

## Overbenutting van persoonlijkheidsvragenlijsten bij competentie management

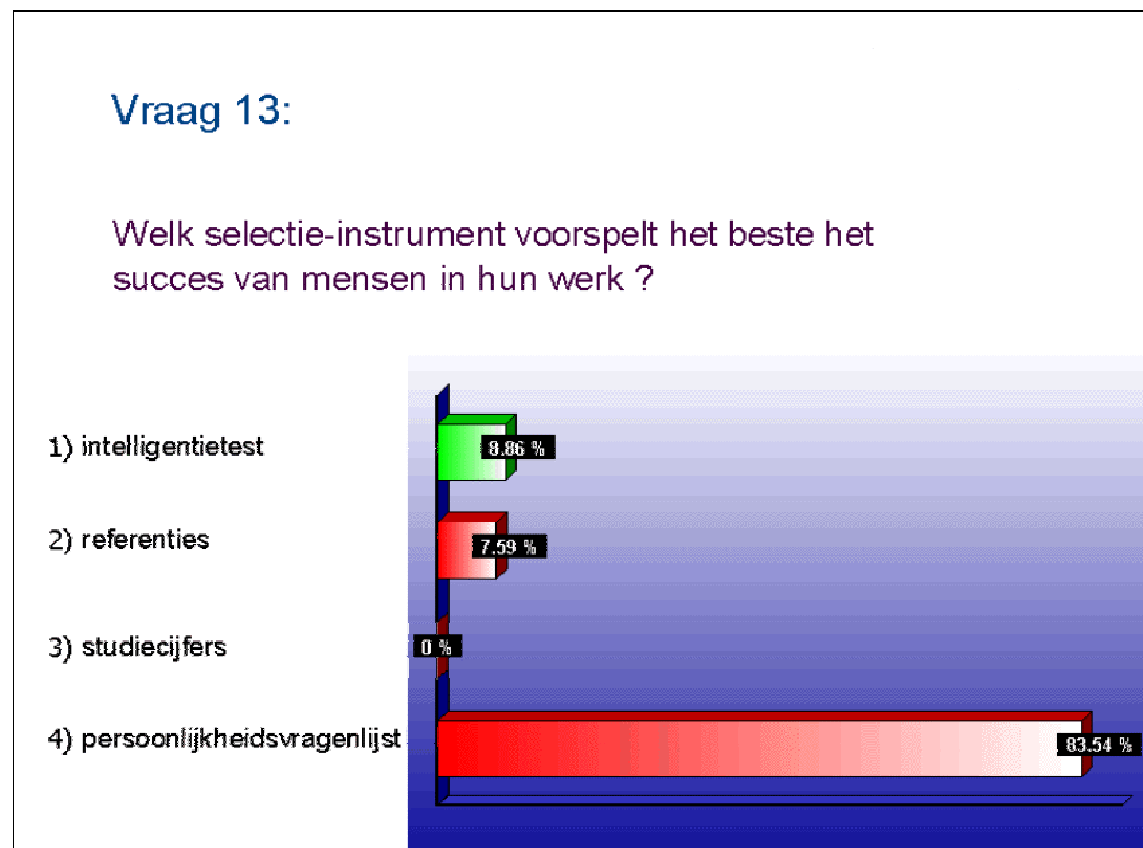
Paul van der Maesen de Sombreff

### Situatie

Een dag in oktober vorig jaar. Een honderdtal managers en adviseurs van een grote overheidsinstelling is bijeen in de School voor de Toekomst in Den Bosch op een congres over Human Resource Management. Eén van de programmaonderdelen is een HR-quiz. Vijftien vragen worden gepresenteerd die gaan over “feiten en effecten” op HRM-gebied. Deelnemers hebben de mogelijkheid om met een stemkastje hun antwoord te geven door te kiezen uit vier alternatieven.

Vraag 13 luidt als volgt: “Welk selectie-instrument voorspelt het beste het succes van mensen in hun werk?” De alternatieven zijn: a) intelligentietest b) referenties c) studiecijfers d) persoonlijkheidsvragenlijst.

Meteen na het beantwoorden van deze vraag volgt het plaatje van de uitkomsten. Dat zag er uit zoals hieronder:



### **Uitkomst**

Ongeveer 83% van de deelnemers denkt dat de persoonlijkheidsvragenlijst het best voorspellende selectie-instrument is, 9% kiest voor de intelligentietest, 7,5% kiest referenties en niemand kiest studiecijfers als de beste predictor.

Deze uitkomst is verrassend als die wordt afgezet tegen de correcte antwoorden. Op basis van empirisch onderzoek is namelijk gebleken dat de intelligentietest verreweg superieur is aan de persoonlijkheidsvragenlijst als voorspeller van succes in het werk.

Er waren in deze HR-quiz wel meer vragen waarvoor discrepanties bleken te bestaan tussen empirisch gefundeerde kennis en opvattingen van de deelnemers. De overwaardering van de persoonlijkheidsvragenlijst en de onderwaardering van de intelligentietest sprong er echter zeer uit. Ditzelfde verschijnsel is ook opgemerkt, door Rynes, Colbert en Brown (2002) die uitkomsten van een HR-quiz presenteerden. We hebben ook geen reden om te veronderstellen dat ditzelfde patroon niet zou optreden bij andere bedrijven en organisaties. Integendeel, mijn stelling is dat er bij Nederlandse bedrijven en organisaties een persoonlijkheidscultus in het competentie-management heerst die tot uiting komt in een niet gerechtvaardigd vertrouwen en gebruik van de persoonlijkheidsvragenlijst.

In de rest van dit artikel wil ik de volgende onderwerpen behandelen. Aan de hand van de *state-of-the-art* van empirisch onderzoek laat ik zien wat beide soorten instrumenten waard zijn. Een vraag die vanuit bedrijfseconomisch oogpunt relevant is betreft de schade die bedrijven lijden door overbenutting van de vragenlijst en onderbenutting van de test. Vervolgens wil ik nagaan wat kan verklaren waarom een intelligentietest effectiever is dan een persoonlijkheidsvragenlijst. Ik probeer verklaringen te geven voor de fascinatie die persoonlijkheidsvragenlijsten bij managers blijkbaar opwekken en voor de afkeer die intelligentietests wellicht oproepen. In de voorlaatste paragraaf beschrijf ik enkele ontwikkelingen die hoopvol stemmen over de mogelijkheid om vragenlijsten meer valide te maken en intelligentietests meer acceptabel. Ik sluit af met een aantal stellingen.

### **Voorspellende waarde**

Er is zeer veel onafhankelijk onderzoek gedaan, vooral in de VS, naar de voorspellende waarde van de twee soorten instrumenten. Die resultaten zijn ook gebundeld in samenvattende onderzoeken, *meta-analyses* genoemd. Uit de meta-analyses komen de beste schattingen van de echte validiteit. Ik baseer me op de resultaten die beschreven worden in een zojuist verschenen handboek voor bedrijfs- en organisatiepsychologie van Borman, Ilgen, Klimoski en Weiner (2003). Voor de intelligentietest baseer ik me op het hoofdstuk van Drasgow, voor de persoonlijkheidsvragenlijst op dat van Hough & Furnham.

De voorspellende waarde van intelligentietests voor "*overall*" werksucces is gemiddeld .51 (zie ook Schmidt & Hunter, 1998), die van persoonlijkheidsvragenlijsten gemiddeld .19. De validiteit van .19 van persoonlijkheidsvragenlijsten komt tot stand door de uitslagen op de "big five"-factoren (Extraversie, Vriendelijkheid, Gewetensvolheid, Stabiliteit en Openstaan voor ervaring) vermenigvuldigd met optimale gewichten bij elkaar op te tellen<sup>1</sup>.

### **Economisch nut**

Het economisch nut van instrumenten is rechtstreeks gekoppeld aan de predictieve validiteit van die instrumenten. Daarom kan men zeggen dat de intelligentietest (met een validiteit van .51) 2,5 keer zoveel nut heeft als de persoonlijkheidsvragenlijst (met een validiteit van ten hoogste .19). Wat betekent het verschil in validiteit tussen .51 en .19 in euro's?

---

<sup>1</sup> Dank aan dr. P.H. Dekker (statisticus bij A&O psychologie aan de VU Amsterdam) voor zijn hulp bij de SPSS-analyse van de samenvattende resultaten in Hough & Furnham (2003).

Empirisch onderzoek heeft inzicht gegeven in de waarde van prestatieverschillen van mensen binnen eenzelfde functie. Binnen een functie van academisch niveau verdient een medewerker die behoort tot de top 20% jaarlijks al vlug € 25.000 meer voor een organisatie dan een gemiddelde medewerker. Een organisatie die potentiële top-20% mensen wil selecteren vergroot met een meer valide methode de kans om die meerwaarde te bereiken. Na selectie met validiteit van .51 is de verwacht meeropbrengst per persoon per jaar ongeveer € 8.000 hoger dan bij selectie met validiteit van .19.<sup>2</sup>.

Als de organisatie de selectieprocedure vijf jaar aanhoudt en per jaar tien nieuwe mensen selecteert die gemiddeld vijf jaar in dienst blijven, bedraagt het verschil in utiliteit ongeveer twee miljoen euro.

Natuurlijk is deze procedure waarin uitsluitend wordt geselecteerd met één instrument onrealistisch, maar het gaat erom te laten zien welke bedragen gemoeid zijn met meer valide selectieprogramma's.

In de titel van dit artikel staat "competentiemanagement". Personeelsselectie is daarvan slechts een deelactiviteit. We zijn echter van mening dat het bij validiteit om een criterium gaat dat net zo goed van toepassing is op de kwaliteit van andere activiteiten die horen tot het competentie management, bijvoorbeeld van "self-assessment" voor loopbaanoriëntatie. Minder valide instrumenten leiden tot minder goede beslissingen van welke aard dan ook.

### **Wat maakt intelligentietests effectief?**

Wat maakt de intelligentietest effectiever dan de vragenlijst?

Een belangrijke oorzaak is dat de intelligentietest meer dan de vragenlijst afgestemd is op de kernvariabele waarin organisaties en bedrijven geïnteresseerd zijn: productiviteit in het werk. Het werken aan een intelligentietest is verwant met werkprocessen die productiviteitsbevorderend werken: snel informatie verwerken en schiften, informatie koppelen aan reeds beschikbare kennis, problemen analyseren en oplossen. Intelligentie is het omgaan met complexiteit, hoe complexer de omgeving hoe hoger de toegevoegde waarde van intelligentie. Over de inhoud van de persoonlijkheidsvragenlijst komen we terug. Naast de inhoudelijke afstemming is de vorm van de intelligentietest fundamenteel verschillend van die van de vragenlijst. De intelligentietest vraagt om objectief scorebare prestaties en is daarom een echte test. Een vragenlijst vraagt naar stijlen van handelen, stelt niet op de proef. In de instructie wordt gezegd dat er geen goede of slechte antwoorden zijn. Daar komt nog bij dat de vragenlijst een zelfbeoordelingsinstrument is. Het is bekend dat enkelvoudige beoordelingen onbetrouwbaar zijn (Hofstee, 1999). Om de betrouwbaarheid op te voeren zou je ook nog beoordelingen van anderen kunnen inwinnen, maar dat is vaak niet eenvoudig en maakt deze methode duur.

De beperkte voorspellende waarde van vragenlijsten lijkt niet te liggen aan de neiging van mensen zich op gunstiger wijze voor te doen dan ze zijn. Of, beter gezegd, aan verschillen tussen mensen in die neiging, want als iedereen in dezelfde mate die neiging had, zou dat immers de validiteit niet noodzakelijk hoeven te beïnvloeden.

### **Wat maakt persoonlijkheidsvragenlijsten fascinerend?**

Persoonlijkheidslijsten fascineren. De vragen worden door iedereen herkend. Er is een sterke interesse bij mensen voor hun karakterstructuur. Ook werkgevers zijn geïnteresseerd. Vragenlijsten lijken heel valide omdat ze in verband te brengen zijn met veel productiviteitsgerelateerde aspecten (initiatief, doorzetten, besluitvaardigheid, durf) én met

---

<sup>2</sup> De berekeningen zijn uitgevoerd met het BCG-utiliteitsmodel (genoemd naar Brogden, Cronbach & Gleser, zie Cascio, 1999; Hofstee, 1999; Van der Maesen, 1992) en uitbreidingen. Verondersteld is beide instrumenten evenveel kosten (namelijk € 200 per persoon) en dat de organisatie 20% van de sollicitanten aanneemt, in volgorde van hun score ("top-down"). Het rendement kan groter of kleiner zijn dan de berekende waarden afhankelijk van de kwaliteit van de werving en de mate waarin de organisatie selectief kan zijn.

aspecten die direct of indirect de productiviteit van de persoon zelf of van die van anderen beïnvloeden (empathie, stabiliteit, sociale invloed, sociabiliteit). Er lijkt met persoonlijkheidsvragenlijsten veel meer te zeggen over een persoon dan met een intelligentietest. Veel organisaties en bedrijven werken met competentie-management met de daaraan verbonden vocabularia van competentie-terminen. Het overgrote deel van die competenties gaat over karaktertrekken, niet over cognitieve capaciteiten. Vragenlijsten, bijvoorbeeld de "big five", zijn gemakkelijk aan die competenties te koppelen. Sommige adviesbureaus leveren vragenlijsten waarmee, als de klant het wil betalen, dikke rapporten zijn uit te draaien.

Vragenlijsten profiteren ook van de emotionaliteitscultus die uit de VS is komen overwaaien. Auteurs als Daniel Goleman beweren dat in het bedrijfsleven emotionele intelligentie veel belangrijker is dan analytische intelligentie. In het kielzog van Goleman zijn allerlei emotionele intelligentietests op de markt gekomen. Nadere inspectie van deze tests leert dat het niet gaat om tests maar om vragenlijsten met uitslagen die herleidbaar zijn tot de "big five".

### **Wat maakt intelligentietests minder populair?**

Die impopulariteit zal zeker liggen aan de schoolsheid van de methode ("sometjes maken"). De taken in een intelligentietest niet lijken op wat er in het werk gebeurt. Voor mensen met werkervaring is de intelligentietest een straf. Er is echter in de loop van de jaren nauwelijks verandering gekomen in de inhoud van intelligentietest. De papier-en-potloodtest en de juffrouw met de stopwatch zijn langzamerhand wel weggeautomatiseerd. Dat de inhoud hetzelfde is gebleven zegt aan de andere kant ook iets over de kracht van het product. Een goed recept moet je niet vervangen. De intelligentietest is wel het meest succesvolle product genoemd dat de toegepaste psychologie heeft voortgebracht.

Serieuze nadelen van de intelligentietest zijn de geringe bruikbaarheid bij oudere mensen en bij mensen van allochtone komaf. Succesvol intelligent handelen is vaak op twee manieren te doen: snel verwerken van informatie of in verband brengen van informatie met reeds aanwezige kennis. In computerterminen is de eerste manier te vergelijken met krachtige hardware (aantal megabytes RAM, snelheid van de processor) en de tweede met kennisbestanden en slimme software die de toegang tot die kennisbestanden regelt. Op oudere leeftijd gaat de hardware achteruit en neemt het aandeel van software in het intelligent handelen toe. Intelligentietests spreken vooral de hardware aan. Oudere mensen zullen daarop minder goed presteren.

Sommige allochtone kandidaten zijn in het nadeel omdat in de test de instructies en opgaven aan de kandidaat worden gecommuniceerd in de vorm van taal.

Rynes, Colbert en Brown (2002) geven in hun artikel aan dat het positieve nieuws over intelligentietests, namelijk hun hoge validiteit, managers, personeelsmensen en ook adviseurs niet bereikt omdat die de academische literatuur niet lezen waarin dit nieuws gepubliceerd wordt. Academics van hun kant komen er niet toe om te publiceren in tijdschriften die mensen uit de praktijk lezen.

### **Ontwikkelingen**

Er blijven interessante vragen over waarop ik een tentatief antwoord wil geven.

In de eerste plaats: is de validiteit van vragenlijsten te verhogen? In elk geval niet door het aantal vragenlijsten te verdubbelen of verdriedubbelen. Er zijn aanwijzingen, ook genoemd door Hough en Furnham, dat er toch verbetermogelijkheden zijn. Ze zoeken het dan in de specialisatie: door vragenlijsten op maat samen te stellen en de inhoud precies af te stemmen op competentiedomeinen is validiteitswinst te bereiken. De auteurs noemen onder andere de domeinen van integriteit en van managementeffectiviteit, met validiteiten van respectievelijk .36 en .42.

Een tweede vraag is of er intelligentietests zijn te maken met een vorm die meer acceptabel is voor sollicitanten en voor werkgevers. Er zijn simulatieproeven op de markt die een hoge lading hebben op cognitieve informatieverwerking en die toch acceptabel zijn, bijvoorbeeld geautomatiseerde In-baskets (postbakoefeningen). Drasgow (2003) ziet veel in situationele

beoordelingstests (Engels: Situational Judgment Tests, SJT's)<sup>3</sup> als mogelijke vervangers van intelligentietests voor de ervaren sollicitant.

### Stellingen

1. Het gebruik van persoonlijkheidsvragenlijsten bij competentie management is niet in overeenstemming met de beperkte validiteit van dat instrument.
2. Het hanteren van uitgebreide competentie vocabulaires werkt het gebruik van persoonlijkheidsvragenlijsten in de hand.
3. De validiteit van intelligentietests is in de wetenschappelijke A&O psychologische literatuur onbetwist. Weinig managers en P&O-adviseurs komen echter met deze literatuur in aanraking.
4. Uitdaging 1: het ontwikkelen van gespecialiseerde vragenlijsten voor speciale (groepen van) functies of voor speciale gedragsdomeinen.
5. Uitdaging 2: het ontwikkelen van intelligentietests die beter geaccepteerd worden dan de traditionele intelligentietests en die de capaciteiten van oudere personen en van personen van allochtone afkomst beter tot hun recht doen komen.

### Referenties

Cascio, W.F. (1999). *Costing human resources. The financial impact of behavior in organizations*, South Western College Publishing

Drasgow, F. (2003). Intelligence and the workplace, In: W.C. Borman, D.R. Ilgen, R.J. Klimoski, & I.B. Weiner (Eds.) *Handbook of Psychology, Volume 12, Industrial and Organizational Psychology*, pp. 107-130, NY: John Wiley & Sons, Inc.

Hofstee, W.K.B. (1999). *Principes van beoordeling. Methodiek en ethiek van selectie, examineren en evaluatie*, Lisse: Swets & Zeitlinger

Hough, L. & Furnham, A. (2003). Use of personality variables in work settings. In: W.C. Borman, D.R. Ilgen, R.J. Klimoski, & I.B. Weiner (Eds.) *Handbook of Psychology, Volume 12, Industrial and Organizational Psychology*, pp. 131-169, NY: John Wiley & Sons, Inc

Maesen de Sombreff, P.E.A.M. van der (1992). *Het rendement van personeelsselectie*, Proefschrift Rijks Universiteit Groningen

Maesen de Sombreff, P.E.A.M. van der, Born, M.Ph., Van Oudenhoven-van der Zee, K.I., & Ruhe, D.R. (2003). Situationele beoordelingstests in de schijnwerpers, *De Psycholoog*, 38, 2, 58-67

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A., and Braverman, E.P. (2001). Predicting job performance using situational judgment tests: A clarification of the literature. *Journal of Applied Psychology*, 86, 4, 730-740.

---

<sup>3</sup> In een situationele beoordelingstest (McDaniel, Morgeson, Finnegan, Campion & Braverman, 2001; Van der Maesen, Born, Van Oudenhoven-Van der Zee & Ruhe, 2003) wordt een praktijkprobleem geschetst en een aantal oplossingen van dat probleem. SJT's worden in toenemende mate in bewegende beelden en geluid (multimediale vorm) gepresenteerd. De taak van de sollicitant bestaat uit het beoordelen van de effectiviteit van die oplossingen. Om de antwoorden te scoren worden die antwoorden vergeleken met de oordelen van ervaren mensen die een goede reputatie hebben op de gemeten vaardigheid. SJT's zijn "hot" in de VS omdat ze zeer valide blijken te zijn, omdat ze geaccepteerd worden als fair en interessant en omdat allochtone kandidaten minder scoreachterstand hebben op SJT's. Ook lijken SJT's geschikt om praktische en sociale intelligentie te benaderen.

Rynes, S.L., Colbert, A.E., & Brown, K.G. (2002). HR professionals beliefs about effective human resource practices: correspondence between research and practice. *Human Resource Management, 41, 2, 149-174*

Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124, 262-274*

**Auteur**

Dr. Paul van der Maesen de Sombreff is A&O-psycholoog en eigenaar van Van der Maesen Advies voor Personeelsmanagement in Den Haag. Het bureau is gespecialiseerd in innovatieve assessmentinstrumenten.

Website: [www.vandermaesen.nl](http://www.vandermaesen.nl)

E-mail: [advies@vandermaesen.com](mailto:advies@vandermaesen.com)