

NOA



vrije Universiteit amsterdam



Oratie

Keuren en Kleuren

Marise Born



Giuseppe Arcimboldo (1527-1593) *Air*, ca. 1566

Omslagillustratie:

Lucas Cranach de Oude (1472-1553) Maarten Luther, 1529

Hans Holbein de Jonge (1497-1543) Johannes Calvijn, ca. 1540

Giuseppe Arcimboldo (1527-1593) *Air*, ca. 1566 (fragment)

Ontwerp:

Eveline Osseweijer

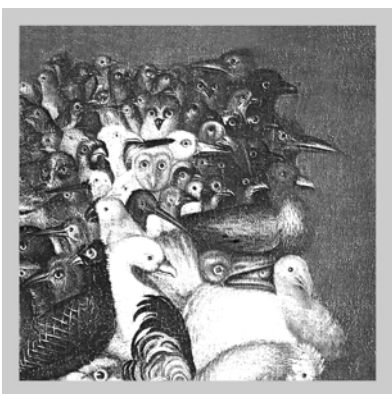
KEUREN EN KLEUREN

Rede

in verkorte vorm uitgesproken
ter gelegenheid van het aanvaarden
van het ambt van bijzonder hoogleraar
aan de Faculteit der Psychologie en Pedagogiek
van de Vrije Universiteit Amsterdam
vanwege de Stichting Het Vrije Universiteitsfonds
met de leeropdracht Arbeids- en personeelspsychologie
in het bijzonder gericht op allochtoon-etnische groepen
op 31 oktober 2008

door

Maria Philippine Born



1. Inleiding

Mijnheer de Rector Magnificus,
Zeer gewaardeerde toehoorders,

De Vrije Universiteit geeft mij onmiskenbaar een nestgevoel. Op deze vertrouwde plek zette ik mijn eerste universitaire schreden. In deze aula promoveerde ik. Mijn echtgenoot

ontmoette ik binnen de traditionele VU-geledingen en op de studentenvereniging ontstonden onze normen en waarden, met Balkenende aan het bier. Voor ons waren allochtonen toen nog Zeeuwen. Onze normen en waarden ontsproten uit de protestants-christelijke traditie. Het waren de normen en waarden die de drijfveer vormden voor de oprichting van deze universiteit.

Binnen de psychologische wetenschap is vooral het gedachtegoed over waarden en normen van sociaal psycholoog Milton Rokeach zeer invloedrijk geweest. In zijn boek *The nature of human values* uit 1973 stelde hij dat waarden de brede en relatief stabiele idealen en doelen zijn die een samenleving of groep als nastrevenswaardig beschouwt. Door het naleven van de erbij horende normen worden deze waarden nagestreefd. Hij zag waarden als het samenbindende concept voor alle wetenschappen die zich met het gedrag van mensen in hun sociale context bezighouden. Van die waarden stelde hij vrijheid en gelijkheid centraal (Wilson, 2004).

In velerlei toonaarden komen beide waarden uit dit oorspronkelijke werk terug in later psychologisch onderzoek. Dit is onder andere het geval in het werk van organisatiepsycholoog Geert Hofstede (1980, 2001), die een omvangrijke studie deed in IBM-vestigingen in meer dan 70 landen. Met behulp van zijn resultaten kon hij culturen aan de hand van vijf dimensies typeren. Er zijn er twee die de kern vormen. De ene is de dimensie *individualisme-collectivisme*, die ernaar verwijst in hoeverre individuen al dan niet geïntegreerd zijn in groepen. De ander is de dimensie *machtsafstand*, die aangeeft in

hoeverre een ongelijke verdeling van macht in een cultuur wordt geaccepteerd (Triandis, 2004). In de dimensie individualisme-collectivisme is Rokeachs waarde van vrijheid duidelijk zichtbaar, en in de dimensie machtsafstand, de mate van ongelijkheid of gelijkheid onder mensen die door een cultuur als normaal wordt beschouwd, is zijn waarde van gelijkheid goed te herkennen.

De psychologie wil het menselijk gedrag kunnen begrijpen en voorspellen. Psychologisch onderzoek laat zien dat waarden het gedrag kunnen voorspellen in een aantal belangrijke domeinen, zoals consumentengedrag en politiek stemgedrag (Bardi, Calogero, & Mullen, 2008). Wat betreft het werkdomein, kunnen werknemers worden bevorderd op grond van hun anciënniteit of op grond van hun individuele werkprestaties. Medewerkers in het Verenigd Koninkrijk, een individualistisch land, staan positiever tegenover beleid voor bevorderingen dat gebaseerd is op individuele werkprestaties dan hun meer gemeenschapsgeoriënteerde Duitse collega's (Fischer & Smith, 2004). In het collectivistische Japan en Taiwan hangt de beslissing om sollicitanten aan te nemen sterker af van de inschatting of deze goed zullen kunnen samenwerken met de reeds in de organisatie aanwezige werknemers dan van hun vakbekwaamheid. Mexico is een maatschappij met een grote machtsongelijkheid. Daar vormt het aantal strategisch goede connecties dat een kandidaat heeft één van de drie belangrijkste criteria om diegene aan te nemen (Huo, Huang, & Napier, 2002). Als laatste voorbeeld noem ik de strikte hiërarchie van de Turkse politie. Daar ligt de toekomst van een sollicitant uiteindelijk slechts in één paar

handen, namelijk die van de hoogste officier in de selectiecommissie (Calgar, 2004).

Welke soms onuitgesproken opvattingen en normen lijken een rol te spelen bij het beoordelen van mensen in ons eigen multiculturele Nederland? In dat land is bijna een kwart van de bevolking allochtoon, en van dat kwart heeft meer dan de helft een niet-westerse afkomst (CBS Statline, 2008). Inmiddels wordt er een bont gezelschap van mensen van diverse etnisch-culturele pluimage voor werkgerelateerde doeleinden gekeurd.

Mijn rede gaat over dit beoordelingsproces en welke factoren daarbij zoal een rol kunnen spelen. Het is van groot maatschappelijk belang dat dit beoordelingsproces niet alleen voor autochtone groepen geschikt is, maar ook voor de aanzienlijke groepen mensen van etnisch-allochtoon afkomst in ons land, en dat testscores en uitspraken over allochtone Nederlanders dezelfde betekenis hebben als voor autochtone Nederlanders.

Met deze rede aanvaard ik graag de bijzondere leerstoel arbeids- en personeelspsychologie, die zich vooral richt op allochtoon-etnische groepen. De arbeids- en personeelspsychologie is onderdeel van de arbeids- en organisatiepsychologie, de psychologie die zich richt op het gedrag van mensen in arbeidsorganisaties (cf. Born, 2007; McGonigle, O'Shea, Curnow, & Brown, 2004). De klassieke kerntaken van testontwikkeling, selectie en assessment zijn het werkgebied van de arbeids- en personeelspsychologie. Het gaat daarbij onder andere om het ontwikkelen van intelligentie- en

persoonlijkheidstests, het trainen van personeel, het meten van werkprestaties en werktevredenheid, en het identificeren van de meest gekwalificeerde medewerkers voor een promotie.

Hoe staat het met de deskundigheid van de arbeids- en personeelspsychologie in deze taken? Kan het imago van vakbekwaamheid op dit domein nog altijd worden waargemaakt, of wordt die deskundigheid langzamerhand obsoleet? Hoe zit het met adviezen die aan sollicitatiecommissies worden gegeven over selectieprocedures in de praktijk van alledag, en die vaak in handen zijn van onbedreven *ad hoc* selecteurs? Mijn rede gaat over een aantal veranderingen in onze werkelijkheid die deze toegepaste tak van de psychologie tot nadenken dwingt. De arbeidsmarkt van nu is in vergelijking met de tijd dat mijn ouders de arbeidsmarkt betraden veel diverser geworden. Bij een veranderende werkelijkheid, moet de arbeids- en personeelspsychologie blijven onderzoeken of haar empirisch onderbouwde bevindingen nog wel geldig zijn. Alleen op die manier kan ze de aspiratie waar blijven maken om verantwoorde metingen en voorspellingen te doen.

Ik spits mijn betoog toe op de etnisch-culturele diversiteit van onze samenleving en de betekenis hiervan voor het beoordelen van mensen voor werkgerelateerde doeleinden. Het accent ligt daarbij vooral op aspecten van werving en selectie, dat wil zeggen, daar waar een differentiatie wordt aangebracht tussen geschikte en minder geschikte mensen. Bij dit aspect van de personeelspsychologie zijn de invloeden van multiculturaliteit bij de beoordeling het sterkst (Andriessen, Dagevos, Nievers, & Boog, 2007).

In deze rede bekijk ik eerst de manieren waarop het beoordelen van mensen van allerlei afkomst samenhangt met waarden en ethische regels in ons land die betrekking hebben op dit beoordelingsproces. Ook zullen we zien hoe dit beoordelingsproces te maken heeft met integriteit en moreel handelen in arbeidsorganisaties.

Ten tweede neem ik een aantal onderdelen van het beoordelingsproces in ogenschouw. Dit thema neemt een centrale plaats in het betoog in want er valt relatief veel over te zeggen. Welke effecten heeft het gebruik van de beoordelingsmethoden waar we ons tot nu toe altijd zo behaaglijk bij voelden in een heterogene groep sollicitanten? Hoe druk moeten we ons bovendien maken over de subjectiviteit die altijd deels in het beoordelingsproces zit, en de daarmee mogelijk gepaard gaande partijdigheid?

Ten derde schenk ik apart aandacht aan het proces van zelfbeoordeling. Meestal bepalen we de persoonlijkheids-eigenschappen van mensen in de psychologie met behulp van zelfbeoordelingen. We kunnen er bij persoonlijkheidsvragenlijsten niet altijd automatisch gerust op zijn dat er geen vertekeningen optreden als mensen uit verschillende etnisch-culturele groepen deze vragenlijsten invullen. Zulke vertekeningen hangen wellicht samen met cultuur-specifieke formuleringen in dit type tests, zoals typisch Nederlandse uitdrukkingen en gezegden die niet goed begrepen worden, maar ook met verschillen in interpretatie van de testvragen (Bochhah, Kort, Seddik, & Van de Vijver, 2001).

Tot slot vertel ik in dit betoog wat er met deze kennis kan worden gedaan en wat ikzelf voor ogen heb.

Uit selectiepsychologisch onderzoek komt systematisch naar voren dat de test die werkprestaties het best voorspelt, de intelligentietest, tegelijkertijd ook de grootste scoreverschillen laat zien tussen de etnisch autochtone en etnisch allochtone groepen. Deze scoreverschillen nemen overigens af in de tweede generatie allochtonen. Er is een aantal concrete strategieën voorhanden die een oplossing kunnen bieden bij het streven naar een divers personeelsbestand zonder dat er concessies hoeven te worden gedaan aan het rendement van personeelsselectie (Ployhart & Holtz, 2008). Over de precieze effecten van iemands etnisch-culturele achtergrond op resultaten op een persoonlijkheidsvragenlijst bestaat tot nu toe echter veel minder duidelijkheid.



2. Integriteit

Van het belang van culturele waarden voor de werkgerelateerde beoordeling van personen gaf ik al enkele voorbeelden. Ik spits het betoog nu toe op morele waarden en normen. Hoe integer wordt er gehandeld in het beoordelingsproces door professionele adviesbureaus en binnen arbeidsorganisaties? Is dat handelen in overeenstemming met relevante morele normen, gedragscodes en waarden zoals die gelden in de maatschappij (McFall, 1987)? Laten we eerst eens

kijken naar de wereld van het professionele adviesbureau, en vervolgens hoe het eraan toegaat in arbeidsorganisaties. We kijken met de blik van de ibis, die eens de morele wetten symboliseerde.

In vervlogen tijden, in september 1962, gingen mijn ouders naar het Instituut van professor Jan Waterink, het Laboratorium voor Toegepaste Psychologie op de hoek van de Vossiusstraat en de Van Baerlestraat. Dit laboratorium herbergde sinds 1946 het psychotechnische, bedrijfs-psychologische en beroepskeuzewerk van Waterink, dat voor die tijd onderdeel was geweest van zijn Psychotechnisch Laboratorium (het tegenwoordige LTP Advies te Amstelveen). Ze zouden die dag een aantal tests ondergaan voor hun uitzending naar Pakistan vanuit de Gereformeerde Kerken en ervoeren dit bezoek als een ware *happening*. Voor de gelegenheid had mijn moeder een roze truitje aan. Het werd voor hen een wonderbaarlijk en magisch bezoek. Langs een wand met uitstekende knoppen moesten ze een touw leggen over een voorgetekend patroon van lijnen. Er werd bekeken of ze dat spoor konden blijven volgen. Daarna deden ze een rangeertest waarbij ze spoortreintjes op de juiste plaats moesten manoeuvreren; daarbij werd ook de stopwatch gehanteerd om te zien hoe lang ze daarover deden. Naast deze praktische tests kregen ze de Rorschach test, daterend uit 1921. In deze opmerkelijke test kregen ze een vlek voorgelegd en moesten ze opschrijven wat ze daarin herkenden. Waar zouden ze toen aan gedacht hebben? Ook was er een gesprek met enkele deskundigen en moesten ze op een serie ethisch-

morele kwesties reageren. Ze hoorden daarna dat ze geschikt werden bevonden en stapten aan boord van een groot schip naar Pakistan. Hoewel ze recht hadden op een gesprek over de uitslag maakten ze geen gebruik van dit recht op informatie in de drukte van dat moment. Op de integriteit van vooral de Rorschachtest valt inmiddels heel wat af te dingen. In het testbeoordelingssysteem van het Nederlands Instituut van Psychologen, het NIP (Evers, Van Vliet-Mulder, & Groot, 2000) wordt de test als eerste in de categorie persoonlijkheidstests besproken. Daaruit blijkt dat hij werkelijk van geen kanten deugt.

Het Instituut van Waterink handelde alleszins oprecht naar de maatstaven van een halve eeuw geleden, maar serieus empirisch onderzoek naar de kwaliteit van tests, laat staan naar hun geschiktheid voor allerlei etnisch-culturele groepen, speelde toen nog nauwelijks een rol.

In de huidige tijd komen praktische tests zoals touwen leggen en spoortreintjes rangeren niet meer voor in de standaard-gereedschapskist van de psycholoog. De Rorschach test is bij de serieuze adviesbureaus gelukkig verdwenen, hoewel het opvalt dat hij bij onze Belgische bureaus nog altijd voorkomt in hun lijstje van meest gebruikte tests (Muñiz & Bartram, 2007). Dat deze tests ondertussen bij ons niet meer worden gebruikt, heeft alles te maken met een aantal strakker geformuleerde rechten van sollicitanten die af te leiden zijn uit de ethische richtlijnen van de Beroepscode voor psychologen van het NIP (2007), de Algemene Standaard Testgebruik van het NIP (2004), en de sollicitatiecode van de Nederlandse Vereniging

voor Personeelsmanagement en Organisatieontwikkeling, de NVP (2006). Naast het recht op informatie, op een vertrouwelijke behandeling van de resultaten, en op het hebben van een eerlijke kans, hebben sollicitanten ook het recht op een instrumenteel doelmatige procedure, dat wil zeggen een procedure met relevante selectiemethoden. Het gebruik van wetenschappelijk niet aanvaardbare en dus ondeugdelijke methoden, zoals de Rorschach test, is ondeskundig en getuigt niet van integer handelen. Het is frappant dat nog maar pas geleden, in 2005, voor het eerst door het NIP in samenwerking met het Landelijk Bureau ter bestrijding van Rassendiscriminatie een aantal broodnodige richtlijnen gepubliceerd zijn voor het gebruik van diagnostische instrumenten bij etnische minderheden (Bochhah, Kort, & Seddik, 2005a). Het blijft echter nog altijd een grote tekortkoming in de Nederlandse situatie blijft echter nog altijd het feit dat voor het overgrote deel van de bestaande Nederlandse tests en vragenlijsten niet of nauwelijks solide empirische informatie beschikbaar is over de toepasbaarheid bij etnische minderheden (Bochhah, Kort, & Seddik, 2005b).

Naast aangescherpte ethische richtlijnen, is ook de kandidaat in het huidige Nederlandse tijdsgewricht veranderd. Deze dwingt de beoordelaar meer dan vroeger tot verantwoording. Deze verandering wordt beschreven door Bas Kok in zijn vorig jaar verschenen debuutroman *Het Lab*, waarvoor hij zich heeft laten inspireren door de geschiedenis van het Instituut van Jan Waterink. Hij schrijft: "We zaten niet langer tegenover onze gesprekspartner, we zaten mét hem aan tafel (p. 156)". Over

een kandidaat die duidelijk slecht op de tests had gepresteerd schrijft hij vervolgens: "Toen ik hem meedeelde dat mijn advies negatief uitviel, trok hij fel van leer. ... 'Dit is allemaal onzin, het kost tijd en geld, flauwekul is het.' ... Ik wil dat u het rapport aanpast" (p. 157). Mondiger en meer assertieve kandidaten passen goed in ons egalitaire land waar men wars is van machtsongelijkheid en waar een kandidaat soms zelfs een bloemetje meeneemt voor de psycholoog als dank voor een positief uitgebracht advies aan een organisatie (Kok, p. 157). Kandidaten kunnen bovendien sinds 1994 hun stem laten horen bij de Commissie Gelijke Behandeling (CGB) en de antidiscriminatiebureaus (ADB's), waar ze een klacht kunnen indienen als ze zich ongelijk behandeld voelen. Hoe snel allochtone kandidaten een klacht in zullen dienen is nog maar de vraag, gezien de gezagsgetrouwheid die in verschillende culturele groepen sterk aanwezig is. Uit onderzoek van Van den Berg en Evers (2006) blijkt dat ongeveer driekwart van de ondervraagden zei dat ze door hen ervaren discriminatie op grond van etnische afkomst nooit meldden.

Professionele adviesbureaus geven hun aanbevelingen gewoonlijk aan een organisatie. Deze zal vervolgens, gebaseerd op de ontvangen adviezen, zelf een keuze uit de sollicitanten maken. Hoe integer wordt er gehandeld binnen arbeidsorganisaties?

In taxonomieën van integriteit en ethische problemen in organisaties komen kwesties voor als omkoping en fraude, zoals het beruchte *Enron* schandaal van een aantal jaren geleden. Ook komen er ethische kwesties in voor die te maken

hebben met het selecteren van personeel (Geva, 2006). In de traditie van Kohlbergs (1969) belangrijke theorie van moreel redeneren, die later door zijn leerling Rest (1975) werd verfijnd, werd onlangs een test ontwikkeld voor het morele oordeelsvermogen van managers, de zogenaamde Measure of Managerial Moral Judgment (Loviscky, Trevino, & Jacobs, 2007). Deze test bevat een scenario over personeelsselectie als één van de zes meest typerende ethische situaties die een manager kan tegenkomen. De meerderheid van bedrijfs-ethische kwesties gaat over eerlijkheid, dat wil zeggen eerlijkheid in procedures en eerlijkheid in de verdeling van middelen, bijvoorbeeld wie de baan krijgt. Antwoorden op deze test worden ondergebracht in Kohlbergs stadia van moreel redeneren. Het kan gaan om zogeheten conventionele redeneringen die uitgaan van argumenten van orde en gezag. Ook kan het gaan om andere, post-conventionele redeneringen die uitgaan van universele ethische waarden. Degenen die redeneren op het post-conventionele niveau bezitten een sterk ontwikkeld waardensysteem dat is gebaseerd op het handelen naar wat het meest eerlijk is vanuit een sociaal-maatschappelijk perspectief. Dit kan bijvoorbeeld betrekking hebben op de mate waarin een lage participatiegraad en werkloosheid onder bepaalde etnische groeperingen onaanvaardbaar is.

Opmerkelijk genoeg blijkt dat het morele redeneerniveau van managers bij beslissingen op het werk wat lager uitvalt dan bij beslissingen die ze in het algemene leven nemen (Weber, 1990). Deze bevinding was onlangs nog zeer herkenbaar in het stukje in de rubriek ik@nrc.nl op de achterpagina van de NRC van 5 september jl. In dit stukje

vertelt de schrijver dat hij in de sollicitatiecommissie zat van een groot bedrijf voor een managementvacature. Hij hoorde de voorzitter van de commissie zeggen dat de kandidaat flexibel, creatief, communicatief, stressbestendig, en een goede teamplayer moest zijn. Hij zag de voorzitter met een vertwijfelde blik naar de dertig centimeter hoge stapel brieven kijken en deze in twee gelijke helften verdelen. De ene stapel van de tafel vegend zei hij: Mensen zonder geluk kunnen we niet gebruiken (J.W. Tamminga, 2008). Het was natuurlijk waarschijnlijk een grap, maar stel dat het echt zo is gegaan, dan werden daarmee de kansen van een groot aantal sollicitanten op een baan binnen een seconde volstrekt tot nul gereduceerd, waaronder wellicht die van de enkele allochtone kandidaat.

Hoe komt het dat anders zo gewone personen soms, bijvoorbeeld in een selectiecommissie, in staat zijn om niet-integer gedrag te vertonen zonder dat ze last hebben van enige vorm van schuldgevoel of zelfcensuur? Volgens de bekende psycholoog Bandura (1990) kan dergelijk gedrag worden verklaard door zogeheten moreel disengagement. Deze term is wat lastig te vertalen en verwijst naar het ontkoppelen van het eigen gedrag en een moreel oordeel daarover. Dit proces van ontkoppelen gaat via een aantal cognitieve mechanismen. Een voorbeeld van zo'n mechanisme is morele rechtvaardiging van het eigen handelen, te herkennen in een uitspraak als 'het contracteren van kinderen in ontwikkelingslanden voorkomt dat ze nog veel gevaarlijker werk moeten doen om hun wanhopige familie financieel te helpen'. Een ander mechanisme is het verdraaien van de gevolgen van het eigen

gedrag, bijvoorbeeld door te zeggen dat 'iemand beledigen diegene absoluut niet zal beschadigen'.

Moreel disengagement kan de vorm krijgen van het moreel uitsluiten van een bepaalde groep (Opatow, 1990a; Opatow, 1990b; Trevino, Weaver, & Reynolds, 2006). Bij morele uitsluiting laat men sommige mensen 'buiten de grenzen vallen waarbinnen morele waarden en normen en overwegingen van eerlijkheid gelden' (Opatow, 1990b). Sommige individuen of groepen worden op deze manier gemakkelijk benadeeld omdat ze niet tot de eigen groep behoren en dus anders zijn, en daardoor sneller buitengesloten worden van morele overwegingen. Hoe snel mensen geneigd zijn om in groepen te denken en over te gaan in wij-versus-zij denken weten we allemaal (Brewer, 2007). Het woord allochtoon maakt dit verschijnsel ogenblikkelijk duidelijk; de term is afkomstig van het Griekse woord *allos*, en betekent *anders*.

Moreel disengagement blijkt vaker voor te komen bij personen die minder empathisch van aard zijn. Deze mensen zijn minder in staat zich in een ander te verplaatsen, en van nature meer wantrouwend tegenover anderen. Het komt ook vaker voor bij mensen die morele kwesties minder belangrijk vinden voor hun zelfbeeld, dat wil zeggen voor hun notie over wie ze zelf zijn (Detert, Trevino, & Sweitzer, 2008).

Mogelijk is het optreden van moreel disengagement één van de verklaringen voor het feit dat het klakkeloos afwijzen van allochtone sollicitanten een groot taboe is – het staat in de Nederlandse taboe-top-honderd op de 40ste plaats (Maassen &

Oosterwijk, 2006) – maar dat het toch gebeurt. Uit psychologisch en sociologisch onderzoek blijft naamdiscriminatie systematisch naar voren komen. Deze vorm van discriminatie bleek pas weer overduidelijk uit een studie van collega Deros (2007). Zij voerde een zogenaamde praktijktest uit. Zij verstuurde 744 fictieve cv's naar 186 bestaande vacatures in het midden- en kleinbedrijf in de Rotterdamse regio, die ofwel van een typisch Marokkaanse ofwel van een typisch Nederlands klinkende kandidaat afkomstig waren. De kandidaten met een Marokkaans klinkende naam ontvingen significant meer afwijzingen (91%) dan de kandidaten met een typisch Nederlands klinkende naam (74%) op een verder volstrekt gelijkwaardig cv. Selecteurs belijden met de mond dat ze geen onderscheid maken. Toch blijken ze in hun selectiedaden wel degelijk blanke sollicitanten te bevoordelen boven zwarte, zelfs als de blanke veroordeeld is voor drugshandel en de zwarte niet, zo blijkt ook weer uit recent Amerikaans onderzoek (Pager, 2003; Pager & Quillian, 2005).

De huidige psychologische modellen over racisme en vooroordeel gaan uit van de aanwezigheid van een zogenaamd aversief racisme. Deze term doelt op een hedendaagse vorm van racisme en vooroordeel waarbij mensen een onaangenaam intern conflict voelen. Enerzijds is er een oprechte behoefte om zichzelf te zien als een niet-bevooroordeeld en onpartijdig iemand. Anderzijds ervaren mensen ongewilde negatieve gevoelens tegenover etnische minderheden (Dovidio & Gaertner, 1998; Nail, Harton, & Barnes, 2008).

Dan toch maar anoniem solliciteren? Het in 2006 gestarte experiment van de gemeente Nijmegen naar anoniem solliciteren, is afgelopen april nota bene alweer gestaakt. Het anoniem solliciteren bleek helaas niet tot het aannemen van relatief meer allochtone kandidaten te leiden. Bij de Belgische federale overheid wordt nog wel steeds anoniem gesolliciteerd en worden de effecten pas op lange termijn bekeken. Ondertussen blijft de discussie tussen voor- en tegenstanders aanhouden.

Tot nu toe sprak ik over moraliteit en integer handelen van diegenen die beoordelingen uitvoeren. We zagen enerzijds dat er veel goede intenties, richtlijnen en gedragsregels zijn ontwikkeld. Anderzijds zagen we ook dat het daadwerkelijk gedrag van beoordelaars soms nog bijzonder weerbarstig lijkt. Bovendien bleek dat over het merendeel van de bestaande Nederlandse tests en vragenlijsten die in beoordelingsprocessen worden gebruikt niet of nauwelijks informatie beschikbaar is over de toepasbaarheid voor etnische minderheden. Het is in mijn ogen daarom belangrijk dat het systeem ter beoordeling van Nederlandstalige tests van de Commissie Testaangelegenheden van het Nederlands Instituut van Psychologen (COTAN), een aparte kwaliteitseis over deze toepasbaarheid opneemt.

In het beoordelingsproces spelen naast ethische regels en integriteit van de betrokkenen ook allerlei andere zaken een rol, zoals kenmerken van de pool van sollicitanten, de gehanteerde selectiemiddelen en verdere factoren waaraan

beoordelaars onderhevig kunnen zijn. We gaan eens wat beter kijken naar de verschillende aspecten van het beoordelingsproces en hoe dat werkt in het multiculturele Nederland.



3. De gekleurde keuring

De groep mensen die voor werkgerelateerde doeleinden wordt beoordeeld is in ons huidige tijdsgewricht veel heterogener dan een halve eeuw geleden. Toch loopt het percentage niet-westerse beroepsgeschikte allochtonen dat daadwerkelijk een betaalde baan heeft met 51,8% in 2007, een stuk achter op het percentage autochtonen met een betaalde baan, namelijk 68,1% in 2007. Bovendien is de werkeloosheid in de groep niet-westerse allochtonen bijna drie keer zo hoog als in de

autochtone groep en bezetten niet-westerse allochtonen over het algemeen lagere functieniveaus (CBS Statline, 2008). Het minst gunstig zijn de kansen daarbij voor Marokkanen, die in ons land onderaan de statushiërarchie van etnisch-culturele minderheden bungelen (Andriessen et al., 2007).

Belangrijke oorzaken voor de achterstand van etnische minderheden op de arbeidsmarkt zijn, zoals bekend, hun gemiddeld lagere opleidingsniveau en het feit dat ze vaker dan autochtonen voortijdig hun opleiding verlaten (Jennissen & Oudhof, 2007). Omdat de werkgelegenheidsgroei in Nederland vooral optreedt in functies voor hoogopgeleiden beperk ik me hier tot het noemen van een paar recente resultaten uit onderzoek samen met collega's die betrekking hebben op het hoger onderwijs.

Dat onderzoek laat systematisch zien dat de studieprestaties van allochtone studenten achterblijven bij die van autochtone studenten. Ze blijven achter zowel in het gemiddelde tentamencijfer als in het percentage behaalde vakken (Meeuwisse, Born, & Severiens, 2008; Meeuwisse, Severiens & Born, 2008). Voortijdige allochtone hbo-verlaters blijken vaker dan autochtone studenten helemaal te stoppen met hun opleiding als ze aangeven te staken vanwege de slechte kwaliteit van de docenten en het onderwijssysteem. Het lijkt erop dat allochtone studenten op zo'n moment niet meer verwachten dat de kwaliteit op een andere opleiding beter zal zijn. Autochtone studenten stappen in dat geval wel naar een andere opleiding over (Meeuwisse, Severiens, & Born, 2008).

De selectiemiddelen

Het academische succes, dat wil zeggen de studieprestaties, is zeker belangrijk want daar letten recruiters op als zij net-afgestudeerden aan het werven zijn voor openstaande vacatures (Cole, Rubin, Feild, & Giles, 2007; Nemanick & Clark, 2002)). Recruiters letten duidelijk ook op enkele andere zaken. Ze kijken naar werkervaring, dat wil zeggen bijbaantjes en praktijkstages, en nevenactiviteiten, zoals het bestuurlijk actief zijn binnen een universitaire vereniging. Uit zulke informatie leiden recruiters af of kandidaten over interpersoonlijke vaardigheden, leiderschapskwaliteiten en een goede werkmotivatie beschikken (Cole et al., 2007).

We vergeleken cv's van 99 autochtone en 88 allochtone studenten aan de Erasmus Universiteit (Born, Derous & Hiemstra, 2008). Op allerlei niet meteen ter zake doende vormaspecten van het cv, zoals de kwaliteit van de layout, waren geen verschillen aanwezig tussen beide groepen studenten. Uit allochtone cv's bleken echter significant lagere studieprestaties, bestuurlijke ervaring en praktijkstage-ervaring naar voren te komen. Op het vlak van werkervaring was er geen verschil. Kortom, op drie belangrijke zaken waar recruiters wel degelijk naar kijken, valt voor allochtonen wat te winnen (zie ook Van Gent, Hello, Odé, Tromp, & Stouten, 2006). Het is hierbij wel hinderlijk voor allochtone hbo-studenten dat het voor hen vaak lastiger blijkt om een praktijkstage te vinden. Zo'n stage is op het hbo dikwijls verplicht (Van Gent et al., 2006).

De vraag moet in dit verband uiteraard gesteld worden hoe valide en functiegerelateerd de kenmerken zijn waar

recruiters zich bij het beoordelen van cv's op richten. Deze functiegerelateerdheid blijkt uit eerder longitudinaal onderzoek dat een empirische relatie kon aantonen (Howard, 1986) tussen enerzijds nevenactiviteiten tijdens de studie en anderzijds interpersoonlijke vaardigheden en werkmotivatie in het latere werk. Niet alleen academische prestaties, maar ook extracurriculaire activiteiten moeten daarom meer aandacht krijgen onder allochtone groepen zodat ze een grotere kans hebben op een baan.

De geluksvogels die door de eerste *screening* komen, gaan door naar een volgende ronde waarin zeer regelmatig, onder het devies 'Een psychologische keuring is onderdeel van de sollicitatieprocedure', ook gestandaardiseerde intelligentie- en persoonlijkheidstests worden ingezet, en soms rollenspellen, en bijna altijd ook een selectiegesprek (Muñiz & Bartram, 2007). Over psychologische assessment bij allochtonen is bijzonder veel gepubliceerd (zie voor een overzicht Evers, Te Nijnhuis, & Van der Flier, 2005).

Het in Nederland meest bekende onderzoek over de toepasbaarheid van beschikbare tests voor allochtone groepen werd bijna twintig jaar geleden uitgevoerd door een werkgroep onder leiding van Willem Hofstee. Deze werkgroep hadden het NIP en het LBR samen ingesteld. In 1990 kon de werkgroep niet anders dan een subjectief oordeel vellen over de inhoud van de items, de afzonderlijke vragen van de tests. Men vond bij veel tests etnocentrische en cultuurspecifieke elementen, zoals moeilijke woorden en alleen autochtonen die

figureerden op plaatjes (Hofstee et al., 1990). Een vervolgonderzoek van Bochhah en anderen (Bochhah, Kort, Seddik, & Van de Vijver, 2001) liet zien dat deze situatie nog altijd niet noemenswaardig is verbeterd. Er zijn wel duidelijke lichtpuntjes, zoals de inspanningen die Stichting NOA verricht voor het ontwikkelen van multicultureel toepasbare tests en vragenlijsten.

In vogelvlucht geven resultaten over verschillen in gemiddelden tussen etnische groepen het volgende beeld te zien. Bij alle selectiemethoden zijn de scores van allochtone kandidaten lager dan die van autochtone kandidaten. De grootste verschillen doen zich voor op de intelligentietests en andere tests die een grote cognitieve component bevatten, de kleinste op persoonlijkheidstests (De Meijer, Born, Terlouw, & Van der Molen, 2006; Dean, Bobko, & Roth, 2008; Ployhart & Holtz, 2008; Roth, Bobko, McFarland, & Buster, 2008). Zodra het selectiegesprek gestructureerder plaatsvindt, worden de scoreverschillen kleiner (Huffcutt & Roth, 1998). In de tweede generatie allochtonen wordt bovendien het verschil met de autochtone groep kleiner (De Meijer et al., 2006; Van den Berg, 2001). Het is belangrijk om na te blijven gaan of deze ontwikkeling zich ook doorzet in de derde generatie.

Het klassieke spanningsveld voor organisaties is het probleem dat de test waarop de verschillen het grootst zijn, namelijk de intelligentietest, tegelijk ook het meest voorspellend is voor later werkgedrag. Het streven naar de meest rendabele selectie en naar een zo divers mogelijk samengesteld personeelsbestand vormt daarom vaak een

dilemma. Toch valt er op grond van een lange lijst onderzoeksresultaten inmiddels een aantal concrete adviezen te geven waarmee beide doelen zo goed mogelijk worden gediend (Ployhart & Holtz, 2008).

Eén van de meest effectieve strategieën om scoreverschillen tussen de etnisch-culturele groepen te reduceren is het gebruik van alternatieven voor de intelligentietests waarop de scoreverschillen tussen de groepen minder groot zijn. Hieronder vallen gestructureerde interviews, assessment opdrachten zoals rollenspellen, en zogenaamde situationele beoordelingstests. Het probleem van deze strategie is dat deze meestal wel tot een minder goede voorspelling van werkgedrag leidt.

De beste strategie die scoreverschillen weet te reduceren en die ook nog eens betere voorspellingen kan opleveren, is het meten bij sollicitanten van het héle arsenaal aan capaciteiten, kennis, vaardigheden en competenties dat voor een functie benodigd is.

Een andere aanpak voor het verminderen van scoreverschillen is het minimaliseren van de benodigde kennis van de Nederlandse taal om een bepaalde test te maken die heel andere competenties dan taalkennis moet meten. Uiteraard moet wel het taalniveau gegarandeerd blijven dat minimaal nodig is voor het uitvoeren van de functie. Bovendien is het bij deze strategie niet in alle gevallen gegarandeerd dat de voorspellende waarde voor het werkgedrag even goed zal blijven. Het effect van de strategie om iemands opleidingsniveau of het gemiddelde behaalde cijferniveau tijdens de opleiding te gebruiken als vervanging voor de selectietests is

minder duidelijk. Scoreverschillen tussen etnisch-culturele groepen fluctueren tussen opleidingsniveaus en de voor-spellende waarde van opleidingsniveau is minder goed dan van de intelligentietest.

Om een diverser personeelsbestand te bevorderen ligt het verder voor de hand om specifiek te werven onder allochtone doelgroepen. Ook is de steun van het management voor diversiteitsbeleid in een organisatie cruciaal. De kans dat allochtone kandidaten een aanbod voor een baan ook zullen accepteren hangt samen met hun percepties van het diversiteitsklimaat in een organisatie (Kravitz, 2008).

In het kader van de testdiagnostiek onder allochtonen kan ik het niet nalaten om even de Nederlandse grenzen te overschrijden en de in het buitenland bekende en unieke Nederlandse inburgeringstoets te noemen. Deze toets wordt sinds 15 maart 2006 verplicht afgenomen. Dit is het gevolg van het in werking treden van de Wet Inburgering Buitenland.

Selectie van mensen met een niet-Nederlandse achtergrond begint tegenwoordig al aan de poort van ons land door middel van de inburgeringstoets, het zogeheten basisexamen inburgering. Het examen wordt per telefoon in het Nederlands afgenomen. De kandidaat zit tijdens het examen achter een computer in de Nederlandse ambassade of een consulaat in het land van herkomst. De wet houdt in dat mensen die naar Nederland komen en in Nederland wonen, in eigen land al moeten aantonen een basiskennisniveau te hebben van de Nederlandse taal en hoe de Nederlandse samenleving in elkaar zit. Het doel is een betere integratie van

alle immigranten. Dit wil zeggen dat het examen een bijdrage moet leveren aan actief burgerschap in ons land door kennis van de Nederlandse taal, bewustzijn van de Nederlandse waarden, en het naleven van de Nederlandse normen (cf. Tijssen, Berntsen, De Jong, & Lennig, 2005, p.2). Nederland is het eerste land ter wereld waar immigranten van buiten de Europese Unie vóór binnenkomst zo'n toets moeten afleggen. Het examen bestaat uit een mondelinge taaltoets en de toets 'Kennis van de Nederlandse samenleving'. In Nederland zelf moet de immigrant zich overigens nog verder verdiepen in de taal en de samenleving.

In het licht van het uiteindelijke doel, namelijk een betere integratie van alle immigranten, valt het op dat sommige immigranten wel, en anderen niet het examen hoeven af te leggen. Immigranten uit bijvoorbeeld Japan, Australië en Vaticaanstad hoeven het examen niet te doen. In mei van dit jaar bracht de *Human Rights Watch*, een internationale niet-gouvernementele organisatie gericht op aanpak van schendingen van mensenrechten, een rapport uit over het examen.

De conclusie luidde dat het gebruik van dit examen discriminerend is tegen specifieke groepen niet-westerse immigranten, in de praktijk vooral Turkse en Marokkaanse gezinsvormers en -herenigers (Human Rights Watch, 2008). Het is, aldus de Human Rights Watch, niet aangetoond dat er een zeer dwingende reden is om individuen met de ene nationaliteit wel en de andere nationaliteit niet het examen af te nemen. Het vaststellen van het ontwikkelingsniveau van het land waaruit de immigrant afkomstig is, is zeker niet hetzelfde

als het toetsen van het vermogen van een individuele immigrant om te integreren.

Naast de wettelijke noodzaak van een eerlijk gebruik van examens en tests, zijn er ook een paar toetstechnische opmerkingen te maken. Net als andere toetsen met grote gevolgen voor individuen, moet de inburgeringstoets aan de belangrijke criteria van validiteit en betrouwbaarheid voldoen. Validiteit betreft de voorspellende waarde van de toets, namelijk of deze inderdaad aantoonbaar bijdraagt aan een betere inburgering. Dat is nog volstrekt onbekend. Wel is onlangs de examendrempel opgehoogd. Deze bleek volgens een rapport van TNO (Kessens & Jacobusse, 2007) te laag. Betrouwbaarheid betreft de vraag of de toets consistente resultaten oplevert. Het mag bijvoorbeeld niet uitmaken of de immigrant door een persoon wordt beoordeeld of automatisch wordt beoordeeld via een computer. Voor een steekproef van 250 kandidaten bleken het computeroordeel en het menselijk oordeel bijvoorbeeld goed inwisselbaar voor de Kennistoets Nederlandse Samenleving. Maar er waren heel weinig kandidaten die een score rond de zak/slaaggrens hadden behaald en er was al helemaal een klein percentage, namelijk 4%, dat voor het examen was gezakt (cf. Kessens & Jacobusse). Met het regeringsbesluit de examendrempel te verhogen, is het dus noodzakelijk om door te gaan met het monitoren of mens en computer inwisselbaar blijven.

Beoordelaarspartijdigheid

Bij selectie-instrumenten waarbij een beoordelaar vragen stelt, luistert, observeert, en de kandidaat vervolgens evalueert,

wordt deze vaak met achterdocht bekeken door kandidaten uit een etnische minderheidsgroep, zeker als het om een autochtone beoordelaar gaat (Andriessen et al., 2007): Hoe onpartijdig is de beoordelaar eigenlijk?

Dat deze vraag niet ten onrechte wordt gesteld blijkt wel uit een klassiek geworden sociaal psychologische studie uit 1974 van de Amerikaan Word en zijn collega's (Word, Zanna, & Cooper, 1974). Deze studie is als één van de essentiële teksten opgenomen in een overzichtswerk over stereotypen en vooroordelen (Stangor, 2000).

De onderzoekers voerden twee experimenten uit. In het eerste experiment beoordeelden blanke universiteitsstudenten in de rol van selecteur elk een zwarte en een blanke kandidaat op hun geschiktheid voor een functie. De kandidaten speelden onder een hoedje met de onderzoekers en waren van tevoren geïnstrueerd in hun rol. Bij zwarte kandidaten bleken de interviewers meer taalfouten te maken, zaten ze wat verder weg en beëindigden ze het interview sneller. Een vervolgsperiment liet zien hoe ondermijnd dergelijk gedrag van interviewers kan werken. In dit tweede experiment werden 30 blanke ongetrainde kandidaten onderworpen aan een selectiegesprek door twee door de onderzoekers geïnstrueerde blanke interviewers. De kandidaten kregen tijdens het selectiegesprek ofwel de 'blanke' behandeling van interviewers ofwel de 'zwarte'. De 'zwarte' behandeling hield in dat er meer taalfouten werden gemaakt, dat de interviewers wat verder weg zaten, en dat zij een korter interview afnamen dan bij de 'blanke' behandeling. Twee beoordelaars die niets van de

bedoelingen van het experiment wisten, kregen een video-opname te zien van de kandidaten maar niet van de interviewers. Zij beoordeelden de kandidaten die de 'zwarte' behandeling hadden ondergaan als minder competent en zenuwachtiger dan de kandidaten die de 'blanke' behandeling hadden gekregen. De conclusie van de onderzoekers was dat het niet-verbale gedrag van blanke interviewers positiever was bij blanke dan zwarte kandidaten, en dat het gedrag van interviewers de mogelijkheid voor kandidaten om competent over te komen wel degelijk kan ondermijnen.

Enige achterdocht ten aanzien van de beoordelaar is dus gerechtvaardigd. Maar als aan een autochtone sollicitant de voorkeur wordt gegeven boven een allochtone sollicitant dan hoeft er natuurlijk niet meteen sprake te zijn van discriminatie. Er kunnen immers zeker competentieverschillen tussen beide kandidaten aanwezig zijn.

Om na te gaan onder welke omstandigheden achterdocht groter kan worden, deden Miller, Turnbull, & McFarland (1989) onderzoek naar de argwaan die mensen hadden over een manager. Deze had de dubieuze reputatie discriminerend te zijn tegen een minderheidsgroep op het werk bij promoties naar een hogere baan. In dit geval ging het om vrouwen. Deelnemers aan het onderzoek moesten één van twee scenario's lezen. Elk scenario beschreef dat de manager een aantal kandidaten voor promotie een toets afnam en dat hij vervolgens meldde dat de hoogste score was behaald door een man. De scenario's waren precies hetzelfde, behalve in het

aantal kandidaten dat aan het examen had meegedaan. In het ene scenario ging het om één man en negen vrouwelijke kandidaten en in het andere om tien mannelijke en negentig vrouwelijke kandidaten. In beide gevallen had de mannelijke kandidaat onder verder gelijke omstandigheden a priori natuurlijk evenveel kans om het hoogste cijfer te behalen.

Toch bleken de deelnemers die het eerste scenario lazen, waarin één man en negen vrouwen meedongen naar de promotie, het allemaal minder te vertrouwen dan in het tweede scenario. Zij vroegen zich af of de manager de toets wel op een eerlijke manier had beoordeeld.

De verklaring voor deze grotere achterdocht heeft te maken met de opmerkelijke manier waarop mensen doorgaans denken. In het eerste scenario was er maar één mannelijke kandidaat aanwezig en dus ook maar één mogelijkheid de mannelijke kandidaat te kiezen. In het tweede scenario met tien mannelijke kandidaten, waren er dus ook tien kansen dat de mannelijke kandidaat het hoogste cijfer kon halen. In de gedachten van de deelnemers die het eerste scenario hadden gelezen, was het uitzonderlijk en dus verdacht dat nou net die ene man tussen de kandidaten het hoogste cijfer had behaald, terwijl in de gedachten van de deelnemers die het tweede scenario hadden gelezen er ook negen andere mannen het hoogste cijfer hadden kunnen behalen. Hierdoor vonden zij het minder vreemd dat een man dit hoogste cijfer had behaald. Anders gezegd, zij konden de uitkomst mentaal gemakkelijker plaatsen en waren daarom minder wantrouwend.

Het mag duidelijk zijn dat er in het eerste scenario niet méér wordt gediscrimineerd door de manager dan in het

tweede geval, en dat er dus geen enkele grond is om in het eerste scenario wantrouwiger te zijn. Mensen hebben echter dikwijls geen goed besef van kansuitspraken.

Er blijven zeker wat andere redenen over om enige argwaan jegens selecteurs te houden. De altijd nog overwegend autochtone selecteurs zijn namelijk onderhevig aan uiteenlopende invloeden van cognitieve en emotionele aard. Ik noem er een paar.

Zo blijkt dat selecteurs die al ervaring hebben met allochtone werknemers gunstiger oordelen over allochtone kandidaten dan diegenen zonder enige ervaring met allochtone kandidaten (Kruisbergen & Veld, 2002). Ook geven selecteurs relatief veel gewicht aan negatieve informatie over een kandidaat. Denk maar aan de aandacht die het 'gat in het cv' steevast krijgt. Selecteurs zijn overigens zeker niet uniek in deze zogenaamde negativiteits*bias* (Vaish, Grossmann, & Woodward, 2008).

Dat een negativiteits*bias* mogelijk een extra probleem kan opleveren voor allochtone kandidaten, heeft te maken met de manier van beoordelen door selecteurs als kandidaten tijdens een selectiegesprek uitleg moeten geven over een eerdere situatie waarin ze hadden gefaald (Silvester, 1997). Selecteurs blijken relatief veel gewicht te hechten aan de verklaring die een kandidaat geeft voor zo'n faalsituatie. Ze oordelen daarbij duidelijk positiever over de kandidaat als deze zijn of haar falen voor eigen rekening neemt en het falen toeschrijft aan factoren die door hem of haar te controleren zijn. Maar als een kandidaat de omstandigheden aanwijst als oorzaak van het

falen, stellen selecteurs die verklaring minder op prijs. Deze neiging van selecteurs zou wel eens ongunstiger uit kunnen pakken voor allochtone kandidaten. Dit komt omdat westerlingen over het algemeen meer dan niet-westerlingen geneigd zijn stabiele kenmerken van mensen, zogeheten disposities, aan te dragen als oorzaak van iemands gedrag (Morris & Peng, 1994). Ook zullen de omstandigheden van allochtonen daadwerkelijk vaker een verklaring vormen voor een gebeurtenis waarin de kandidaat faalde, bijvoorbeeld omstandigheden als zware verantwoordelijkheden voor familieleden en een financieel lastige situatie. Daarbij komt dat allochtonen zich tijdens sollicitaties meestal tamelijk afwachtend en bescheiden opstellen, terwijl juist van kandidaten wordt verwacht dat ze hun eigen rol bij het beschrijven van een gebeurtenis benadrukken (Van Gent et al., 2006).

Eén vraag over de beoordelaar die op vele lippen is blijven liggen hebben we nog altijd niet besproken: Geven autochtone selecteurs autochtone kandidaten eigenlijk hogere beoordelingen dan allochtone kandidaten tijdens het selectiegesprek alleen al vanwege etnisch-demografische gelijkheid. Geldt dit ook vanwege gelijkheid met de kandidaat die ze waarnemen in de zin van zaken als houding, normen en waarden? Bij velen bestaat intuïtief het idee dat demografische en waargenomen gelijkheid beide een effect zullen hebben op beoordelingen van sollicitanten. Niet lang geleden echter hebben een aantal Amerikaanse onderzoekers door het gebruik van goede statistische analysemethoden aangetoond dat er in de praktijk bij deugdelijk gestructureerde

selectiegesprekken geen effecten te vinden waren van demografische gelijkheid (Sacco, Scheu, Ryan, & Schmitt, 2003). In ons eigen grootschalige onderzoek bij de Nederlandse politie kwamen we onlangs tot dezelfde conclusie: Ondanks onze vermoedens vooraf, bleken er in de praktijk van de politiselectie geen effecten aanwezig te zijn van demografische gelijkheid, en evenmin van waargenomen gelijkheid. Ook hier was sprake van goed gestructureerde selectiegesprekken en rollenspellen, en waren de assessoren en psychologen grondig getraind. In deze studie konden ook allochtone beoordelaars worden onderzocht, voor wie evenmin effecten van gelijkheid werden gevonden (De Meijer, Born, Van Loon, & Van der Molen, 2008).

Effecten van het gehele beoordelingsproces

Voordat we nauwkeuriger gaan kijken naar de manier waarop mensen zichzelf beoordelen en presenteren, is het van belang ons te realiseren dat een bespreking van het selectieproces meer omvat dan alleen een reflectie op de psychologische tests, het selectiegesprek, en eventuele andere selectiemiddelen. Kenmerken van de totale procedure en de context waarin de procedure zich afspeelt kunnen ook gerelateerd zijn aan de cultureel-etnische achtergrond van kandidaten. Ik geef er een paar. Eén bevinding, afkomstig uit de Verenigde Staten, is dat allochtone kandidaten zichzelf sneller dan autochtone kandidaten terugtrekken uit langshepende selectieprocedures die soms vele maanden in beslag kunnen nemen (Arvey, Gordon, Massengill, & Mussio, 1975). Als het bedrijfsmanagement voor een belangrijk deel zichtbaar afkomstig is uit de

etnische minderheid, blijken kandidaten uit deze minderheid aanmerkelijk positiever over de selectiemethoden dan autochtone kandidaten. Dit resultaat steekt af tegen de gebruikelijke bevinding dat kandidaten uit etnische minderheden juist negatiever zijn over selectiemethoden dan autochtone kandidaten (Becton, Feild, Giles, & Jones-Farmer, 2008).

Bij de meeste beoordelingsprocedures die in de praktijk plaatsvinden, integreert een menselijke beoordelaar de verzamelde informatie over kandidaten uit tests, vragenlijsten en een gesprek. Op grond van deze integratie komt de beoordelaar tot een voorspelling over het toekomstige werkgedrag en een advies over de geschiktheid van de kandidaat. Dit wordt ook wel klinische voorspelling genoemd (Meehl, 1954). In onderzoek bij de politie kwamen we er pas geleden achter dat bij de integratie van informatie tot een voorspelling en een advies, door autochtone selecteurs meer, waaronder dus ook meer irrelevante gegevens werden gebruikt. Bovendien bleken zij ook wat meer waarde te hechten aan de informatie die afkomstig was van anderen dan aan de informatie die zij zelf hadden verzameld (De Meijer, Born, Van Zielst, & Van der Molen, 2007). Onze verklaring hiervoor is dat autochtone selecteurs zich wellicht wat minder zeker voelen over het beoordelen van allochtone kandidaten. Ze zijn daarom geneigd zich te baseren op meer gegevens dan gewoonlijk, en ook meer op gegevens die door anderen zijn verzameld. Deze bevinding pleit in mijn ogen voor verdere expertiseontwikkeling van beoordelaars en een meer heterogene samenstelling van het beoordelaarsbestand, dat wil zeggen dat zowel autochtone als allochtone beoordelaars worden ingeschakeld.

Kortom, het is zaak ook aandacht te geven aan factoren uit de ruimere context van het beoordelingsproces. Als dat niet gebeurt, bedrijven we struisvogelpolitiek en blijven we blind voor factoren die mogelijk een verschillend effect hebben voor allochtone en autochtone kandidaten.

We namen hiervoor enkele kenmerken door van verschillende selectietests en de bredere context van de beoordelingsprocedure. Ook gingen we enkele mogelijke verklaringen na voor het soms partijdige en gekleurde gedrag van de selecteur, vooral als het om een ongeoefende gelegenheidsselecteur gaat die geen systematische voorschriften en procedures hanteert. Om ongewenste effecten van dergelijk ongeoefend gedrag tegen te gaan, is het belangrijk om personeelsbeoordelingen en –beslissingen zo transparant en gestructureerd mogelijk te laten plaatsvinden en er verantwoording over te verlangen. Het kan zinvol zijn een professioneel adviesbureau in te schakelen.

We verlaten het beoordelingsproces waarbij anderen als beoordelaar optreden, om nu aandacht te geven aan het proces van zelfbeoordeling.



4. De gekleurde zelfbeoordeling

Wat weten we eigenlijk van de zichzelf beoordelende sollicitant, die tegenwoordig een zeer heterogene origine heeft? Is er iets bekend over de wijze waarop deze zich, vaak bij een professioneel adviesbureau, presenteert via het invullen van persoonlijkheidsvragenlijsten?

Een oude kwestie die altijd al om persoonlijkheidsvragenlijsten heen hangt is de vraag of kandidaten dergelijke

vragenlijsten op een sociaal wenselijke manier invullen. Er hangt veel af van hun vragenlijstscores, want het gaat per slot meestal om het krijgen van een baan. In assessmentsituaties willen psychologen dan ook kunnen achterhalen of iemands zelfbeoordeling inderdaad is beïnvloed door sociaal wenselijk antwoordgedrag. Verschillen tussen kandidaten in de neiging tot zulk gedrag moeten worden geëlimineerd omdat ze een stoorzender kunnen vormen. Deze verschillen zijn niet van belang bij het meten van individuele persoonlijkheidskenmerken aan de hand van de vragenlijst.

Los van deze overwegingen is er in de tussentijd een levendige discussie op gang gekomen in de cross-culturele psychologie. De discussie betreft de vraag of het zich beter willen voordoen bij zelfevaluaties misschien een typisch westers verschijnsel is, of dat het juist om een universeel menselijk verschijnsel gaat. Ieder mens heeft toch de behoefte aan een positief zelfbeeld? Sommige studieresultaten wijzen erop dat het om een typisch westers fenomeen zou gaan (Heine, 2005) maar andere vormen juist bewijzen voor de universaliteit ervan (Sedikides, Gaertner, & Toguchi, 2003).

Deze onderzoekslijn kreeg een interessante wending door de vraag of het gepresenteerde zelfbeeld niet zou afhangen van het type persoonlijkheidskenmerk dat wenselijk is in een bepaalde cultuur (Sedikides, Gaertner, & Vevea, 2005; Sedikides, Gaertner, & Vevea, 2007). In de westerse cultuur zijn gewenste kenmerken vaak gerelateerd aan iemands individualisme, zoals originaliteit, onafhankelijkheid, uniekheid, zelfverzekerdheid en zelfexpressie. Samengevat

zou dit het 'ideale zelf' kunnen worden genoemd (cf. Higgins, 1987). In collectivistische culturen is een ander type attributen gewenst, waaronder conformisme, gerichtheid op het maken van compromissen, coöperatief gedrag, loyaliteit, respectvolheid, en zelfopoffering. Samengevat zou dit het 'normatieve zelf' genoemd kunnen worden, dat wil zeggen het onberispelijke zelf door de ogen van de sociale groep (cf. Higgins). In een Zuid-Afrikaanse studie vonden we bevestiging voor deze gedachtegang over culturele verschillen in wenselijke persoonlijkheidskenmerken onder etnisch Afrikaanse en etnisch Europese groepen (Born, Kevenaar, & Myburgh, 2008).

De conclusie die uit cross-culturele studies lijkt te komen is dat mogelijke verschillen tussen culturele groepen in sociaal wenselijke antwoordtendenties niet zozeer moeten worden uitgeschakeld als hinderlijk, maar dat deze tendenties kunnen worden opgevat als indicatoren van belangrijke waarden in een cultuur (Smith, 2004).

Op een fundamenteeler niveau speelt de kwestie of stabiele persoonlijkheidstrekken in niet-westerse, collectivistische culturen wel even belangrijk en even voorspellend zijn voor gedrag als in westerse, individualistische culturen, waar personen worden gezien als autonome entiteiten, elk met een kenmerkende set aan eigenschappen. Daarom komen er cross-culturele onderzoeksinitiatieven op gang waarin de relatie tussen persoonlijkheid en gedrag wordt bestudeerd en waarin ook pogingen worden gedaan om iemands persoonlijkheid af te leiden uit gedragsobservaties (onder andere met de

zogenaamde Act Frequency Approach; zie bijvoorbeeld Church, Katigbak, Miramontes, Del Prado, & Cabrera, 2007).

Een intrigerend bericht over persoonlijkheidsmeting aan de hand van gedragsobservatie komt verder bij het Nederlands Instituut voor Ecologie vandaan (Kraan, 2007). Aan dit instituut blijkt al sinds enige tijd een gedragstest voor de persoonlijkheid te worden gebruikt door de onderzoeksgroep van Pieter Drent. Deze onderzoekers volgen al negen jaar lang dezelfde groep koolmezen in het veld. Elke nieuwe vogel ondergaat eenmalig een psychologische test in het lab. Deze test houdt in dat de dieren in een voor hen onbekende omgeving worden gebracht om te kijken hoe ze daarmee omgaan. Koolmezen verschillen in de manier waarop ze op een nieuwe uitdaging of iets vreemds reageren. Sommige zijn dan roekeloos en andere voorzichtiger. Er blijken extraverte en introverte types te zijn. Na de psychologische test in het lab laten de onderzoekers de vogel weer vrij. De test hoeft niet herhaald te worden, want de resultaten blijken toch elke keer hetzelfde zijn. Deze zomer werd in het tijdschrift *Science* de bevinding gepubliceerd dat in Engeland de koolmezen goed blijken te gedijen door de stijgende temperaturen (Charmantier, McCleery, Cole, Perrins, Kruuk, & Sheldon, 2008). Nederlandse mezen blijken zich juist niet goed aan te passen aan de klimaatverandering. Er werd overigens geen verband gelegd met persoonlijkheidsverschillen tussen Engelse en Nederlandse mezen (Inman, 2008).

Net als in ecologisch onderzoek lijkt het ook voor psychologen belangrijk zich naast persoonlijkheidsonderzoek

met acculturatie en adaptatiemechanismen bezig te houden. Ook om scores op psychologische tests goed te kunnen interpreteren is het nodig na te gaan in hoeverre er sprake is van acculturatie bij een individuele kandidaat, zoals kennis van de taal en van de Nederlandse gewoonten en opvattingen (Poortinga, 2001). Voor dit doel bestaan acculturatievragenlijsten.

We keren van cross-culturele kwesties inmiddels weer terug naar vergelijkingen tussen etnisch-culturele groepen binnen een land. Eerder gaf ik al aan dat de verschillen tussen etnisch-culturele groepen op persoonlijkheidsvragenlijsten klein lijken te zijn (cf. De Meijer et al., 2006). Tegelijkertijd is er nog nauwelijks systematisch onderzoek gedaan naar mogelijke groepsverschillen op persoonlijkheidsvragenlijsten.

Pas voor het eerst in deze herfst is een omvangrijke overzichtsstudie verschenen waarin vier grote etnisch-culturele groepen in de Verenigde Staten op 44 persoonlijkheidsvragenlijsten zijn vergeleken. Hieruit blijkt dat de grootte van het groepsverschil op een persoonlijkheidsschaal sterk samenhangt met de keuze van de twee etnische groepen die met elkaar worden vergeleken en met de schaal die het betrof (Foldes, Duehr, & Ones, 2008). Als bijvoorbeeld zwarte Amerikanen werden vergeleken met Aziatische Amerikanen dan was er geen verschil in ordelijkheid, maar scoorde de zwarte Amerikaanse groep aanmerkelijk hoger op emotionele stabiliteit. In ons eigen onderzoek bij de Nederlandse politie bleken allochtone kandidaten hoger te scoren op consciëntieusheid dan autochtone kandidaten maar lager op andere

Big Five persoonlijkheidsfactoren (De Meijer, Born, Terlouw, & Van der Molen, 2006).

Dergelijk onderzoek geeft aan dat we momenteel nog geen goed inzicht hebben in voorkomende groepsverschillen in persoonlijkheid. Bovendien moeten we ook serieus rekening houden met de mogelijkheid dat systematische verschillen tussen groepen zijn gemaskeerd. Het promotieonderzoek van Anita de Vries dat aan deze leerstoel is verbonden, tracht de effecten van een aantal factoren op persoonlijkheidsscores te ontmaskeren. Het gaat daarbij onder andere om mogelijke effecten van de culturele groep waarmee iemand zich vergelijkt als deze zichzelf beoordeelt. Het is dan ook riskant om op dit moment te concluderen dat er nauwelijks of geen verschillen aanwezig zijn.

Voorwaardelijk voor een goede inhoudelijke interpretatie van groepsverschillen is overigens uiteraard dat de uitspraken in een persoonlijkheidsvragenlijst op de bedoelde manier worden begrepen door de verschillende etnisch-culturele groepen. Het kan daartoe heel zinvol zijn om personen te ondervragen, terwijl zij de uitspraken lezen en beantwoorden. Dit wordt cognitief pre-testen genoemd (NapolesSpringer, SantoyoOlsson, O'Brien, & Stewart, 2006).

Het dominante persoonlijkheidsmodel in de hedendaagse psychologie is het zogenaamde vijf factor model dat bestaat uit vijf grote persoonlijkheidsfactoren zoals Extraversie en Neuroticisme, en dat is ontstaan in onderzoek onder westerse populaties. Inmiddels is er een persoonlijkheidsmodel in opkomst dat is ontstaan in China en dat ook steeds meer belangstelling in de westerse wereld aan het krijgen is

(Cheung et al., 2001; Cheung, Cheung, Leung, Ward, & Leong, 2003; Lin & Church, 2004). Een kenmerk van dit model is dat er een belangrijke aparte factor in wordt onderscheiden die interpersoonlijke relaties en gerelateerdheid vertegenwoordigt. Deze factor representeert de gerichtheid op harmonieuze relaties en op het naleven van sociale normen en regels. Wellicht kan dit model, of andere in niet-westerse landen ontwikkelde persoonlijkheidsmodellen, diensten bewijzen in multiculturele westerse samenlevingen.

We zagen dat er vele ideeën en noties bestaan over de uitleg van verschillen tussen etnisch-culturele groepen op het gebied van persoonlijkheid (zie ook Van Leest, 2001). Hier moet systematischer onderzoek naar worden gedaan.



5. Slot

Het wordt tijd om voor vandaag de balans op te maken en ons op de toekomst te richten. Wat brengt die ons en welke vraagstukken zijn daarin het belangrijkste? Ik wil een paar sporen uitzetten. De meest prangende kwestie daarbij is de vraag of het beoordelingsproces bij selectie met de huidige middelen geschikt is voor een divers samengestelde kandidatenpool. We zagen dat vooral op cognitieve tests het scoreverschil tussen allochtone en autochtone groepen soms

sterk in het nadeel van de allochtone groepen kan uitvallen, terwijl deze tests over het algemeen juist één van de meest voorspellende zijn voor later werkgedrag.

Er zijn een paar manieren te noemen om een meer divers personeelsbestand te bereiken zonder dat er concessies hoeven te worden gedaan aan het selectierendement. Het gebruik van professioneel gestructureerde selectiegesprekken in combinatie met andere middelen zoals rollenspellen en situationele beoordelingstests is daar een van. Het gestructureerde selectiegesprek behoort namelijk naast algemene intelligentie tot de beste voorspellers van later werkgedrag. Een andere methode is ervoor zorgen dat tests die niet bestemd zijn om iemands Nederlandse taalkennis te meten, maar erop uit zijn om andersoortige competenties te meten, bijvoorbeeld integriteit, inderdaad niet onverhoopt toch iemands taalkennis meten. Zo'n strategie moet wel het minimale taalniveau in de gaten houden dat nodig is voor een bepaalde functie. Ook zagen we dat bij beoordelingsprocessen waarbij niet-professionals zijn betrokken, allerlei factoren een rol kunnen spelen die ten onrechte allochtone kandidaten benadelen. Beelden over een bepaalde etnisch-culturele groep die ongunstig zijn spelen soms nog door in de beoordeling van een individuele sollicitant uit deze groep.

Waar we veel minder van weten is of scores op persoonlijkheidsvragenlijsten een goed beeld leveren van persoonlijkheidskenmerken van allochtone kandidaten, en hoe gevonden verschillen tussen etnisch-culturele groepen moeten worden geïnterpreteerd.

De urgentie van een goede beoordeling van personen staat boven kijf. Ons tijdperk is meritocratischer geworden dan vroeger en stelt daardoor hogere eisen aan de manier waarop personen op de arbeidsmarkt op hun geschiktheid worden beoordeeld. Tegelijkertijd is het zwaarder geworden om goed aan deze eisen te voldoen. De arbeidsmarkt zelf is immers internationaler en diverser geworden door de gevarieerde groepen die er zijn neergestreken. Dwars door deze veranderingen heen spelen de grote technologische ontwikkelingen in de maatschappij, die ook een effect hebben op de inhoud van het werk en op het psychologisch meetinstrumentarium. Een dringend vraagstuk daarbij is of een technologisch vernieuwd meetinstrumentarium, dat tegenwoordig bestaat uit tests die via het internet worden afgenomen, maar ook uit games en webcamtests, niet onbedoeld een groep pechvogels op de arbeidsmarkt treft die minder ervaring hebben in computer- en internettoepassingen.

Zijn deze ontwikkelingen reden om onze beoordelingsmethodieken te hervormen? Als we daarmee doelen op het verder verbeteren van deze methodieken dan is het antwoord een eenvoudig 'ja'. Ik wil daartoe een aantal stellingen te berde brengen. Luther heeft reeds in de zestiende eeuw laten zien dat het afkopen van verantwoordelijkheden via een aflat geen oplossing kan bieden, maar dat er inspanningen moeten worden geleverd:

1. Veel ontwikkelde tests en vragenlijsten rapporteren niet of nauwelijks over de etnisch-culturele samenstelling van hun normen. De Commissie Testaangelegenheden van het

Nederlands Instituut van Psychologen (COTAN) dient in het beoordelingssysteem representativiteit in etnische samenstelling van de normgroepen op te nemen als aparte beoordelingscategorie ter bepaling van de kwaliteit van normen. In dit beoordelingssysteem dient bovendien als aparte kwaliteitseis onpartijdigheid jegens etnisch-culturele groepen te worden opgenomen. Hier horen zaken bij als een beoordeling op kwetsende inhoud, en onderzoek naar de vergelijkbaarheid van de onderliggende teststructuur als er vermoedens bestaan dat de interpretatie van bepaalde begrippen en vragen verschilt tussen etnisch-culturele groepen en ook binnen deze groepen, afhankelijk van de mate van acculturatie aan de Nederlandse cultuur.

2. Bij beoordelingsprocedures in arbeidsorganisaties heerst dikwijls een gebrek aan voorschriften en procedures en zijn de betrokkenen vaak ongeoefende selecteurs. Om het effect van onjuiste beelden over etnische minderheden te verminderen die bij selecteurs en beoordelaars in arbeidsorganisaties bestaan, moet explicieter organisatiebeleid worden geformuleerd. Dit beleid dient te verlangen om personeelsbeslissingen systematischer en transparanter te laten plaatsvinden en te verantwoorden (Messick & Bazerman, 1996). Maak ook zoveel mogelijk gebruik van professionele adviesbureaus. Onderzoek lijkt namelijk uit te wijzen dat discriminatoir gedrag daar minder kans heeft.

3. Er moet systematischer onderzoek gedaan worden naar verschillen en overeenkomsten tussen etnisch-culturele groepen op het gebied van persoonlijkheidsvragenlijsten. Dit gebied is nog vrij onontgonnen.

Levendig herinner ik me hoe in maart 2006 na de oratie van Wilma Resing een groepje Amsterdamse collega's bij mij neerstreek, hoe het groepje na wat geklets weer was verdwenen, maar hoe Nico Bleichrodt naast me was blijven zitten en na een tijdje vroeg of ik zijn leerstoel zou willen bezetten nu hij met emeritaat zou gaan. Ik ben de Stichting Nederlands Onderzoekscentrum Arbeidsmarkt en Allochtonen (NOA) in de persoon van directeur Remko van den Berg, en de afdeling Arbeids- en Organisationspsychologie van de Faculteit Psychologie en Pedagogiek in de persoon van Henk van der Flier bijzonder erkentelijk voor het vertrouwen dat ze mij hebben geschonken door het mogelijk maken en aanbieden van deze Leerstoel, die vanwege de Stichting Het Vrije Universiteitsfonds in het leven is geroepen. Via deze leerstoel kan psychologisch onderzoek een belangrijke betekenis hebben voor de huidige maatschappij. De verbinding met Stichting NOA is een prettige omdat NOA gericht is op het ontwikkelen van tests en vragenlijsten die multicultureel toepasbaar zijn.

Hoe cruciaal een zorgvuldige beoordeling is in deze multiculturele maatschappij, valt af te leiden uit ons verre verleden, namelijk uit de waarschijnlijk oudste ons bekende Nederlandstalige zin uit de elfde eeuw. We kennen deze zin wellicht allemaal: *Hebban olla uogala nestas hagunnan hinase hi(c) (a)nda thu uuat unbidan uue nu* (Hebben alle vogels nesten begonnen, behalve ik en jij. Waarop wachten we nu?). Eén van de interpretaties van deze zin is dat deze het verlangen van de schrijver uitdrukt om te worden opgenomen in de gemeenschap. Eén van de cruciale manieren voor allochtoon-etnische

groepen om te worden opgenomen in de Nederlandse gemeenschap is het verkrijgen van een goede opleiding en een goede baan. Ik hoop met het onderzoek dat aan deze leerstoel is gerelateerd hieraan een bijdrage te kunnen leveren.

Een belangrijke taak binnen mijn leeropdracht is het begeleiden van promotieonderzoek. Daarom is Anita de Vries afgelopen januari doortastend gestart met een onderzoek gericht op persoonlijkheidsmeting in cultureel diverse groepen. Dat ik in dit project heel plezierig samenwerk met twee De Vriezen – de co-promotor bij dit project heet Reinout de Vries – heeft vast te maken met het feit dat mijn wortels in Friesland liggen.

Ik ben erg trots dat ik het werk van Nico Bleichrodt per januari 2007 mocht gaan voortzetten. Samen met Henk van der Flier begeleidde Nico mij op mijn allereerste schreden in de wetenschap. Bij mijn eerste internationale publicatie in 1987 waren zij mijn co-auteurs. Die publicatie in de *Journal of Cross-Cultural Psychology* (Born, Bleichrodt, & Van der Flier, 1987) is achteraf gezien een verre voorbode geweest van deze leerstoel, die ik twintig jaar later zou gaan bezetten.

Graag wil ik met name Pieter Drenth en Jen Algera noemen, die mij als promotoren coachten naar het doctoraat in 1995. Na zeven jaren afwezigheid ben ik sinds vorig jaar weer verbonden aan mijn Alma Mater, de Vrije Universiteit. Het vertrouwde VU-gevoel werd in dat jaar meteen versterkt door Rector Magnificus Lex Bouter toen hij me uitnodigde op de kennismakingsthee. De leerstoel aan de VU vormt een harmonieus geheel met de bijzondere leerstoel die ik sinds

2006 bezet aan de Erasmus Universiteit Rotterdam en die zich bevindt op het gebied van de Personeelspsychologie, in het bijzonder de selectiepsychologie. Er is sprake van een goede kruisbestuiving. Dit gaat ver terug. Al in de 16de eeuw was er eensgezindheid over een aantal belangrijke thema's tussen Luther, Calvijn en Erasmus, hoewel Erasmus liever de fluwelen hervormingsweg beliep dan de revolutionaire.

Voor de voorbereidingen van vandaag ben ik Loes Palplatz en Anna Brinkman zeer erkentelijk. Loes kwam precies op tijd achter het mysterie van de dubbele boeking van deze aula voor vandaag. Anna kreeg het voor elkaar om in één dag tijd alle uitnodigingen klaar te hebben. Ook Maud Beuker Andreae bedank ik graag voor haar inspanning om een aantal CBS-statistieken boven water te krijgen.

U hoefde vandaag niet alleen naar mij te luisteren maar kon ook enkele andere vogels observeren, die mijn Rotterdamse collega Eveline Osseweijer in haar vrije tijd tot een kleurige presentatie heeft samengesteld. Graag wil ik Henk van der Molen hartelijk bedanken voor het geven van feedback op een eerdere versie van de tekst. Mijn ouders Pieter en Else Born dank ik voor hun enthousiasme en het gemak waarmee ze een delicaat Pakistaans vogeltje uit hun 60-er jaren archief wisten op te vissen. Verrast werd ik door mijn dochter Lily's prachtige veelkleurige geschilderde inbreng die ze me in augustus onverwachts aanleverde. "Alsjeblieft mam, gebruik maar in je oratie". Dank je wel, liefje! Chris, je relaxte houding tegenover het alweer voorbereiden van een oratie werkte verfrissend. Naast jou is het heel prettig werken.

Geachte aanwezigen, graag dank ik u tot slot voor de belangstelling die u vandaag heeft getoond. En pikt u zo dadelijk een lekker toastje mee! Ik heb gezegd.

Referenties

- Andriessen, I., Dagevos, J., Nievers, E., & Boog, I. (2007). *Discriminatiemonitor niet-westerse allochtonen op de arbeidsmarkt 2007*. Den Haag/Rotterdam: Sociaal en Cultureel Planbureau.
- Arvey, R. D., Gordon, M. E., Massengill, D. P., & Mussio, S. J. (1975). Differential dropout rates of minority and majority job candidates due to "time lags" between selection procedures. *Personnel Psychology*, 28(2), 175-180.
- Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues*, 46(1), 27-46.
- Bardi, A., Calogero, R. M., & Mullen, B. (2008). A new archival approach to the study of values and value-behavior relations: Validation of the value lexicon. *Journal of Applied Psychology*, 93(3), 483-497.
- Becton, J. B., Feild, H. S., Giles, W. F., & Jones-Farmer, A. (2008). Racial differences in promotion candidate performance and reactions to selection procedures: A field study in a diverse top-management context. *Journal of Organizational Behavior*, 29(3), 265-285.
- Bochhah, N., Kort, W., & Seddik, H. (2005a). *Richtlijnen gebruik diagnostische instrumenten bij etnische minderheden*. Rotterdam/ Amsterdam: LBR/ NIP.
- Bochhah, N., Kort, W., & Seddik, H. (2005b). *Toepasbaarheid van enkele psychologische tests bij personeelsbeoordeling bij etnische minderheden*. Rotterdam/ Amsterdam: LBR/ NIP.
- Bochhah, N., Kort, W., Seddik, H., & Van de Vijver, F. (2001). *Deskundigen over het testen van etnische minderheden*. Rotterdam/ Amsterdam: LBR/ NIP.
- Born, M.Ph. (2007) *Vruchten van selectie*. Oratie. Rotterdam: Erasmus Universiteit Rotterdam.
- Born, M. Ph., Bleichrodt, N., & Van der Flier, H. (1987). Cross-cultural comparison of sex-related differences on intelligence tests: A meta-analysis. *Journal of Cross-Cultural Psychology*, 18(3), 283-314.
- Born, M. Ph., Deros, E., & Hiemstra, A. (2008). *Ethnicity and graduate resume content*. Manuscript in preparation.
- Born, M. Ph., Kevenaar, M., & Myburgh, W. (2008, July). Self concepts and response styles in personality testing across cultures. In Born, M.Ph. & Cheung, F. (Chairs). *Testing and assessment in emerging and developing countries I: The current state*. Symposium conducted at the 39th International Congress of Psychology, Berlin.
- Brewer, M. B. (2007). The importance of being "we": Human nature and intergroup relations. *American Psychologist*, 62(8), 728-738.
- Calgar, A. (2004). Recruitment in the Turkish police. *Policing & Society*, 14(4), 348-364.
- CBS Statline (2008). <http://www.cbs.nl/nl-NL/menu/themas/dossiers/allochtonen/nieuws/default.htm>. Verkregen op 2 oktober 2008.
- Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B., & Sheldon, B. C.. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, 320 (5877), 800-803.
- Cheung, F. M., Cheung, S. F., Leung, K., Ward, C., & Leong, F. (2003). The English version of the Chinese personality assessment inventory. *Journal of Cross-Cultural Psychology*, 34(4), 433-452.
- Cheung, F. M., Leung, K., Zhang, J., Sun, H., Gan, Y., Song, W., et al. (2001). Indigenous Chinese personality constructs: Is the five-factor model complete? *Journal of Cross-Cultural Psychology*, 32(4), 407-433.
- Church, A. T., Katigbak, M. S., Miramontes, L. G., Del Prado, A. M., & Cabrera, H. F. (2007). Culture and the behavioural manifestations of traits: An application of the act frequency approach. *European Journal of Personality*, 21(4), 389-417.
- Cole, M. S., Rubin, R. S., Feild, H. S., & Giles, W. F. (2007). Recruiters' perceptions and use of applicant resume information: Screening the recent graduate. *Applied Psychology-an International Review-Psychologie Appliquee-Revue Internationale*, 56(2), 319-343.
- De Meijer, L. A. L., Born, M. P., Terlouw, G., & Van der Molen, H. T. (2006). Applicant and method factors related to ethnic score differences in personnel selection: A study at the Dutch police. *Human Performance*, 19(3), 219-251.

- De Meijer, L.A.L., Born, M.Ph., Van Loon, H., & Van der Molen, H. T. (2008). *Through the eyes of the assessor: Demographic and perceived similarity with regard to score differences between ethnically diverse applicants*. Manuscript under review.
- De Meijer, L. A. L., Born, M. P., Van Zielst, J., & Van der Molen, H. T. (2007). Analyzing judgments of ethnically diverse applicants during personnel selection: A study at the Dutch police. *International Journal of Selection and Assessment*, 15(2), 139-152.
- Dean, M. A., Bobko, P., & Roth, P. L. (2008). Ethnic and gender subgroup differences in assessment center ratings: A meta-analysis. *Journal of Applied Psychology*, 93(3), 685-691.
- Derous, E. (2007). Naamsdiscriminatie bij cv-screening. *Tijdschrift voor Arbeidsvraagstukken*, 23 (4), 365-379.
- Detert, J. R., Trevino, L. K., & Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: A study of antecedents and outcomes. *Journal of Applied Psychology*, 93(2), 374-391.
- Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J.L. Eberhardt & S.T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3-32). Thousand Oaks, CA: Sage.
- Evers, A. V. A. M., Te Nijenhuis, J., & Van der Flier, H. (2005). Ethnic bias and fairness in personnel selection: Evidence and consequences. In A. V. A. M. Evers, O. Smit-Voskuyl, and N. Anderson (Eds.). *The Blackwell handbook of personnel selection*. (pp. 306-328). Oxford, UK: Blackwell.
- Evers, A., Van Vliet-Mulder, J. C., & Groot, C. J. (2000). *Documentatie van tests en testresearch in Nederland* (Deel I). Assen: Van Gorcum.
- Fischer, R., & Smith, P. B. (2004). Values and organizational justice - performance- and seniority-based allocation criteria in the United Kingdom and Germany. *Journal of Cross-Cultural Psychology*, 35(6), 669-688.
- Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five US racial groups. *Personnel Psychology*, 61(3), 579-616.
- Geva, A. (2006). A typology of moral problems in business: A framework for ethical management. *Journal of Business Ethics*, 69(2), 133-147.
- Heine, S. J. (2005). Where is the evidence for pancultural self-enhancement? A reply to Sedikides, Gaertner, and Toguchi (2003). *Journal of Personality and Social Psychology*, 89(4), 531-538.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3), 319-340.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
- Hofstede, G. (2001). *Cultures consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.
- Hofstee, W. K. B., Campbell, W. H., Eppink, A., Evers, A., Joe, R. C., Van de Koppel, J. M. H., Zweers, H., Choenni, C. E. S., & Van der Zwan, T. J. (1990). *Toepasbaarheid van psychologische tests bij allochtonen*. Utrecht: Landelijk Bureau Racismebestrijding. LBR-serie nr.11
- Howard, A. (1986). College experiences and managerial performance. *Journal of Applied Psychology*, 71(3), 530-552.
- Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology*, 83(2), 179-189.
- Human Rights Watch (2008). *The Netherlands: Discrimination in the name of integration; Migrants' rights under the Integration Abroad Act*. New York: Auteur.
- Huo, Y. P., Huang, H. J., & Napier, N. K. (2002). Divergence or convergence: A cross-national comparison of personnel selection practices. *Human Resource Management*, 41(1), 31-44.
- Inman, M. (2008). *Great tits enjoying the warmer weather – so far*. NewScientist.com news service. <http://environment.newscientist.com/article/dn13857>. Verkregen op 2 oktober 2008.
- Jennissen, R. P. W., & Oudhof, J. (Red.) (2007). *Ontwikkelingen in de maatschappelijke participatie van allochtonen. Een theoretische verdieping en een thematische verbreding van de Integratiekaart 2006*. Den Haag: WODC.

- Kessens, J. M., & Jacobusse, G. W. (2007). *Onderzoek naar de kwaliteit van het inburgeringsexamen buitenland*. Soesterberg: TNO Defensie en Veiligheid.
- Kohlberg, L. (1969). Stage and sequence: The cognitive development approach to socialization. In D. A. Goslin (Ed.). *Handbook of socialization theory* (pp. 347-480). Chicago, Ill: Rand McNally.
- Kok, B. (2007). *Het lab*. Amsterdam: Olive Press.
- Kraan, L. (2007). *Psychologische test voor vogels*. <http://noorderlicht.vpro.nl/artikelen/35812651/>. Verkregen op 1 oktober 2008.
- Kravitz, D. A. (2008). The diversity-validity dilemma: Beyond selection-the role of affirmative action. *Personnel Psychology*, 61(1), 173-193.
- Kruisbergen, E. W., & Veld, Th. (2002). *Een gekleurd beeld. Over beelden, beoordeling en selectie van jonge allochtone werknemers*. Assen: Van Gorcum.
- Lin, E. J., & Church, A. T. (2004). Are indigenous Chinese personality dimensions culture-specific? An investigation of the Chinese personality assessment inventory in Chinese American and European American samples. *Journal of Cross-Cultural Psychology*, 35(5), 586-605.
- Lovinsky, G. E., Trevino, L. K., & Jacobs, R. R. (2007). Assessing managers' ethical decision-making: An objective measure of managerial moral judgment. *Journal of Business Ethics*, 73(3), 263-285.
- Maassen, M., & Oosterwijk, F. (2006). *Taboes. 100 gevoelens waar Nederlanders zich voor schamen*. Amsterdam: Balans.
- McFall, L. (1987). Integrity. *Ethics: An International Journal of Social, Political, and Legal Philosophy*, 87.98 5-20.
- McGonigle, T. P., O'Shea, P. G., Curnow, C. K., & Brown, K. G. (2004, April). *An empirical comparison of I/O psychology's internal and external identities*. Paper presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology, April 2004, Chicago, Illinois.
- Meehl, P. E. (1954). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. (1954, Reprinted 1996). *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. xviii, 149 Pp. Lanham, MD: Jason Aronson.
- Meeuwisse, M., Born, M.Ph., & Severiens, S. (2008). *Familie-(bij)baan-vrije tijd-school interferentie en facilitatie voor studenten van diverse etnische achtergronden. Enkele resultaten op basis van een kleine pilot*. Tussenrapportage College van Bestuur. Intern rapport. Rotterdam: Erasmus Universiteit.
- Meeuwisse, M., Severiens, S., & Born, M.Ph. (2008). *Reasons for withdrawal from higher vocational education. A comparison of ethnic minority and majority non-completers*. Manuscript submitted for publication.
- Messick, D. M., & Bazerman, M. H. (1996). Ethical leadership and the psychology of decision making. *Sloan Management Review*, 37(2), 9-22.
- Miller, D. T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. *Journal of Personality & Social Psychology*, 57(4), 581-589.
- Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, 67(6), 949-971.
- Muñiz, J., & Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, 12, 206-219.
- Nail, P. R., Harton, H. C., & Barnes, A. (2008). A test of Dovidio and Gaertner's integrated model of racism. *North American Journal of Psychology*, 10(1), 197-220.
- NapolesSpringer, A. M., SantoyoOlsson, J., O'Brien, H., & Stewart, A. L. (2006). Using cognitive interviews to develop surveys in diverse populations. *Medical Care*, 44(11, Suppl 3), S21-S30.
- Nederlands Instituut van Psychologen. (2007). *Beroepscode voor psychologen 2007*. Amsterdam: Auteur.
- Nederlands Instituut van Psychologen. (2004). *Algemene Standaard Testgebruik*. Amsterdam: Auteur.
- Nederlandse Vereniging voor Personeelsmanagement en Organisatieontwikkeling. (2006). *NVP Sollicitatiecode* (Brochure). Nieuwegein: Auteur.
- Nemanick, R. C. J., & Clark, E. M. (2002). The differential effects of extracurricular activities on attributions in resume evaluation.

- International Journal of Selection and Assessment*, 10(3), 206-217.
- Opotow, S. (1990a). Deterring moral exclusion. *Journal of Social Issues*, 46(1), 173-182.
- Opotow, S. (1990b). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, 46(1), 1-20.
- Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, 108(5), 937-975.
- Pager, D., & Quillian, L. (2005). Walking the talk? what employers say versus what they do. *American Sociological Review*, 70(3), 355-380.
- Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology*, 61, 153-172.
- Poortinga, Y. H. (2001). Het gebruik van psychologische tests bij allochtonen: Probleembeschrijving. In N. Bleichrodt & F. J. R. van de Vijver (Red.), *Diagnostiek bij allochtonen. Mogelijkheden en beperkingen van psychologische tests* (pp. 15-42) Lisse: Swets & Zeitlinger.
- Rest, J. R. (1975). Longitudinal study of the defining issues test of moral judgment: A strategy for analyzing developmental change. *Developmental Psychology*, 11(6), 738-748.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Roth, P., Bobko, P., McFarland, L., & Buster, M. (2008). Work sample tests in personnel selection: A meta-analysis of black-white differences in overall and exercise scores. *Personnel Psychology*, 61(3), 637-661.
- Sacco, J. M., Scheu, C. R., Ryan, A. M., & Schmitt, N. (2003). An investigation of race and sex similarity effects in interviews: A multilevel approach to relational demography. *Journal of Applied Psychology*, 88(5), 852-865.
- Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, 84(1), 60-79.
- Sedikides, C., Gaertner, L., & Vevea, J. L. (2005). Pancultural self-enhancement reloaded: A meta-analytic reply to Heine (2005). *Journal of Personality and Social Psychology*, 89(4), 539-551.
- Sedikides, C., Gaertner, L., & Vevea, J. L. (2007). Inclusion of theory-relevant moderators yield the same conclusions as Sedikides, Gaertner, and Vevea (2005): A meta-analytical reply to Heine, Kitayama, and Hamamura (2007). *Asian Journal of Social Psychology*, 10(2), 59-67.
- Silvester, J. (1997). Spoken attributions and candidate success in graduate recruitment interviews. *Journal of Occupational and Organizational Psychology*, 70(1), 61-73.
- Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, 35(1), 50-61.
- Stangor, C. (2000). Stereotypes and prejudice: Essential readings. (2000). *Stereotypes and Prejudice: Essential Readings*. xii, 490 Pp. New York, NY, US: Psychology Press.
- Tijssen, M., Berntsen, A., De Jong, J.H.A.L., & Lennig, M. (2005). *Verantwoording Inburgeringsexamen Kennis van de Nederlandse Samenleving*. Den Bosch: CINOP.
- Trevino, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management*, 32(6), 951-990.
- Triandis, H. C. (2004). The many dimensions of culture - academic commentary. *Academy of Management Executive*, 18(1), 88-93.
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134(3), 383-403.
- Van den Berg, R. (2001). *Psychodiagnostiek bij kandidaten met verschillende culturele achtergronden*. Academisch proefschrift. Amsterdam: Vrije Universiteit.
- Van den Berg, H., & Evers, J. (2006). Discriminatie-ervaringen 2005. Een onderzoek naar ervaringen met discriminatie op grond van land van herkomst, geloof en (huids)kleur. In I. Boog, J. van Donselaar, D. Houtzager, P.R.Rodrigues en R. Schriemer (Red.), *Monitor rassendiscriminatie 2005* (pp.15-47). Rotterdam: Landelijk Bureau ter bestrijding van Rassendiscriminatie.
- Van Gent, M.J., Hello, E., Odé, A.W.M., Tromp, E., & Stouten, J. (2006). *Hogeropgeleide allochtonen op weg naar werk: Successen en belemmeringen*. Den Haag: Raad voor Werk en Inkomen.
- Van Leest, P.F. (2001). *Persoonlijkheidsmeting bij allochtone volwassenen*.

In N. Bleichrodt & F. van de Vijver (Red.). *Diagnostiek bij allochtonen. Mogelijkheden en beperkingen van psychologische tests* (pp. 143-156).

Lisse: Swets & Zeitlinger.

Weber, J. (1990). Managers moral reasoning - assessing their responses to 3 moral dilemmas. *Human Relations*, 43(7), 687-702.

Wilson, M. S. (2004). Values and political ideology: Rokeach's two-value model in a proportional representation environment. *New Zealand Journal of Psychology*, 33(3), 155-162.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, 10(2), 109-120.

