Door de contacten tussen deze wetenschappers en consulenten van Performance Coaching, genieten wij het voorrecht deze ook in onze trainingen en coaching te mogen inzetten.

De N-IAS meet gedragsrepertoire op 3 wijzen:

1. “hoe gedraagt u zich volgens uzelf tegenover anderen?”
2. “hoe zou u zich graag gedragen?”
3. “wat vinden anderen van het gedrag dat u vertoont ten aanzien van hen?”

De totale test bestaat dus uit 3 vragenlijsten die telkens dezelfde adjectieven bevragen.

N-IHS en F-IHS in ontwikkeling

De N-IHS meet gedragsrepertoire op dezelfde manier, maar nu bestaan de vragenlijsten uit zinnen die “handelingen” beschrijven. Van deze test bestaat er ook een Franstalige versie, de **F-IHS**. Beide tests bevinden zich (2004-2005) echter nog in de onderzoeksfase en is nog niet genormeerd; om te vergelijken gebruiken we voorlopig de Nederlandstalige database. De afnamen zullen worden gebruikt voor de verdere ontwikkeling van het instrument zelf en voor de validering van de N-IAS.

Thomas-Killmann Instrument (TKI)

Dit instrument is zowat het meest gebruikte instrument om de conflictstijl van mensen na te gaan.

Er zijn duidelijk verwantschappen en parallellen met het circumplex model (zie andere rubriek). Als men de twee hoofdasen van het circumplex neemt (en men dus 4 basistypes overhoudt), kan men de Thomas-Killman conflictstijlen er zo bovenop leggen.

Er zijn een paar bedenkingen te maken bij het model en de test:

- De theorie dat men elke stijl evenwichtig moet kunnen hanteren is slechts ten dele juist, zo blijkt uit diverse bronnen. Naast de situationele aspecten die aanleiding geven tot conflicten (moeheid, mate van belang, machtsrelatie met de andere conflictpartij...) spelen de **dispositionele aspecten** (de aanleg van de persoon of persoonlijkheidstype) eveneens een belangrijke rol. **Als mensen een aanleg hebben voor bepaalde stijlen (dispositioneel) dan zijn er slechts twee stijlen die een positief effect hebben op de eigen gezondheid ("collaborating" of "competing") en slechts één stijl die systematisch betere resultaten oplevert ("collaborating"), met name dát krijgen wat men zelf belangrijk vindt.**
- Er dient bovendien een belangrijk onderscheid te worden gemaakt tussen "taakconflicten" en "relatieconflicten". Taakconflicten kunnen positieve nieuwe denkpijlers opleveren¹⁴, terwijl relatieconflicten leiden tot een verslechtering van de werksfeer en de prestaties.
- Dr. A. Rahim stelde vast dat er **in organisaties 3 verschillende situaties** zijn die elk leiden tot een andere keuze voor conflictbeheer:
 - een conflict met mijn gelijke
 - een conflict met mijn meerdere
 - een conflict met een ondergeschikte
 Daardoor moet er volgens hem dan ook met een meer gedifferentieerde schaal worden gewerkt.
- De **betrouwbaarheid van dit instrument is (onaanvaardbaar) lager**¹⁵ in vergelijking met andere instrumenten.
- De test maakt gebruik van de ipsatieve vragenlijsten, met telkens een geforceerde keuze tussen twee stellingen. Ipsatieve vragenlijsten leveren heel wat problemen op (zie elders in dit rapport)¹⁶

¹⁴ De Dreu & Van de Vliert, 1997; A. Nauta, 1996 ("Oog om oog en baas boven baas")

¹⁵ Thomas-Killmann: interne consistentie: 0,43 tot 0,71; test-hertest betrouwbaarheid van 0,61 tot 0,68 * onder 0,70 is laag, zie begin van deze tekst * ROCI II van Rahim: interne consistentie: 0,72 tot 0,76; test-hertest betrouwbaarheid van 0,60 tot 0,83

¹⁶ Ondermeer:

- Tenopyr, M.L., "Artifactual reliability of forced-choice scales", Journal of applied psychology (1988), 73, 749-751.
- Baron, H, "Strengths and Limitations of Ipsative Measurement, Journal of Occupational and Organizational psychology (1996), 69, 49-56
- Meade, A.W; "Psychometric problems and issues involved with Creating and using ipsative measures for selection?" Journal of Occupational and Organizational Psychology (2004), 77, 531-552 *The British Psychological Society*

Eindconclusie

Mensen hebben van nature een grote behoefte zichzelf of de dingen die rondom hen gebeuren te verklaren. Dit verklaart de bijna vruchteloze zoektocht naar “mirakels”, het geloof in goeroes die niet gespeend zijn van enig wetenschappelijk inzicht enz.

Deze zoektocht naar “dé waarheid” of een “wonderlijke ontdekking” maakt mensen spijtig genoeg kwetsbaar voor bedrog en misleiding. Spijtig genoeg worden velen vaak (bijvoorbeeld het NLP-verhaal) in het ootje genomen door “briljante” marketeers en verkopers.

Kan het inzetten van tests nuttig zijn?

Hier kan zeker affirmatief op geantwoord worden want vele tests vergroten het inzicht in zichzelf, anderen, en de relatie met anderen (bijvoorbeeld het vergroten van de “functionele ruimte” uit het Johari-Window of het vergroten van het zelfbewustzijn als één van de peilers van emotionele intelligentie).

Het gebruik van tests is echter altijd contextgebonden: zoekt iemand daadwerkelijk eerlijke informatie over zichzelf of heeft deze persoon een andere doelstelling voor ogen (vb zich goed voor te doen voor een sollicitatie, een promotie, een assessment, zichzelf te vergoelijken...).

Tests zijn dus in het algemeen slechts waardevol voor de zorgvuldige gebruiker en mits toegediend en uitgelegd door een zorgvuldige coach, trainer of therapeut.

Het type van tests en hun wenselijkheid

Voor het bedrijfsleven zijn wij meer gewonnen voor het inzetten van tests die het interpersoonlijke gedrag in kaart brengen, en minder gewonnen voor “verklarende” persoonlijkheidstests. Voor zogenaamde “type” tests zijn wij helemaal niet gewonnen.

Bovendien werken wij ook het liefst met “multi-rater” tests, gelet op de eerder uiteengezette problematieken met zelfbeoordelingen.

Persoonlijkheidstests die “bouwstenen” meten zetten wij alleen maar in zeer specifieke gevallen in (vb bij coaching van leidinggevende, om oorzaken te helpen vinden voor bijvoorbeeld autoritair gedrag).

Het nut van vele “persoonlijkheidstests” zoals de “types” – vooral bij gebruik in trainingscontext - is immers meestal alleen maar dat **ze een dialoog tussen mensen op gang brengen en dat ze een soort van bewustwording teweeg brengen dat andere mensen anders in elkaar zitten. De praktische vertaling (transfer) naar de werksituatie of het wijzigen van gedrag is quasi steeds afwezig (op een**

hardnekkige “believer” na). Dat is zeker al een probleem voor de bedrijven die een meerwaarde wensen (en niet alleen maar een “nice to have” willen aanbieden aan hun medewerkers).

De vraag is bovendien of er werkelijk dergelijk (vaak duur) materiaal voor nodig is dat bovendien qua onderbouwing vaak zo weinig blijkt te kloppen.

Er is bovendien een groot gevaar voor stereotypering bij bepaalde tests, zelfs wanneer deze tests zorgvuldig worden uitgelegd en gebruikt. Dit risico wordt des te groter naarmate je te maken hebt met pseudo-wetenschappelijke tests die vaak door leken (zogenaamde practitioners) worden toegediend.

Wat is “wetenschappelijk”?

“Scientific research is systematic, controlled, empirical, and critical investigation of natural phenomena guided by theory and hypotheses about the presumed relations among such phenomena” (Verhallen, 1999)

Er wordt vaak smalend gedaan over wetenschap als zijnde dat de ene wetenschapper de andere tegenspreekt, of dat je met “statistiek alles kan bewijzen”.

Eén van de grote verschillen met “pseudo-wetenschappen” is juist dat wetenschappers hun eigen verklaringsmodellen **permanent en openbaar ter discussie stellen**. In tegenstelling tot “pseudo-wetenschappen” en het gros van “commerciële modellen” is er vrijheid en openbaarheid in publicatie. Daarom ook dat de meeste “ernstige” wetenschappelijke tijdschriften zorgen dat een publicatie eerst onderworpen wordt aan “peer-review”; collega-wetenschappers die het artikel lezen en beoordelen. Vandaar de strenge methodologie en de kritieken op elke bewering die zo typisch zijn aan écht wetenschappelijk onderzoek.

Vooraleer iets als écht wetenschappelijk wordt beschouwd onder wetenschappers is het nodig dat het onderzoek **repliceerbaar**ⁱⁱⁱ en **falsificeerbaar**^{iv} is. Bovendien praten wetenschappers meestal in termen van “de voorlopige stand van de wetenschap”.

De gevaren van pseudo-wetenschappelijke tests of methodes

Laat ons even van start gaan door Dr. Rory Coker, van de universiteit van Texas te citeren:

“Het is helaas bittere noodzaak voor iedereen om wetenschap en pseudo-wetenschap te leren onderscheiden. Iedereen moet in staat zijn om betrouwbare informatie te zoeken en te herkennen. Behoorlijk wat goed opgeleide mensen denken dat pseudo-wetenschap te idioot is om gevaarlijk te zijn, eerder een bron van amusement dan van zorgen. Die houding getuigt van weinig doorzicht. Pseudo-wetenschap kan extreem gevaarlijk zijn. Als ze doordringt in de politiek verschaft ze argumenten om etnische zuiveringen goed te praten. Als ze doordringt in het onderwijs staat ze wetenschap en kritisch denken in de weg. Als ze doordringt in de gezondheidszorg leidt dit tot onnodig lijden en de dood van duizenden. Als ze doordringt in de godsdienst krijgen we fanatisme, onverdraagzaamheid en heilige oorlog. Als ze doordringt in communicatie en media dan wordt het nagenoeg onmogelijk om betrouwbare informatie te vinden en worden belangrijke beslissingen genomen op basis van compleet irrationele argumenten.”

In sommige gevallen – vooral gevallen waarin mensen wanhopig zijn omdat ze bijvoorbeeld terminaal ziek zijn – zijn mensen vatbaar voor manipulatie.

In Nederland werd begin februari 2004 door de Inspectie voor de Gezondheidszorg (IGZ) een klacht ingediend tegen het “genezend medium” Jomanda om haar strafrechtelijk te vervolgen. De inspectie vond dat Jomanda actrice Sylvia Millecam heeft “misleid” door te zeggen dat zij geen borstkanker had. Millecam overleed in augustus 2001 aan die ziekte nadat zij reguliere medische hulp afwees na contacten met ondermeer Jomanda. KWF Kankerbestrijding onderschrijft het oordeel van de Inspectie voor de Gezondheidszorg dat een medische diagnose als kanker uitsluitend door erkende artsen gesteld moet worden: “In de praktijk blijkt dat alternatieve behandelaars zich op het exclusieve terrein van reguliere artsen begeven. Daardoor krijgen kwetsbare en vaak wanhopige patiënten foutieve informatie over de oorzaken van kanker. Bovendien krijgen zij behandelingen voorgeschreven die geen bewezen effect op kanker hebben. Patiënten ontlenen hieraan valse hoop. Als zij vervolgens alleen kiezen voor niet bewezen behandelingen kunnen zij in een levensbedreigende situatie komen.”

Met psychologische tests loopt het natuurlijk niet zo’n vaart – van levensbedreigende fouten hebben we nog niet gehoord – maar gevaarloos is het toch niet: vooral typologieën geven wel eens aanleiding tot een duidelijke stereotypering of tot “typearrogantie” (mijn stijl is beter dan die andere, ik ben extravert en dat is beter dan introvert...). Het effect dat een aantal tests daarmee beoogt, namelijk respect krijgen voor elkaars anders-zijn, wordt daarmee dus niet bereikt, integendeel.

“In het algemeen staan pseudo-wetenschappers onverschillig tegenover HOE bewijsmateriaal wordt verkregen. De nadruk ligt hoegenaamd niet op betekenisvolle, gecontroleerde en herhaalde wetenschappelijke experimenten. Wat gebruikt wordt, zijn onmogelijk na te trekken verslagen van ooggetuigen, verhaaltjes, “horen zeggen”, geruchten en twijfelachtige anekdoten. Naar wetenschappelijke literatuur wordt zelden verwezen. Over het algemeen geeft de pseudo-wetenschappen GEEN ENKEL geldig bewijs voor wat hij beweert.” (Bron: Rory Coker, Universiteit Texas, en bewerkt door Dirk Ploegaerts)

Bij pseudo-wetenschappelijke tests of methodes dient men ook op te letten dat men niet in de valkuilen trapt die door briljante (maar doortrapte) marketeers worden aangelegd: hiervoor verwijzen wij naar psycholoog Anthony R. Pratkanis, Phf.D., die 9 strategieën heeft geïdentificeerd¹⁷ om pseudo-wetenschappelijke zaken te verkopen (zoals verklaringen die bedoeld zijn om het vertrouwen te winnen in de aard van “ondersteund door meer dan 100 studies”)

Ook dient gewaarschuwd voor één van de valkuilen van ons denken, namelijk “foutief attribueren”: het is niet omdat je jezelf herkent in de test of omdat veel mensen de test hebben afgelegd, dat de test ook “juist” of “waardevol” is (net zomin als het feit dat de meeste mensen te hard rijden, dit dan ook wel zal toegestaan zijn of een verstandige keuze is...). **In de wetenschappelijke psychologie wordt de neiging van mensen om zichzelf te herkennen in een brede algemene persoonsomschrijving ook wel aangeduid als het ‘Barnum-effect’ – genoemd naar de “grootste circusman aller tijden P.T. Barnum”, die ook wel bekend staat als een meesterlijke psychologische manipulator.**

Mensen zijn doorgaans zeer tevreden met een algemene karakterschets wanneer ze menen dat die speciaal voor hen bedoeld is. Dit zogenoemde Barnum-effect trad duidelijk aan het licht in een aardig experiment van de neurochirurg Dick Zeilstra. Hij plaatste een advertentie in *De Telegraaf* waarin hij namens het fictieve bureau Grafospect voor fl. 17,50 een handschriftanalyse aanbood. Iedereen die zo’n analyse

¹⁷ Pratkanis AR. How to sell a pseudoscience, *Skeptical Inquirer* 19(4):1925, 1995

aanvraag, ontving dezelfde karakterbeschrijving plus een begeleidende enquête, die door driekwart van de aanvragers werd ingevuld. Uit de enquête bleek dat iedereen zich grotendeels en soms volledig in de karakterschets kon herkennen. De meerderheid was van mening dat de analyse hen meer inzicht had gegeven in de eigen persoonlijkheid.

Eén of meerdere tests?

Alle tests meten zonder uitzondering slechts een deelaspect van onze persoonlijkheid of ons gedrag; het is gevaarlijk één test te gebruiken en daar allerhande “psychotherapeutische” verklaringen uit af te leiden die er in werkelijkheid niet uit af te leiden vallen -zo dient men ook bijvoorbeeld gerespecteerde test als de IAS enkel te gebruiken om interpersoonlijk gedrag in kaart te brengen, en is het geenszins de bedoeling te verregaande “psychoanalytische” conclusies te gaan trekken over de oorzaken of verklaringen voor dit gedrag.

Het is belangrijk tests alleen maar in te zetten voor wat ze bedoeld zijn en niets meer.

Aanbevelingen

Naast de aanbevelingen gegeven in de aanhef van dit document, verwijzen wij naar het feit dat er **in de academische wereld veel tests en modellen vrij toegankelijk zijn**. Het is nu eenmaal een kenmerk van “wetenschappelijkheid” dat men zijn kennis deelt en dat men zijn onderzoek publiceert.

Met deze modellen en tests worden dezelfde trainingsresultaten geboekt (vb zelfinzicht en respect voor anderen). Daarnaast bieden deze modellen en tests bijkomende voordelen ten opzichte van louter commerciële of pseudo-wetenschappelijke modellen:

- Ze zijn vaak goedkoper of zelfs vrij te gebruiken (vb circumplextesten).
- Ze bieden tegelijk bewezen accuraatheid en wetenschappelijke onderbouwing, en worden vaak opnieuw onderzocht en verder verfijnd.
- Ze zetten mensen niet in “vakjes”: ze tonen stuk voor stuk aan dat mensen zeer complex zijn en erg gevarieerde persoonlijkheden hebben. Mensen hebben vaak wel tendensen, maar “stereotypes” zijn het zeer zelden.
- Ze leiden niet tot het zogenaamde “barnum”-effect noch tot “typearrogantie”.

Met een gerust hart kunnen we onze klanten dus aanbevelen inzicht te verschaffen in het eigen gedrag of persoonlijkheidstrekken het gedrag van anderen zonder gebruik te maken van relatief dure tests. Bovendien is er voor elke “onwetenschappelijke” wel een “wetenschappelijke” test; zo kan men “juistheid” combineren met “leereffecten”.

ⁱ Bron: Nederlands Tweelingen Register (NTR), 2001.

ⁱⁱ Bron: Nederlands Tweelingen Register (NTR), 2001; de resultaten werden bekomen door bij meer dan 40.000 Nederlandse kinderen de “Child Behavior Checklist” af te nemen. Ook tweelingen werden in dit onderzoek ruim ingesloten.

ⁱⁱⁱ **Repliceerbaar** = door andere onderzoekers (in andere labo's indien het onderzoek betrekking heeft op laboratoriumproeven) herhaalbaar en daarbij de onderzoeksresultaten van de eerdere studies bevestigend.

^{iv} Een (wetenschappelijke of andere) theorie wordt gezegd **falsificeerbaar** te zijn als het mogelijk is een experiment uit te voeren waardoor bij een bepaalde uitkomst die theorie moet worden verworpen

