



Emotionele intelligentie als criterium bij werving en selectie

door

Barend Koch
Van der Maesen | Koch HRM-advies

Verslag van de 5^{de} Research Meeting of the Dutch-Flemish Network on Recruitment and Selection

Het Dutch-Flemish Network on Recruitment and Selection bestaat uit Vlaamse en Nederlandse wetenschappers en praktijkmensen uit het werkveld van werving en selectie. Het netwerk stelt zich ten doel om de kwaliteit van het onderzoek te verbeteren, de onderlinge samenwerking te bevorderen en een brug tussen wetenschap en praktijk te slaan.

Zie ook: <http://www.eur.nl/fsw/psy/onderzoek/onderzoekswebsites/rsr/>

Inhoud

Deel I : Emotionele Intelligentie	3
Wetenschappelijk onderzoek over Emotionele intelligentie	3
<i>Geschiedenis</i>	3
<i>Kritiek</i>	3
<i>Prestatietests in plaats van vragenlijsten</i>	4
<i>Onderzoekresultaten</i>	4
<i>Emotionele intelligentie: Jekyll & Hyde</i>	5
Deel II : Actueel onderzoek op het gebied van werving en selectie	7
1. Is er één algemene factor om persoonlijkheid te beschrijven?	7
2. Niveaus van persoonlijkheidsmeting	9
3. De verbeterde predictie van ander-oordelen ten opzichte van zelfoordelen	10
4. De aantrekkelijkheid van organisaties voor werkzoekenden	11
5. Het opsporen van 'faking' en ander afwijkend antwoordgedrag	12
6. Toetsen van functiekennis	14
Deel III : Algemene beschouwing van alle presentaties	16
Verantwoording	19

WETENSCHAPPELIJK ONDERZOEK OVER EMOTIONELE INTELLIGENTIE

**Stéphane Côté - Rotman School of
Management, University of Toronto, Canada**

Stéphane Côté onderzoekt hoe emotionele vaardigheden van invloed zijn op de prestaties en het welzijn van mensen. Hij publiceert boeken en wetenschappelijke artikelen en traint daarnaast managers en verzorgt onderwijs. Hij publiceerde onder meer over onderzoek waaruit blijkt dat mensen die hun emoties beter weten te reguleren, in het algemeen gelukkiger zijn en meer status en inkomen hebben dan mensen die minder goed met hun emoties om kunnen gaan.

GESCHIEDENIS

Het begrip Emotionele Intelligentie (EI) werd voor het eerst genoemd in een publicatie van Hunt uit 1928. Salovey en Mayer pakten het construct vanaf 1990 serieus op en definiëren het nader als het vermogen om emoties te herkennen, te begrijpen en vervolgens te reguleren, zowel bij zichzelf als bij anderen. In 1995 volgt de doorbraak bij het grote publiek als Daniel Goleman met zijn bestseller 'Emotionele Intelligentie' het begrip populariseert. Côté noemt de gevolgen 'good, bad and ugly'. Hij is blij met de aandacht voor het onderwerp ('good'), maar de keerzijde is dat jan-en-alleman ermee aan de haal gaat ('bad') en dat soms de grootste onzin wordt uitgekraamd ('ugly'). Zo citeert hij een topmanager die ten onrechte beweert dat Emotionele Intelligentie twee keer zoveel bijdraagt aan werkprestaties als intelligentie en technische vaardigheden bij elkaar. In werkelijkheid ligt dit veel genuanceerder zoals later in de presentatie wordt toegelicht.

KRITIEK

Kritiek vanuit wetenschappelijke hoek is er volop. Emotionele Intelligentie is als construct niet valide, het is oude wijn in nieuwe zakken, het drijft meer op goede marketing dan op overtuigende onderzoeksresultaten en het voegt niets toe aan de voorspelling van arbeidsprestaties. Veel van deze kritiek is volgens Côté terug te voeren op de wijze waarop Emotionele Intelligentie wordt gemeten, namelijk met vragenlijsten waarin respondenten zichzelf beoordelen. Een dergelijke meting is immers niets anders dan die van de reguliere persoonlijkheidsvragenlijsten, terwijl de termen Emotionele Intelligentie en EQ suggereren dat het, net als bij cognitieve intelligentie en IQ, om prestatietests zou gaan.



PRESTATIETESTS IN PLAATS VAN VRAGENLIJSTEN

Prestatietests waarmee emotionele intelligentie wordt gemeten zijn er inmiddels wel. De Mayer-Salovey-Caruso Emotional Intelligence Test (Mayer et al., 2003), kortweg MSCEIT, is daarvan de bekendste. De MSCEIT heeft vier schalen die gaan over het herkennen, faciliteren, begrijpen en reguleren van emoties. Kenmerkend voor de scoring is dat respondenten niet wordt gevraagd om zichzelf te beoordelen. In plaats daarvan beoordelen respondenten verschillende stimuli en situatieschetsen waarin emoties centraal staan. De oordelen van de respondenten worden vergeleken met die van experts. Een hogere overeenkomst met de expertoordelen leidt tot een hogere score op Emotionele Intelligentie. Deze aanpak lijkt sterk op de methode van de situationele beoordelingstests (Engels: Situational Judgment Test, afgekort als SJT). In SJT's wordt steeds een situatie gepresenteerd gevolgd door een aantal reacties die door de respondent beoordeeld moeten worden. De context is in de MSCEIT toegespitst op het omgaan met emoties.

Onlangs ontwikkelden Richard Roberts en Carolyn MacCann twee SJT's die expliciet zijn bedoeld om Emotionele Intelligentie te meten, namelijk de Situational Test of Emotional Understanding (STEU) en Situational Test of Emotional Management (STEM) (MacCann & Roberts, 2008). Nele Libbrecht (2011) promoveerde in januari van dit jaar bij Filip Lievens (beide Universiteit Gent en lid van het netwerk) op de waarde van dit type alternatieve metingen voor Emotionele Intelligentie.

ONDERZOEKSRESULTATEN

Côté presenteert een aantal studies waarin Emotionele Intelligentie met prestatietests in plaats van met vragenlijsten is gemeten. De resultaten samengevat:

- Emotionele Intelligentie hangt sterk samen met Cognitieve Intelligentie. Côté

rapporteert een onderzoek waarin hij een correlatiecoëfficiënt vindt van .47 (n=175).

- Beide soorten intelligentie voorspellen arbeidsprestaties.
- Emotionele Intelligentie voorspelt vooral prestaties die hier theoretisch aan zijn gerelateerd.
- Interpersoonlijke prestaties worden beter voorspeld dan intellectuele prestaties.
- Emotionele Intelligentie levert extra voorspellende waarde bovenop Cognitieve Intelligentie en ook op andere traditionele voorspellers zoals persoonlijkheid.
- Wat betreft arbeidsprestaties lijkt het er op dat personen die laag scoren op Cognitieve Intelligentie dat voor een groot deel kunnen compenseren met een hoge score op Emotionele Intelligentie.



EMOTIONELE INTELLIGENTIE: JEKYLL & HYDE

Een hoge EI-score zegt wel iets over iemands inzicht in emoties maar niet over hoe dat inzicht in het dagelijks leven wordt gebruikt. Volgens Côté kan Emotionele Intelligentie ten goede worden ingezet, zoals bij het oplossen van conflicten en het motiveren van anderen. Aan de andere kant is het ook mogelijk om emotionele intelligentie in het eigenbelang aan te wenden om anderen te manipuleren. Zowel sociaal wenselijke als sociaal onwenselijke persoonlijkheidseigenschappen kunnen op die manier door emotionele intelligentie worden versterkt. Voor de selectiepraktijk betekent dit volgens Côté dat het extra risicovol is als kandidaten met een egoïstisch

persoonlijkheidsprofiel ook een hoge EI-score hebben.

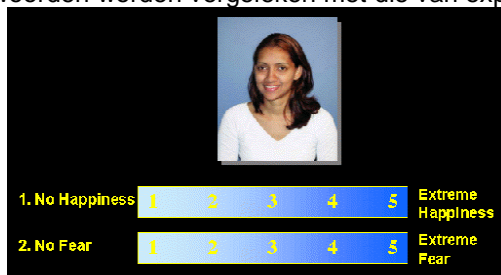
Gaat Emotionele Intelligentie over alle emoties of kunnen die nog van elkaar worden gescheiden? Want is het niet zo dat iemand met de ene emotie beter om kan gaan dan met een andere? Dat iemand boosheid bijvoorbeeld goed weet te reguleren maar angst niet? Het is dan niet alleen de vraag óf je emotioneel intelligent bent, maar ook in welke emoties je intelligent bent. Côté stelt: 'De verschillende emoties worden samengenomen vanwege de hoge onderlinge correlaties, maar het uitsplitsen naar verschillende emoties kan mogelijk leiden tot verdere nuancering'.



Ter verduidelijking van wat de MSCEIT inhoudt worden op de volgende bladzijde de verschillende schalen kort beschreven en geïllustreerd met een voorbeeld.

I. Emoties herkennen

Foto's met gezichtsuitdrukkingen worden getoond. Vervolgens moet de respondent de emotie die wordt uitgedrukt duiden op een vijfpuntsschaal, bijvoorbeeld van zeer ongelukkig (1) tot extreem gelukkig (5). De antwoorden worden vergeleken met die van experts. Hoe meer overeenkomst, hoe hoger de score.



II. Faciliteren

Welke stemming helpt je als je voor het eerst je schoonfamilie gaat ontmoeten?

- | | | | |
|--------------|--------|-----------|-------------|
| - Vreugde | nuttig | 1 2 3 4 5 | niet nuttig |
| - Spanning | nuttig | 1 2 3 4 5 | niet nuttig |
| - Verrassing | nuttig | 1 2 3 4 5 | niet nuttig |

III. Emoties begrijpen

Respondenten krijgen korte situatieschetsen te lezen en moeten daarbij aangeven welke emoties de hoofdpersoon zal voelen. Bijvoorbeeld:

Frank keek op tegen de hoeveelheid werk die hem die dag stond te wachten. Toen zijn baas hem nog eens een extra opdracht kwam geven voelde hij zich:

- Zelfverzekerd
- Overdonderd
- Beschaamd
- Zenuwachtig

IV. Emoties reguleren

Korte situatiebeschrijvingen worden aangeboden waarin emoties of stemmingen centraal staan, gevolgd door een aantal acties. De respondent wordt gevraagd om te beoordelen hoe effectief de acties in die situatie zijn. De oordelen van de respondent worden ook hier afgezet tegen de oordelen van experts.

Carla kwam terug van vakantie en voelde zich ontspannen en tevreden. Hoe goed helpen Carla de volgende acties om in die stemming te blijven.

Een lijst maken met dingen die ze nog thuis moet doen:

Niet effectief 1 2 3 4 5 *Zeer effectief*

Nadenken waar en wanneer ze de volgende keer op vakantie zal gaan:

Niet effectief 1 2 3 4 5 *Zeer effectief*

Dit gevoel negeren omdat ze dit toch niet lang vast kan houden:

Niet effectief 1 2 3 4 5 *Zeer effectief*

1. IS ER ÉÉN ALGEMENE FACTOR OM PERSOONLIJKHEID TE BESCHRIJVEN?

Dimitri van der Linden - Erasmus Universiteit Rotterdam

Eén algemene factor om onze persoonlijkheid te beschrijven? Het voelt bijna aan als een provocatie. De discussies rond dit controversiële onderwerp kunnen dan ook hoog oplopen. Dimitri van der Linden (Erasmus Universiteit Rotterdam) stelt zichzelf de vraag of de ultieme reductie tot één algemene g-factor mogelijk is, zoals dat ook bij cognitieve capaciteiten het geval is. Reductie tot een enkele persoonlijkheidsfactor die allerlei wenselijke eigenschappen in zich verenigt. Er bestaat ook al een naam voor: GFP, wat staat voor General Factor of Personality.



Welke basale persoonlijkheidsfactoren kunnen we onderscheiden? Het meest gangbare antwoord is dat dit de inmiddels welbekende dimensies uit het Big Five model zijn: Extraversie, Vriendelijkheid, Zorgvuldigheid, Emotionele Stabiliteit en Openstaan voor Ervaringen. Naast deze bekende vijf zijn er ook andere persoonlijkheidsmodellen

waarin minder of juist meer factoren worden onderscheiden, bijvoorbeeld Eysenck's Big Three, (Extraversie, Neuroticisme, Psychoticisme) het HEXACO-model met zes factoren (Big Five + Integriteit) of het 16-PF model van Catell met wel 16 factoren (sociaal, intelligent, vrolijk, dominant, ontspannen, radicaal, zelfgedisciplineerd, onafhankelijk, emotioneel stabiel, consciëntieus, timide, gevoelig, naïef, praktisch, predispositie voor schuld, etcetera).

Controversieel of niet, in zijn presentatie onderbouwt Van der Linden het bestaan van een algemene factor binnen het persoonlijkheidsdomein met statistische onderzoeksresultaten. De logische vervolgvraag op deze statistische bevinding is welke betekenis aan zo'n algemene factor moet worden toegekend. Van der Linden zoekt het antwoord in een combinatie van sociaal wenselijke eigenschappen ('De ene persoon is nu eenmaal sociaal wenselijker dan de ander', zo drukte Willem Hofstee het eens uit). Daarnaast wijst hij op hoge correlaties (.65 -.77) met Sociale en Emotionele Intelligentie (zoals gemeten met zelfbeoordelingsvragenlijsten) en de inhoudelijke overeenkomsten die deze constructen vertonen. Ook de interpretatie die Van der Linden aan GFP geeft sluit aan op de interpretatie die Côté aan Emotionele Intelligentie geeft: wenselijke persoonlijkheidseigenschappen zoals vriendelijkheid worden nog eens extra opgetrokken door een hoge score op GFP. Ook zijn er verbanden gevonden met verschillende gedrags- en prestatie-maten. Personen met hoge scores op GFP zijn bijvoorbeeld populairder en leveren hogere arbeidsprestaties. Hoe de inzet van GFP zijn beslag moet gaan krijgen in de selectiepraktijk is vooralsnog niet duidelijk geworden.

Reductie tot één algemene score betekent dat het aantal vragenlijstitems kan worden gereduceerd om tot een betrouwbare meting te komen. Tegelijkertijd biedt één score weinig aanknopingspunten voor feedback. Hoe je een kandidaat met een lage GFP-score op een

begrijpelijke en acceptabele manier uitlegt wat dit voor die persoon betekent zal voor veel selectiepsychologen geen eenvoudige opgave zijn.

Uitbreiding van het aantal persoonlijkheidsschalen in plaats van reductie is een ander scenario. Het is voorstelbaar dat een GFP-score wordt berekend uit de items van een bestaande persoonlijkheidsvragenlijst en dat deze score in de rapportage als extra score wordt toegevoegd aan de reeds bestaande schalen. De waarde van zo'n extra schaal hangt niet alleen af van onderzoeksresultaten en de uitkomsten van het wetenschappelijk debat, maar ook van de vraag of de psycholoog die het persoonlijkheidsprofiel moet interpreteren de GFP-score op een zinvolle manier weet te duiden. Wellicht is er ook op dit vlak voor van der Linden c.s. nog veel missiewerk te verrichten.



2. NIVEAUS VAN PERSOONLIJKHEIDMETING

Thomas Sitser - TestGroup Consulting, Erasmus Universiteit Rotterdam

Gebruikers van persoonlijkheidsvragenlijsten hebben vaak de indruk dat persoonlijkheidsscores kunnen voorspellen hoe succesvol een sollicitant zal zijn. Voor het gemak gaan ze vaak voorbij aan de vraag wat 'succesvol' is. Dit kun je breder of smaller definiëren. Een verkoper kan in brede zin succesvol zijn omdat zijn leidinggevende vindt dat hij 'in het algemeen' goed functioneert. Een smallere maat voor succes is bijvoorbeeld de mate waarin hij zijn administratie op orde heeft.

Sitser onderscheidt vier niveaus van maten (criteria) voor succes. Van breed naar smal zijn dat:

- 1) Algemene niet-functiespecifieke beoordeling
- 2) Algemene beoordeling, functiespecifiek (bijvoorbeeld 'verkoper')
- 3) Specifieke beoordeling op aspecten binnen een functie (bijvoorbeeld 'het onderhouden van relaties met klanten' of 'het afhandelen van administratie')
- 4) Objectieve maten die niet zijn gebaseerd op beoordeling, maar op telbare gegevens die relevant zijn voor de functie (bijvoorbeeld 'het aantal nieuwe klanten dat iemand heeft weten aan te trekken in een bepaalde tijdsperiode')

Net als maten van succes kun je ook persoonlijkheidseigenschappen breder of smaller definiëren.

In zijn onderzoek onderscheidt hij drie niveaus van persoonlijkheidsmeting. Van breed naar smal:

- 1) General Factor of Personality (GFP)
- 2) Big Five (Extraversie, Vriendelijkheid, Zorgvuldigheid, Emotionele Stabiliteit en Openstaan voor Ervaringen)
- 3) 13 smalle en specifieke persoonlijkheidsfacetten (bijvoorbeeld stessbestendigheid, detailgerichtheid)

Het gaat Sitser er om te achterhalen met welk niveau van persoonlijkheidsmeting je de beste voorspelling kunt doen voor elk van de genoemde criteriumniveaus.

Uit de onderzoeksresultaten die hij laat zien wordt zijn veronderstelling in grote lijnen bevestigd dat het niveau van de voorspeller aan moet sluiten op het niveau van het criterium. Brede succesmaten kunnen het beste worden voorspeld met brede persoonlijkheidsfactoren. GFP zou geschikter zijn om overallbeoordelingen te voorspellen dan welk persoonlijkheidsfacet dan ook. Smalle persoonlijkheidsfacetten zijn juist beter geschikt om specifieke succesmaten te voorspellen. De mate waarin iemand zijn administratie op orde heeft kan bijvoorbeeld beter met een persoonlijkheidsfacet als 'detailgerichtheid' worden voorspeld dan met de brede GFP.

DISCUSSIE

De conclusie, die ook in de discussie werd besproken, is dat het belangrijk is om eerst een beslissing te nemen van wát je wilt voorspellen voordat je kiest welke predictoren je in gaat zetten en op welk niveau deze moeten worden gemeten.



3. DE VERBETERDE PREDICTIE VAN ANDER-OORDELEN TEN OPZICHTE VAN ZELFOORDELEN

Ute Hülshager - Universiteit van Maastricht

Auteurs: U.R. Hülshager, Brian S. Connelly
(University of Toronto)

Ute Hülshager (Universiteit Maastricht) begint haar presentatie met de vaststelling dat een aantal vooraanstaande onderzoekers zoals Fred Morgeson, Michael Campion en Neal Schmitt het gebruik van persoonlijkheidsvragenlijsten ter discussie stelt, met name wanneer deze voor selectiedoeleinden worden ingezet. Belangrijkste kritiekpunt is dat persoonlijkheidsmeting door de bank genomen een zeer matige voorspeller is gebleken voor toekomstige arbeidsprestaties. Daarbij wordt wel eens over het hoofd gezien dat de overgrote meerderheid van onderzoek naar de voorspellende waarde van persoonlijkheid betrekking heeft op metingen via zelfbeoordeling. Validiteitstudies naar persoonlijkheidsmeting waarbij de beoordeling door anderen wordt gedaan laten duidelijk hogere correlaties zien met arbeidsprestaties.



Voor de bevinding dat ander-oordelen arbeidsprestaties beter voorspellen dan persoonlijkheidsmeting via zelfbeoordeling geeft Hülshager twee alternatieve verklaringen:

1) De "Narrower scope"-verklaring:
In de meeste studies worden de ander-oordelen uitgevoerd door collega's. Deze collega's kennen

de beoordeelde persoon vooral binnen de werkcontext en kunnen hem of haar dan goed in dit licht beoordelen. Omdat de blik van de collega's zich focust op gedrag op het werk, zo is de redenering, zijn hun oordelen bij uitstek in overeenstemming met de context waarin de arbeidsprestaties plaatsvinden.

2) De "Clearer lens"-verklaring:

In het algemeen is het oordeel van anderen objectiever en accurater dan wanneer personen een oordeel over zichzelf geven. Als beoordelaar en beoordeelde samenvallen kan dat een minder zuiver en vertroebeld beeld opleveren.

Hülshager toetst beide verklaringen door te kijken naar de oordelen van bekenden die juist geen deel uitmaken van de werkkring van de hoofdpersoon. Uit de onderzoeksresultaten blijkt dat niet alleen de oordelen van collega's leiden tot een betere voorspelling van arbeidsprestaties, maar dat dit ook het geval is als die oordelen worden overgelaten aan familie, vrienden en partners van de beoordeelde persoon. De voorspellende waarde van de persoonlijkheidsmeting is dan minstens zo goed.

Hülshager concludeert dat haar bevindingen slecht nieuws zijn voor de "narrow scope"-verklaring. Persoonlijkheidsbeoordelingen door anderen binnen de werkcontext leiden in haar onderzoek niet tot betere voorspellingen van arbeidsprestaties dan beoordelingen door personen buiten de werkcontext. De verbeterde predictie van ander-oordelen ten opzichte van zelfoordelen is volgens haar helemaal toe te schrijven aan de "clearer lens"-verklaring: Of persoonlijkheidsoordelen nu binnen of buiten de werkcontext plaatsvinden, beoordelingen door anderen leiden tot betere voorspellingen van arbeidsprestaties dan wanneer deze van de persoon zelf afkomstig zijn.

Uit de bevinding dat persoonlijkheidsoordelen door bekenden buiten de werkkring een goede voorspellende waarde hebben volgt de logische discussievraag hoe dit dan praktisch moet worden vormgegeven. In een echte selectiesituatie spelen

er immers andere belangen mee dan in de onderzoekssituatie waarin de dataverzameling van het gepresenteerde onderzoek heeft plaatsgevonden. Voorlopige aanbeveling is om vrienden en bekenden alleen te betrekken in selectievraagstukken als er voor de kandidaat minder op het spel staat, bijvoorbeeld bij herplaatsing.

4. DE AANTREKKELIJKHEID VAN ORGANISATIES VOOR WERKZOEKENDEN

Greet Van Hoye (Universiteit Gent)

Auteurs: G. Van Hoye, D.B. Turban (University of Missouri)

Greet van Hoye (Universiteit Gent) onderzoekt of werkzoekenden zich meer tot een organisatie voelen aangetrokken wanneer de mensen die er werken passen bij hun eigen persoonlijkheid. Dit is een relevante onderzoeksvraag omdat organisaties steeds vaker de eigen werknemers in de werving betrekken, bijvoorbeeld door ze op de website aan het woord te laten in testimonials.

In haar onderzoek heeft van Hoye zich toegelegd op de persoonlijkheidseigenschappen extraversie, vriendelijkheid en consciëntieusheid. Zij veronderstelt dat werkzoekenden die hoog scoren op deze eigenschappen zich meer aangetrokken voelen tot organisaties met extraverte, vriendelijke en consciëntieuze werknemers dan werkzoekenden die laag scoren op die eigenschappen.

Deze intuïtief aansprekende veronderstelling onderbouwt ze met de 'trait activation' theorie van Tett (Tett & Burnett, 2003). Volgens deze theorie komen persoonlijkheidstrekken tot uiting in situaties die relevant zijn voor die trek, en omgekeerd zoeken mensen ook actief naar situaties (in dit geval hun werkomgeving) waarin zij hun persoonlijkheid tot uiting kunnen laten komen (Tett & Murphy, 2002). Deze laatste

veronderstelling sluit weer goed aan bij de attraction-selection-attrition (asa)-theorie van Schneider (Schneider 1987, 1995).

Uit de resultaten blijkt dat er voor de eigenschappen vriendelijkheid en consciëntieusheid ondersteuning is voor haar hypothese. Vriendelijke sollicitanten denken met andere woorden dat er een betere match is als de werknemers ook vriendelijk zijn. Hetzelfde mechanisme gaat op voor consciëntieuze sollicitanten die te maken krijgen met consciëntieuze werknemers van een organisatie.

Van Hoye adviseert organisaties die vriendelijke en consciëntieuze mensen willen aantrekken om vooral testimonials van werknemers te gebruiken waarin deze twee eigenschappen worden benadrukt.



Voor de eigenschap extraversie werd er opvallend genoeg geen effect gevonden. Extraverte sollicitanten worden volgens de onderzoeksresultaten niet aangetrokken door extraverte uitingen in testimonials. Een mogelijke verklaring die Van Hoye in de discussie geeft is dat een omgeving met introverte collega's ook aantrekkelijk kan zijn voor extraverte sollicitanten. Omdat zij in zo'n omgeving gemakkelijk aan het woord zullen komen, kunnen zij hun extraversie immers goed tot uiting laten komen.

DISCUSSIE

In de discussie werd verder vooral aandacht besteed aan de vraag of het wenselijk is om meer mensen aan te trekken met een zelfde persoonlijkheidsprofiel of dat het niet beter is om complementaire eigenschappen binnen de organisatie te halen. Gewaarschuwd wordt voor de gevolgen die Benjamin Schneider als de 'dark side of fit' beschrijft: extreme homogeniteit, inperking van creatieve processen, strategische kortzichtigheid, gebrek aan demografische diversiteit en het onvermogen van organisaties zich aan te passen.



Ook werd in de discussie de vraag gesteld hoe ethisch en geloofwaardig het is wanneer testimonials zo zijn geselecteerd dat ze een vertekend beeld geven van de mensen in de organisatie. Wanneer een organisatie geen realistische weergave biedt van de mensen die er rondlopen, dan leidt dat ongetwijfeld tot teleurstellingen bij sollicitanten. Dit zal weer zijn weerslag hebben op de organisatie. 'Valse' testimonials zullen op die manier eerder een averechts hebben waardoor de wal het schip keert.

5. HET OPSPOREN VAN 'FAKING' EN ANDER AFWIJKEND ANTWOORDGEDRAG

Iris Egberink - Universiteit Twente, thans Rijksuniversiteit Groningen

Iris Egberink ging in haar presentatie in op de bruikbaarheid van verschillende methoden om afwijkend antwoordgedrag bij persoonlijkheidsvragenlijsten op te sporen.

In de selectiecontext wordt bij afwijkend antwoordgedrag vaak gedacht aan het bewust faken door een sollicitant om op die manier een baan te krijgen, maar er bestaan ook andere vormen zoals willekeurige antwoordpatronen door ongemotiveerde respondenten of onbedoelde antwoordpatronen waarbij veelvuldig voor het schaal midden wordt gekozen.



Egberink onderscheidt in haar presentatie drie typen van opsporingsmethoden:

- Validiteitsschalen: Dit zijn speciale schalen die aan een test worden toegevoegd. Voorbeelden zijn schalen voor inconsistent antwoordgedrag of schalen voor sociaal wenselijk antwoordgedrag. Laatstgenoemde schalen bestaan uit items zoals: "Ik vertel nooit een leugentje om eigen bestwil" of "Al mijn gewoontes zijn goed en wenselijk". Hoe vaker de

respondent met deze items instemt, hoe hoger de score op de sociaal wenselijkheidsschaal. In de selectiecontext is deze opsporingsmethode weinig succesvol omdat sociaal wenselijk antwoordgedrag en daadwerkelijk gedrag meestal niet goed van elkaar zijn te onderscheiden (zie bijvoorbeeld Schmitt, N. & Oswald, (2006). Egberink merkt overigens op dat sociaal wenselijkheidsschalen in het algemeen sterk samenhangen met de eerder genoemde General Factor of Personality (GFP) uit de presentatie van Van der Linden.

- Person-fit methode: Een tweede methode om fouten op te sporen is van toepassing op tests die zijn ontwikkeld met de zogenaamde Item Response Theorie. Wanneer een kandidaat een vraag als: 'Ik ben zuinig op mijn spullen' met 'nee' beantwoordt dan zou het vreemd zijn als diegene de vraag 'ik ben *extreem* zuinig op mijn spullen' met 'ja' beantwoordt. Zo'n inconsistentie, in dit voorbeeld sterk gechargeerd, heet een Guttman-fout. Door het aantal Guttman-fouten te tellen kun je iets zeggen over de wijze waarop de test of vragenlijst is ingevuld. Een antwoordpatroon met veel Guttman-fouten betekent dat iemand niet consistent heeft geantwoord op vragen. De scores van die persoon kunnen dan ook niet goed worden geïnterpreteerd als verwijzend naar het concept, in dit voorbeeld is dat 'consciëntieusheid', wat de schaal wil meten.
- Overige statistieken: Deze bestaan uit het onderzoeken van het schaalgebruik van de respondent. De ene respondent heeft een voorkeur voor de middencategorie terwijl een ander graag de extremen opzoekt. Probleem van dit soort afwijkend antwoordgedrag is dat het vergelijking met andere personen in de weg staat.

Egberink had de beschikking over een databestand met persoonlijkheidsscores op de bekende Big Five dimensies, aangevuld met een sociaal wenselijkheidsschaal en een schaal voor inconsistent antwoordgedrag. Eén groep van 2217 respondenten, de sollicitantengroep, had de persoonlijkheidsvragenlijst ingevuld met het oog

op een toekomstige baan. Bij de andere groep met 1488 respondenten stond er minder op het spel omdat zij de vragenlijst in het kader van hun eigen loopbaanontwikkeling invulden. Aan het bestand werden data toegevoegd waarin afwijkende antwoordpatronen werden gesimuleerd.



Op het bestand werden de verschillende detectiemethoden losgelaten en onderzocht. De resultaten laten het volgende zien:

- De verschillende opsporingsmethoden detecteren verschillende soorten van afwijkend antwoordgedrag.
- De groep met sollicitanten laat vaker dan de loopbaanontwikkelingsgroep extreem sociaal wenselijk antwoordgedrag (minimaal antwoord vier op een vijf-punts Likertschaal) zien. Bij de loopbaanontwikkelingsgroep komen vaker willekeurige antwoordpatronen voor.
- Extreem antwoordgedrag in sociaal wenselijke richting kan het beste worden gedetecteerd met een sociaal wenselijkheidsschaal.
- Willekeurig antwoordgedrag kan het beste worden opgespoord met de eerder genoemde person-fit methode of met een schaal voor inconsistent antwoordgedrag.

DISCUSSIE

In de discussie wordt besproken dat de genoemde opsporingsmethoden weliswaar een indicatie kunnen geven, maar dat ze niet waterdicht zijn. Sollicitanten zouden niet op basis van een sociaal wenselijkheidsscore uit de selectieprocedure mogen worden gegooid. Aangeraden wordt om indicaties van 'faking' altijd te toetsen aan een interview.

6. TOETSEN VAN FUNCTIEKENNIS

Stefan Mol - Amsterdam Business School

Stefan Mol breekt in zijn presentatie een lans voor het gebruik van tests, of eigenlijk toetsen, die functiekenis meten (Engels: job knowledge tests).

Kennistoetsen kunnen verschillende vormen aannemen. Kenmerkend is in ieder geval dat deze bestaat uit vragen en opdrachten die relevant zijn voor de functie.



Hieronder een voorbeeldvraag uit een kennistoets voor dokters-assistenten.

De vriend van mijnheer Nielsen, 50 jaar, belt dat mijnheer Nielsen sinds tien minuten een heftig snoerende pijn op de borst voelt. Hij is bleek en transpireert. De dokters-assistent concludeert dat hier sprake is van urgentie U1.

- A) Dit is correct
- B) Dit is niet correct

In de bekende meta-analyse van Schmidt en Hunter (1998) komen job knowledge tests naar voren als één van de beste voorspellers voor arbeidsprestaties. Toch krijgt dit type tests weinig aandacht in het onderzoeksveld van de selectiepsychologie. Misschien komt dit wel omdat job knowledge tests sterk correleren met cognitieve capaciteitentests en daarom weinig kunnen toevoegen aan de voorspellende waarde op arbeidsprestaties.

Maar -zo betoogt Mol in navolging van diezelfde Schmidt en Hunter- de reden dat intelligente mensen beter presteren in hun werk is dat zij zich functiekenis beter en sneller eigen maken. De belangrijkste *directe* determinant voor arbeidsprestaties is dus niet intelligentie, maar functiekenis. In andere woorden voegt hij daar aan toe: "Je vliegt toch liever mee met een piloot die het besturingssysteem kent dan met een piloot die alleen maar intelligent is?".

Aan de hand van het model van Binning en Barrett (1992) laat Mol de verschillende soorten validiteit zien van job knowledge tests. Functiekenis is niet alleen een goede voorspeller voor arbeidsprestaties (criteriumvaliditeit), het staat ook rechtstreeks in relatie tot deze prestaties (constructvaliditeit). Verder worden job knowledge tests als relevant en eerlijk ervaren (indruksvaliditeit), met name als de vragen een goede afspiegeling zijn van de benodigde functiekenis (inhoudsvaliditeit).

Praktisch punt bij selectie is dat job knowledge tests niet automatisch kunnen worden ingezet bij externe sollicitanten. Volgens Schmidt en Hunter moet een respondent eerst relevante scholing of werkervaring hebben opgedaan voordat deze

kennis getoetst kan worden. Mol voegt hier aan toe dat deze beperking tegenwoordig niet vaak meer opgaat omdat voor vrijwel elke functie werkervaring of scholing wordt verondersteld.

Een ander aandachtspunt is dat het ontwikkelen van job knowledge tests een arbeidsintensieve en dus dure klus is. Voor elke functie en voor elk niveau moeten immers aparte vragen worden gegenereerd en gevalideerd. Aan de hand van een praktijkvoorbeeld toont Mol hoe dit proces met een gecomputeriseerd systeem efficiënt kan worden ingericht.

Conclusie is dat binnen het domein van de selectiepsychologie nog veel winst valt te behalen met deskundig ontwikkelde job knowledge tests, vooral als de ontwikkelkosten beperkt kunnen worden en wanneer de toetsen bij interne selectievraagstukken met ervaren kandidaten worden ingezet.

Bij de ontwikkeling van kennistoetsen lijkt een belangrijke taak te zijn weggelegd voor mensen uit de organisatie zelf. Omdat zij beschikken over de benodigde functiekennis en ook omdat zij weten wat relevant is en wat niet, zal de input voor toetsvragen van ervaren medewerkers moeten komen. Onder deskundige begeleiding kunnen zij ook zelf zorgen voor een groot deel van de toetsontwikkeling, wat in de meeste gevallen ook kostenbesparend werkt. Ook het afnemen van de toetsen kan uitstekend door de betreffende organisatie zelf worden uitgevoerd. In de rol van adviseur op gebied van o.a. itemontwikkeling, toetsamenstelling, scoring, normering, afnamecondities en terugkoppeling kan de selectiepsycholoog beslist van toegevoegde waarde zijn. Wanneer deze adviezen onderbouwd kunnen worden met toegepast onderzoek naar job knowledge tests dan zal deze waarde verder toenemen.



DEEL III :ALGEMENE BESCHOUWING VAN ALLE PRESENTATIES

De nadruk lag op selectie en op persoonlijkheid. Zes van de zeven presentaties gaan over selectie van personeel, één presentatie gaat over werving. Vijf van de zeven presentaties gaan over persoonlijkheid of persoonlijkheidsmeting. De problematiek van metingen met zelfoordelen is een aantal keren expliciet aan de orde geweest. Ute Hülshager stelde zelfoordelen aan de orde met haar betoog over de hogere voorspellende waarde van ander-oordelen in persoonlijkheidsvragenlijsten. Iris Egberink deed hetzelfde met haar verhaal over detectiemethoden naar het (extreem) sociaal wenselijk invullen van vragenlijsten en ander afwijkend antwoordgedrag. In het kader van persoonlijkheidsmeting is deze problematiek, namelijk de kritiek op persoonlijkheidsmeting via zelfbeoordeling, in de Dutch-Flemish meetings al eerder besproken, onder anderen in lezingen van Neil Christiansen in 2008 en van Robert Tett in 2009. Nico Smid (Smid, 2010) besteedde in zijn verslag van de vorige editie dan ook ruim aandacht aan dit vraagstuk. Ook gastspreker Stephane Côté roert het onderwerp van zelfbeoordeling aan en hij doet dat met een goede reden. Onderzoekers die zich net als hij bezig houden met Emotionele Intelligentie maken namelijk duidelijk onderscheid tussen EI via zelfbeoordeling en EI via prestatietests. Dit onderscheid gaat zelfs zo ver dat er in de literatuur sprake is van twee verschillende constructen die allebei het label 'Emotionele Intelligentie' dragen (zie bijvoorbeeld Joseph & Newman, 2010).

EI VIA ZELFBEOORDELING OF VIA PRESTATIETEST

Aan de ene kant is er het 'ability-model' waarin EI tot het intelligentiedomein wordt gerekend. Het gaat om het vermogen op een intelligente manier met emoties om te gaan door deze te herkennen, te begrijpen en te reguleren. Verstand en emotie beïnvloeden elkaar wederzijds. Verondersteld

wordt dat er samenhang bestaat met cognitieve intelligentie. EI wordt in het ability-model meestal gemeten met prestatietests. Aan de andere kant is er het 'mixed model', ook wel 'personality model' genoemd, waarin EI niet wordt geclassificeerd als intelligentie, maar meer als een combinatie van verschillende (non-cognitieve) vaardigheden en persoonlijkheidseigenschappen die er toe bijdragen dat iemand succesvol met omgevingsfactoren om kan gaan. Het gaat bijvoorbeeld om assertiviteit, inlevingsvermogen, flexibiliteit, stresstolerantie, impulscontrole en optimisme. Cognitieve en Emotionele Intelligentie worden binnen dit model als onafhankelijk van elkaar gezien. EI wordt in het mixed model gemeten met (zelfbeoordelings)vragenlijsten.



Conceptueel lijkt de General Factor of Personality (GFP) uit de presentatie van Van der Linden het meest aan te sluiten op het laatstgenoemde 'mixed model', al is het alleen maar vanwege de vragenlijstmethode waar GFP uit voortkomt. Zowel GFP als het mixed model-EI steunen op zelfoordelen en ze verenigen beide een groot aantal sociaal wenselijke eigenschappen in zich. Hoge scores hierop blijken samen te hangen met betere arbeidsprestaties, al is nog niet helemaal duidelijk welk mechanisme hier precies aan te grondslag ligt. De sociaal wenselijkheidsschalen uit de presentatie van Iris Egberink hebben dezelfde eigenschappen. Het hoeft dan ook geen

verbazing te wekken dat er sterke onderlinge correlaties bestaan. Ondanks de inspanningen die op dit vlak worden gedaan blijft het zoeken naar een eenduidige interpretatie van de genoemde concepten. Om deze reden werd op een gegeven moment tijdens de discussie gevraagd hoeveel GFP's er nu eigenlijk zijn.

Ondanks de problemen die er aan kleven, blijven zelfbeoordelingen binnen de selectiecontext een populaire methode. Dat heeft vooral te maken met economische motieven. Gebruik maken van oordelen door anderen (360-graden methode) is nu eenmaal arbeidsintensiever en is meestal ook niet wenselijk bij selectiebeslissingen vanwege de belangen die meespelen. In de loop der jaren zijn er verschillende oplossingen aangedragen om problemen met zelfoordelen tegen te gaan en om de aldus verkregen scores beter te kunnen duiden. In sommige gevallen is het middel erger dan de kwaal, zoals het gebruik van gedwongen-keuze vragen (hier valt meer over te zeggen, maar in de kern komt het er op neer dat je verwrongen antwoorden en scoreprofielen krijgt). Andere oplossingen zijn zondermeer aan te bevelen zoals het toetsen van scores aan een interview. Echter, het feit dat deze extra ingrepen nodig worden geacht zegt ook iets over de beperking van de methode.



Zoals gezegd spelen deze problemen niet bij prestatietests die EI het 'ability-model' meten. Wat dat betreft kunnen deze tests worden beschouwd

als een stap voorwaarts, maar dit betekent niet dat er geen beperkingen aan zitten.

Eén daarvan is dat de metingen iets zeggen over het inzicht dat de kandidaat heeft op het vlak van Emotionele Intelligentie, maar dat het nog niets zegt over hoe een kandidaat dit inzicht toepast in zijn gedrag. Voor zover het gaat om het herkennen en begrijpen van emoties hoeft dat geen probleem te zijn omdat dit voornamelijk een interne aangelegenheid is, maar bij het reguleren van emoties komt ook gedrag om de hoek kijken. Wéét een kandidaat bijvoorbeeld dat hij in een bepaalde situatie tot tien moet tellen om tot bezinning te komen, of dóet hij dat ook?

Het toevoegen van een gedragscomponent aan prestatietests voor Emotionele Intelligentie kan beslist van toegevoegde waarde zijn. Er moet wel rekening mee worden gehouden dat dit extra tijd en kosten met zich meebrengt omdat er beoordelingen door anderen van dat gedrag aan te pas moeten komen.

Dankzij technologische ontwikkelingen kunnen deze kosten sterk worden gereduceerd terwijl ook het gestandaardiseerde karakter van een test behouden blijft. Het is bijvoorbeeld mogelijk om kandidaten te laten reageren op een aantal gefilmde situaties waarin emoties centraal staan. Hun reacties worden via de webcam opgenomen en kunnen online worden beoordeeld. Janneke Oostrom deed onderzoek naar deze methode en toonde aan dat dit betrouwbare en valide metingen op kan leveren (Oostrom, 2010). Het zou interessant zijn om te onderzoeken of deze methode ook is in te zetten voor het meten van Emotionele Intelligentie en daarmee een bijdrage kan leveren aan de constructvaliditeit en de predictieve validiteit. Tijdens het afsluitende diner toonde Stéphane Côté zich alleszins geïnteresseerd in deze aanpak.

Referenties:

Binning JF, Barrett GV. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, 74,478-494.

Connelly Brian S. and Hülshager Ute R. (2011). A Narrower Scope or a Clearer Lens for Personality? Examining Sources of Observers' Advantages over Self-Reports For Predicting Performance *Journal of Personality* . Accepted manuscript

Joseph & Newman (2010) Emotional intelligence: an integrative meta-analysis and cascading model. *J Appl Psychol*. 2010 Jan ;95(1):54-78.

Kismihók, G., Vas, R. and Mol, S.T. (in press) 'An innovative ontology-driven system supporting personnel selection: the OntoHR case', *Int. J. Knowledge and Learning*, Vol. X, No. Y, pp.000–000.

Libbrecht N. (2011) Emotional intelligence tests: The value of alternative measurement approaches. Dissertation Ghent University.

Linden, D. van der, Bakker, A. B. & Serlie, A. W. (2011). The general factor of personality selection and assessment samples. *Personality and Individual Differences*, 641-645.

Linden, D. van der, Te Nijenhuis, J., Cremers, M. & Ven, C. van de (2011). General factors of personality in six datasets and a criterion-related validity study at the Netherlands armed forces. *International Journal of Selection and Assessment*, 19, 157-169.

Linden, D. van der, Nijenhuis, J. te & Bakker, A.B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44, 315-327.

MacCann, C., & Roberts, R. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, 8, 540–551.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3, 97-105.

Oostrom, J.K., Born, M.P., Serlie, A.W. & Molen, H.T. van der (2010). Webcam Testing: Validation of an innovative open-ended multimedia test. *European Journal of Work and Organizational Psychology*, 19(5), 532-550.

Schmidt, F.L. & J.H. Hunter (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124 pp. 262-274.

Schmitt, N. & Oswald (2006) The impact of corrections for faking on the validity of noncognitive measures in selection settings. *J Appl Psychol*. 2006 May ;91(3):613-21.

Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437-453.

Schneider, B., Goldstein, H.W. & Smith, D.B. (1995). The ASA Framework: An Update. *Personnel Psychology*, 48, 747-779.

Smid N. (2010) Persoonlijkheid in Organisaties. Meten is weten? Een studiedag. NSVP

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88*, 500–517.

Tett, R. P., & Murphy, P. J. (2002). Personality and situations in coworker preference: Similarity and complementary in worker compatibility. *Journal of Business and Psychology, 17*, 223-241.

VERANTWOORDING

Het Dutch-Flemish Network on Recruitment and Selection bestaat uit Vlaamse en Nederlandse wetenschappers en praktijkmensen uit het werkveld van werving en selectie. Het netwerk stelt zich ten doel om de kwaliteit van het onderzoek te verbeteren, de onderlinge samenwerking te bevorderen en een brug tussen wetenschap en praktijk te slaan. Voor meer informatie zie:

<http://www.eur.nl/fsw/psy/onderzoek/onderzoekswbsites/rsr/>

In dit document vindt u een samenvatting van de presentaties die zijn gehouden van de voordrachten die zijn gehouden tijdens de 5^e Research Meeting van het Nederlands-Vlaamse Netwerk voor Werving en Selectie. Het is geschreven voor de geïnteresseerde wetenschapper of professional die inzicht wil krijgen in het onderzoek dat momenteel plaatsvindt op het gebied van werving en selectie. Veel van het onderzoek dat in Gent werd besproken werd en wordt door de leden van het netwerk verder uitgewerkt en heeft inmiddels zijn beslag gekregen in een aantal publicaties (bijvoorbeeld Van der Linden, 2011; Kismihók, G., Vas, R. and Mol, S.T. (in press); Connelly & Hülshager (in press)). De actuele onderzoeksvragen die aan bod komen inspireren de discussies over verdere professionalisering van de werving- en selectiepraktijk. De nadruk bij het verslag ligt op de presentatie van de gastspreker, Stephane Côté, die de vraag behandelde hoe emotionele intelligentie van sollicitanten kan worden gemeten om voorspellingen te doen over het latere werkgedrag.

De opmaak en verspreiding van dit document is mede mogelijk gemaakt door de NSvP. De Nederlandse Stichting voor Psychotechniek (NSvP) ondersteunt de ontwikkeling wetenschappelijke methoden en praktijkkennis die bijdragen aan optimale inzet, betere motivatie en betrokkenheid van mensen in de werksituatie - www.innovatiefinwerk.nl.

