

# Persoonlijkheid als voorspeller van werkprestatie en contraproductief werkgedrag

## Het belang van specifieke persoonlijkheidsmetingen\*

Anita de Vries, Reinout E. de Vries, Marise Ph. Born & Remko H. van den Berg\*\*

*Dit artikel heeft ten eerste tot doel inzicht te geven in de vraag welke persoonlijkheidsdimensies werkprestatie en contraproductief werkgedrag voorspellen. Een tweede doel is om vast te stellen hoe de voorspellende waarde van persoonlijkheid verhoogd kan worden. Hierbij is gekeken naar drie factoren: het gebruik van specifieke persoonlijkheidsschalen, het toevoegen van een specifieke context aan de persoonlijkheidsvragen en het beïnvloeden van de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest. Uit de resultaten komt naar voren dat de voorspellingskracht van een persoonlijkheidstest verbetert wanneer gekeken wordt naar de specifieke onderliggende schalen van de brede dimensies en wanneer een relevante context aan de persoonlijkheidsitems, in dit geval 'op het werk', wordt toegevoegd. Daarnaast suggereren de bevindingen dat de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest, invloed heeft op het soort antwoorden dat diegene geeft. Dit onderzoek benadrukt hiermee het belang van meer specifieke persoonlijkheidsmetingen om de voorspelling van werkprestatie en contraproductief werkgedrag te verbeteren.*

## 1 Inleiding

*Bent u bij afspraken precies op tijd? Blijft u bij problemen rustig? Vindt u het leuk om veel aandacht van een groep mensen te krijgen? Geeft u het eerlijk toe als u een fout heeft gemaakt? Vindt u het leuk om iets origineels te bedenken? Vragen andere mensen u vaak om hulp?*

Uw antwoorden op deze vragen laten iets zien over uw persoonlijkheid: de kenmerken en eigenschappen die u maken zoals u bent. Meerdere studies hebben aangetoond dat iemands persoonlijkheid een relatie vertoont met zijn of haar prestaties op school en op het werk (bijv. Barrick, Mount & Judge, 2001; Poropat,

\* Dit artikel is gebaseerd op De Vries, A. (2012). *Specificity in personality measurement*. Proefschrift: VU Universiteit Amsterdam.

\*\* Anita de Vries was werkzaam bij psychologisch adviesbureau NOA. Correspondentieadres: NOA, Jollemanhof 14 A, 1019 GW Amsterdam, tel. 020-5040800, info@noa-vu.nl. Reinout E. de Vries is verbonden aan de Vrije Universiteit Amsterdam, afdeling Sociale en Organisationspsychologie. Marise Ph. Born bekleedt de door NOA gefinancierde leerstoel 'Arbeids- en Personeelspsychologie in het bijzonder gericht op allochtoon-etnische groepen' aan de Vrije Universiteit Amsterdam, afdeling Sociale en Organisationspsychologie. Remko H. van den Berg is werkzaam bij psychologisch adviesbureau NOA.

2009). Als gevolg hiervan worden in selectiesituaties vaak persoonlijkheidstests gebruikt met als doel het voorspellen van toekomstige prestaties.

Hoewel intelligentie regelmatig als een belangrijke voorspeller van studie- en werkprestatie wordt gezien, is de laatste tijd uit onderzoek gebleken dat, in het bijzonder in een schoolsetting waar de verschillen in intelligentie relatief klein zijn, persoonlijkheid toegevoegde waarde heeft als het gaat om het voorspellen van studieprestatie (bijv. Conard, 2006). Een verklaring voor deze voorspellende waarde van persoonlijkheid is dat intelligentie vooral bepaalt wat iemand maximaal aankan qua niveau, terwijl persoonlijkheid meer inzicht geeft in de prestaties die iemand daadwerkelijk zal laten zien (Furnham & Chamorro-Premuzic, 2004). Persoonlijkheidstests worden dan ook regelmatig gebruikt in selectie- en intakeassessments.

De meeste persoonlijkheidstests zijn gebaseerd op het bekende Big Five-persoonlijkheidsmodel. Dit model is gebaseerd op lexicaal onderzoek naar het gebruik van persoonsbeschrijvende termen in verschillende talen en bestaat uit de volgende vijf brede persoonlijkheidsdimensies: Consciëntieusheid, Extraversie, Vriendelijkheid, Emotionele Stabiliteit en Openheid voor Ervaringen (Goldberg, 1990). Heranalyses van het lexicale onderzoek dat eerder heeft geleid tot het Big Five-model, laten echter zien dat niet vijf, maar zes persoonlijkheidsdimensies onderscheiden kunnen worden. Het relatief nieuwe persoonlijkheidsmodel, dat bekendstaat als het HEXACO-model, bestaat uit varianten van de Big Five-persoonlijkheidsdimensies en uit een zesde dimensie: Integriteit (Ashton & Lee, 2007). Binnen het HEXACO-model heeft Integriteit betrekking op de mate waarin personen hebzucht vermijden, en oprecht, rechtvaardig en bescheiden zijn.

Hoewel de belangrijkste toevoeging van het HEXACO-model ten opzichte van het Big Five-model de toevoeging van de Integriteitsdimensie is, suggereren sommige onderzoekers dat Consciëntieusheid en Integriteit overlappende constructen zijn (bijv. Schmidt & Hunter, 1998). Ons inziens is deze gedachtegang ten eerste gebaseerd op een verkeerde interpretatie van eerdere bevindingen, waarbij een onderscheid gemaakt moet worden tussen open integriteitstests (*overt integrity tests*) en integriteitstests gebaseerd op persoonlijkheid (*covert integrity tests*). Een open integriteitstest bestaat uit concrete vragen over het aantal integriteitsschendingen, terwijl een covert integriteitstest bepaalde persoonlijkheidskenmerken meet, gerelateerd aan bijvoorbeeld betrouwbaarheid, die een (positieve of negatieve) relatie vertonen met contraproductief gedrag. Consciëntieusheid blijkt voornamelijk met covert integriteitstests te correleren (Schmidt & Hunter, 1998), terwijl Integriteit als onderdeel van het HEXACO-model sterke relaties vertoont met open integriteitstests (Lee, Ashton & De Vries, 2005). Ten tweede hebben studies aangetoond dat Consciëntieusheid en Integriteit binnen het HEXACO-model twee onafhankelijke persoonlijkheidsconstructen zijn ( $r < .16$ ), terwijl Consciëntieusheid in het HEXACO-model bijna niet te onderscheiden is van Consciëntieusheid in het Big Five-model ( $r = .82$ , zie bijvoorbeeld De Vries, De Vries, De Hoogh & Feij, 2009). Bovengenoemde bevindingen laten zien dat er niet veel bewijs is voor een sterke conceptuele overlap tussen Consciëntieusheid en Integriteit binnen het HEXACO-model. Eerder onderzoek heeft verder aangetoond dat, binnen het HEXACO-model, de persoonlijkheidsdimensie Integriteit – ten

opzichte van de andere persoonlijkheidsdimensies – in het bijzonder een significante bijdrage levert aan de voorspelling van onethisch werkgedrag, zoals diefstal, vandalisme en alcoholgebruik op de werkvloer (Lee, Ashton & De Vries, 2005).

In Nederland zijn op basis van het HEXACO-model twee vragenlijsten ontwikkeld die de zes persoonlijkheidsdimensies meten en die centraal staan in het huidige onderzoek: de gereviseerde HEXACO Persoonlijkheidsvragenlijst (HEXACO-PI-R; De Vries, Ashton & Lee, 2009) en de Multiculturele Persoonlijkheidstest – Big Six (MPT-BS; De Vries, De Vries & Born, 2011; NOA, 2009).

Vanwege het veelvuldige gebruik van persoonlijkheidstests in selectie- en begeleidingssituaties is het belangrijk om onderzoek te (blijven) doen naar de voorspellende waarde van persoonlijkheid én naar factoren die de voorspellende waarde van persoonlijkheid kunnen beïnvloeden. Met dit doel voor ogen richt het huidige onderzoek, dat gerapporteerd is in het proefschrift van De Vries (2012), zich op de zogeheten specificiteit van persoonlijkheidsmetingen. Ten eerste is onderzocht of persoonlijkheid studie- en werkprestatie beter kan voorspellen als gekeken wordt naar de specifieke onderliggende schalen van de brede dimensies (*bandwidth-fidelity dilemma*; Cronbach & Gleser, 1965). Hierbij is ook aandacht besteed aan etniciteit; eerder onderzoek heeft bijvoorbeeld aangetoond dat verschillen tussen culturele groepen op persoonlijkheidsvragenlijsten klein lijken te zijn (bijv. Ployhart & Holtz, 2008), maar er is nog nauwelijks onderzoek gedaan naar mogelijke groepsverschillen op de specifieke onderliggende schalen. Ten tweede is onderzocht of door het toevoegen van een specifieke context aan de persoonlijkheidsitems de voorspellingskracht van persoonlijkheid verbetert (*frame-of-reference effect*; Lievens, De Corte & Schollaert, 2008). Ten derde is nagegaan of de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest, effect heeft op het soort antwoorden dat iemand geeft (*reference-group effect*; Heine, Lehman, Peng & Greenholtz, 2002).

Als blijkt dat het gebruik van specifieke onderliggende persoonlijkheidsschalen en het toevoegen van een context inderdaad de voorspellingskracht van persoonlijkheid verbeteren, dan kan deze kennis gebruikt worden bij de inzet én de constructie van persoonlijkheidsinstrumenten in selectiesituaties. Wanneer er daadwerkelijk groepsverschillen zijn op specifieke persoonlijkheidsschalen en wanneer het aanwijzen van een specifieke vergelijkingsgroep inderdaad invloed heeft op de vraag welke persoonlijkheidskenmerken gerapporteerd worden, dan is het aan te raden hier rekening mee te houden bij de inzet van persoonlijkheidstests in de praktijk.

In de volgende paragrafen zal eerst een overzicht worden gegeven van de bestaande literatuur met betrekking tot persoonlijkheid als voorspeller van studie- en werkprestatie, het bandwidth-fidelity dilemma, het frame-of-reference effect en het reference-group effect. Per onderwerp zullen vervolgens hypothesen geformuleerd worden die aan de hand van meerdere studies getoetst zijn.

## 2 Persoonlijkheid als voorspeller van studie- en werkprestaties

Een belangrijke vraag binnen het onderwijs en binnen organisaties is welke factoren een relatie vertonen met studie- en werkprestaties. Onlangs heeft onderzoek in Nederland aangetoond dat naast de persoonlijkheidsdimensie Consciëntieusheid ook Integriteit een positieve voorspeller is van studieprestatie (De Vries, De Vries et al., 2011; De Vries, Van den Berg, Born & De Vries, 2011). Op het niveau van de smalle onderliggende schalen (ook wel facetten genoemd) bleek dat studenten die een voorkeur hebben voor structuur, hard werken, duidelijke doelen stellen (facetten van Consciëntieusheid), bescheiden zijn en weinig waarde hechten aan materieel bezit en status (facetten van Integriteit), over het algemeen meer succesvol zijn op studiegebied dan studenten die deze eigenschappen in mindere mate bezitten.

Het onderzoek van De Vries en collega's (2011) is voor zover bekend het eerste onderzoek dat een relatie heeft aangetoond tussen de zesde, nieuwe persoonlijkheidsdimensie Integriteit en studieprestatie. Een interessante vraag is wat de reden is dat Integriteit een positieve relatie vertoont met studieprestatie. Deze bevinding zou verklaard kunnen worden door de gevonden significante en positieve relatie tussen het Integriteitsfacet Hebzuchtvermijding en studieprestatie (zie De Vries, De Vries et al., 2011). Studenten die laag scoren op het facet Hebzuchtvermijding, hechten veel waarde aan materieel bezit en status en geven daarom wellicht de voorkeur aan een bijbaan om zodoende geld te kunnen verdienen voor de aanschaf van luxe goederen. Dit kan tot gevolg hebben dat deze studenten minder tijd hebben om te studeren. Eerder onderzoek heeft bevestigd dat de hoeveelheid tijd die studenten besteden aan een bijbaan, een nadelig effect heeft op studieprestatie (Butler, 2007).

Als het gaat om prestaties en gedrag op de werkvloer, heeft eerder onderzoek zich grotendeels gericht op persoonlijkheid als voorspeller van werkprestatie en van contraproductief werkgedrag (Rotundo & Sackett, 2002). Onder werkprestatie wordt veelal het algemeen functioneren van de werknemer verstaan (Borman & Motowidlo, 1997). De Big Five-dimensie Consciëntieusheid blijkt consequent een valide en stabiele voorspeller van prestaties op het werk te zijn (Barrick & Mount, 1991; Hurtz & Donovan, 2000). Dit betekent dat, ongeacht het beroep, consciëntieuze medewerkers over het algemeen beter presteren dan minder consciëntieuze collega's. Emotionele Stabiliteit kwam in de meta-analyse van Barrick et al. (2001) tevens als algemene voorspeller van werkprestatie naar voren, hoewel de relatie tussen deze dimensie aan de ene kant en specifieke prestatiecriteria en beroepen aan de andere kant minder consistent was dan in het geval van Consciëntieusheid. Terwijl de andere Big Five-persoonlijkheidsdimensies (Extraversie, Vriendelijkheid en Openheid voor Ervaringen) ook relaties laten zien met werkprestatie, blijkt dat deze relaties variëren afhankelijk van het beroep. Extraversie blijkt bijvoorbeeld voornamelijk een voorspeller te zijn van werkprestatie in beroepen waar het verkopen van diensten en goederen centraal staat (Barrick & Mount, 1991).

Verder is uit eerder onderzoek naar voren gekomen dat – naast Consciëntieusheid – Integriteit een van de beste voorspellers is van werkprestaties (Ones & Vis-

wesvaran, 2001; Schmidt & Hunter, 1998). Echter, in deze onderzoeken is Integriteit gemeten aan de hand van de eerdergenoemde open integriteitstests, waarbij rechtstreeks naar integriteitsschendingen wordt gevraagd. In een recente studie hebben Johnson, Rowatt en Petrini (2011) als eerste een significante en positieve relatie laten zien tussen werkprestatie en Integriteit als persoonlijkheidskenmerk binnen het HEXACO-model. Deze resultaten zijn echter gebaseerd op een steekproef bestaande uit medewerkers die uitsluitend in de zorg werkten. Een specifiek doel van het huidige onderzoek is om te onderzoeken of de persoonlijkheidsdimensie Integriteit werkprestatie verklaart ongeacht het beroep, of dat onderzocht moet worden of de relatie tussen Integriteit en werkprestaties afhankelijk is van de specifieke kenmerken van een beroep.

Eerder onderzoek heeft aangetoond dat persoonlijkheid behalve voor werkprestaties ook een voorspeller is van contraproductief werkgedrag. Dit zijn opzettelijke of onopzettelijke handelingen die schadelijk zijn voor een organisatie of voor leden van een organisatie (Sackett & DeVore, 2001). Gedacht kan worden aan gedragingen zoals diefstal, misbruik van informatie, misbruik van tijd en middelen, onveilig gedrag, verzuim, slechte kwaliteit van het werk en alcohol- en drugsmisbruik (Gruys & Sackett, 2003). Uit eerder onderzoek blijkt dat de Big Five-dimensie Consciëntieusheid als sterkste (negatieve) voorspeller van contraproductief werkgedrag naar voren kwam (Salgado, 2002). Bovendien blijkt uit ander onderzoek dat, binnen het HEXACO-model, voornamelijk Integriteit variantie verklaart in contraproductieve werkgedragingen. Integriteit is bijvoorbeeld een negatieve voorspeller van antisociaal gedrag gericht op organisaties (Lee, Ashton & Shin, 2005) en van delinquentie op de werkplek (Lee, Ashton & De Vries, 2005). Met andere woorden, consciëntieuze en integere werknemers blijken over het algemeen minder contraproductieve gedragingen op de werkvloer te laten zien dan werknemers die deze eigenschappen in mindere mate bezitten. Op basis van bovengenoemde bevindingen zijn in het huidige onderzoek de volgende hypothesen geformuleerd:

*Hypothese 1a:* Consciëntieusheid en Integriteit vertonen een positieve relatie met werkprestatie.

*Hypothese 1b:* Consciëntieusheid en Integriteit vertonen een negatieve relatie met contraproductief werkgedrag.

Om bovenstaande hypothesen te testen zijn twee afzonderlijke studies uitgevoerd. Voor de eerste studie hebben we gebruikgemaakt van een sneeuwbaltechniek om allerlei werknemers te benaderen via e-mail, telefoon en individuele contacten. Werknemers die bereid waren deel te nemen ( $N = 238$ ,  $M_{leeftijd} = 32.9$ ,  $SD = 11.1$ , 47.9% vrouw), vulden online een vragenlijst in bestaande uit persoonsgegevens, de gereviseerde HEXACO Persoonlijkheidsvragenlijst (HEXACO-PI-R) en vragen over contraproductief werkgedrag. Vervolgens werd aan de leidinggeevenden gevraagd om de werkprestatie van de werknemers te beoordelen door online een aantal vragen te beantwoorden, zoals of de desbetreffende werknemer laag, gemiddeld of hoog presteerde in vergelijking met andere werknemers in vergelijkbare functies. Het opleidingsniveau van de werknemers varieerde van middelbare

school (6.8%), lbo (3.8%), mbo (25.6%), hbo (37.8%) tot wo (26.1%) niveau. Voor de tweede studie ( $N = 254$ ,  $M_{leeftijd} = 33.9$ ,  $SD = 11.1$ , 54.3% vrouw) hebben we een vergelijkbare aanpak gekozen, met als belangrijkste verschil dat de deelnemers een andere persoonlijkheidsvragenlijst hebben ingevuld, namelijk de Multiculturele Persoonlijkheidstest – Big Six (MPT-BS). Het opleidingsniveau van de werknemers varieerde van middelbare school (4.7%), lbo (0.8%), mbo (25.6%), hbo (41.7%) tot wo (27.2%) niveau. De deelnemers van beide studies waren werkzaam in allerlei beroepssectoren (van administratie, IT, financiële dienstverlening tot de medische sector) en hadden verschillende beroepen, zoals magazijnmedewerker, leraar/lerares, adviseur, verpleegkundige en journalist.

De resultaten van beide studies laten zien dat Consciëntieusheid positief en significant gerelateerd is aan werkprestatie ( $r = .21$ ,  $p < .01$  respectievelijk  $r = .23$ ,  $p < .01$ ). Dit is in overeenstemming met eerdere studies die de relatie hebben onderzocht tussen persoonlijkheid en werkprestatie (bijv. Barrick & Mount, 1991; Hurtz & Donovan, 2000). De resultaten op facetniveau laten zien dat de Consciëntieusheidsfacetten IJver ( $r = .27$ ,  $p < .01$ ), Perfectionisme ( $r = .19$ ,  $p < .01$ ), Doorzettingsvermogen ( $r = .21$ ,  $p < .01$ ) en Prestatiemotivatie ( $r = .35$ ,  $p < .01$ ) de sterkste positieve relaties vertoonden met werkprestatie. Dit betekent dat werknemers die hard werken, die oog hebben voor detail, die afmaken waaraan ze begonnen zijn en die een zo goed mogelijke prestatie willen leveren, over het algemeen hogere beoordelingen kregen van hun leidinggevendenden. In beide studies werd echter geen bewijs gevonden voor de veronderstelling dat Integriteit een relatie vertoont met prestaties op het werk. Geconcludeerd kan worden dat hypothese 1a deels aangenomen kan worden: Consciëntieusheid vertoont een significante en positieve relatie met werkprestatie, maar Integriteit vertoont deze relatie niet.

Met betrekking tot contraproductief werkgedrag bleek dat Consciëntieusheid ( $r = -.37$ ,  $p < .01$  respectievelijk  $r = -.43$ ,  $p < .01$ ) en Integriteit ( $r = -.47$ ,  $p < .01$  respectievelijk  $r = -.33$ ,  $p < .01$ ) significant en negatief gecorreleerd waren aan dergelijk werkgedrag (hypothese 1b wordt hiermee aangenomen): werknemers die een voorkeur hebben voor regels, die een gestructureerde stijl van werken aannemen, die hun impulsen weten te remmen en die eerlijk zijn tegenover anderen, zullen over het algemeen minder contraproductieve werkgedragingen laten zien dan werknemers die minder consciëntieus en integer zijn. Op facetniveau bleek tevens dat een aantal facetten van Openheid en Extraversie voorspellend was: werknemers die op zoek zijn naar spanning, die creatief zijn, graag in de schijnwerpers staan en die openstaan voor schijnbaar vreemde of radicale ideeën, hebben over het algemeen *meer* kans om contraproductieve werkgedragingen te vertonen dan andere werknemers.

Het doel van deze studies was om de relatie tussen de relatief nieuwe persoonlijkheidsdimensie Integriteit aan de ene kant en werkprestatie en contraproductief werkgedrag aan de andere kant te onderzoeken. Geconcludeerd kan worden dat Integriteit vooral een relatie vertoont met contraproductief werkgedrag en *niet* met werkprestatie. Dit laatste is tegen de verwachting in, omdat recent onderzoek heeft aangetoond dat hoe meer integer, eerlijk en bescheiden een werknemer is, hoe hoger zijn of haar prestaties op het werk zijn (Johnson et al., 2011).

Hierbij moet echter worden opgemerkt dat het onderzoek van Johnson en collega's zich richtte op werknemers uit een specifieke beroepssector, namelijk de zorgverlening. In het huidige onderzoek is echter gekeken naar de relatie tussen persoonlijkheid en werkprestatie onder werknemers uit verschillende beroepssectoren. Men dient dus voorzichtig te zijn met de conclusie dat Integriteit een algemene voorspeller is van prestatie op het werk. Het zou kunnen zijn dat Integriteit vooral van belang is voor het functioneren van werknemers in organisaties die zorg verlenen aan andere mensen, maar dat werknemers met minder integriteit, eerlijkheid en bescheidenheid wellicht beter presteren in organisaties waar zelfpromotie belangrijk is, zoals in organisaties waar activiteiten gericht op de verkoop centraal staan. Vervolgonderzoek zou kunnen nagaan of de relatie van Integriteit met werkprestaties afhankelijk is van de specifieke kenmerken van een beroep.

### 3 Het 'bandwidth-fidelity dilemma'

De hierboven genoemde onderzoeksresultaten tonen aan dat persoonlijkheid een relatie laat zien met contraproductief werkgedrag en prestaties op de werkvloer. Volgens eerder onderzoek kan persoonlijkheid bepaalde (werk)criteria, zoals werkprestatie, beter voorspellen als gekeken wordt naar de specifieke onderliggende schalen van de brede dimensies (bijv. Paunonen, Rothstein & Jackson, 1999). Persoonlijkheidsmodellen zijn meestal hiërarchisch gestructureerd. Het laagste niveau bestaat uit smalle, specifieke persoonlijkheidsstreken die gecombineerd kunnen worden om bredere persoonlijkheidsfactoren te vormen, zoals de Big Five- of HEXACO-dimensies (bijv. Goldberg, 1990). Onderzoekers hebben de keuze om zich te richten op de brede, heterogene persoonlijkheidsfactoren – d.w.z. persoonlijkheidsfactoren met een grote bandbreedte ('bandwidth') – of op de smallere, homogene persoonlijkheidsfacetten – d.w.z. persoonlijkheidsfacetten die een trek nauwkeurig meten (met 'high fidelity'). Consciëntieusheid als samengestelde persoonlijkheidsfactor omvat bijvoorbeeld een breder en abstracter domein dan een onderliggend facet, zoals het Consciëntieusheidsfacet Zelfdiscipline, dat een specifiekere concept omvat en gedrag nauwkeuriger meet. Het bovenstaande staat bekend als het 'bandwidth-fidelity dilemma' (Cronbach & Gleser, 1965).

Voorstanders van het gebruik van de brede persoonlijkheidsdimensies zijn van mening dat de meeste criteria, zoals werkprestaties, niet eenduidige concepten zijn en dat soortgelijke samengestelde voorspellers nodig zijn om de voorspellende waarde te maximaliseren (Ones & Viswesvaran, 1996). Prestatiecriteria zouden voorspeld moeten worden door brede persoonlijkheidsdimensies, omdat smallere eigenschappen meestal gemeten worden met minder items, wat resulteert in lagere betrouwbaarheden van deze schalen. Dit zal tot gevolg hebben dat de gevonden correlaties worden afgezwakt en de relaties tussen persoonlijkheid en werkprestaties worden onderschat (bijv. Ones & Viswesvaran, 1996).

Vanuit een ander oogpunt hebben onderzoekers gepleit voor het gebruik van de smalle facetten. De brede persoonlijkheidsfactoren zouden minder nuttig zijn

voor het voorspellen van prestatie en gedrag door de heterogeniteit van hun onderliggende facetten. Zo kan het zijn dat een aantal facetten sterke correlaties vertoont met een bepaald criterium terwijl andere facetten, behorend tot dezelfde dimensie, nauwelijks of helemaal niet gerelateerd zijn aan het criterium. Dit verzwakt vervolgens de predictieve validiteit van de gehele factor (O'Neill & Allen, 2011). Er wordt betoogd dat een facet bestaat uit psychometrisch overeenkomende items en dat een facet daarom afzonderlijk een unieke bijdrage kan leveren aan de voorspelling van gedrag (Paunonen et al., 1999). Tot slot, sommige onderzoekers zijn van mening dat het gebruik van brede persoonlijkheidsdimensies leidt tot een verlies van gedetailleerde informatie: smalle facetten kunnen diepere inzichten geven in de persoonlijkheid van een individu en daardoor zorgen voor meer inzicht in de relaties tussen persoonlijkheid en (werk)gedrag (bijv. Tett, Steele & Beauregard, 2003).

Een aantal studies heeft tot nu toe aangetoond dat de afzonderlijke facetten van de brede persoonlijkheidsfactoren samen betere voorspellers zijn dan de brede factoren zelf (bijv. Ashton, 1998; Paunonen & Ashton, 2001; Paunonen, Haddock, Forsterling & Keinonen, 2003). Christiansen en Robie (2011) benadrukken daarnaast het belang van studies die zich concentreren op zowel de brede als de smalle persoonlijkheidskenmerken, aangezien tot nu toe relatief weinig studies verschenen zijn die dit in combinatie gedaan hebben. Daarom zal het huidige onderzoek zich richten op zowel de brede als smalle persoonlijkheidskenmerken en is de volgende hypothese opgesteld:

*Hypothese 2:* De onderliggende facetten van de brede persoonlijkheidsfactoren zijn samen beter in staat werkprestatie en contraproductief werkgedrag te voorspellen dan de brede factoren zelf.

Meerdere empirische studies in het proefschrift van De Vries (2012) laten zien dat bij het voorspellen van allerlei complexe criteria, zoals studie- en werkprestaties en contraproductief gedrag, de hoeveelheid verklaarde variantie door de smalle onderliggende schalen hoger was dan door de brede persoonlijkheidsdimensies (hypothese 2 wordt hiermee aangenomen). Dit heeft onder andere te maken met het feit dat het ene facet een sterkere correlatie vertoonde met een bepaald criterium terwijl een ander facet, behorend tot dezelfde dimensie, minder of helemaal niet gerelateerd was aan het criterium. Hierdoor werd de predictieve validiteit van de gehele factor afgezwakt. In het onderzoek naar het voorspellen van werkprestatie kwam bijvoorbeeld naar voren dat het Consciëntieusheidsfacet Behoefte aan regels en zekerheid geen enkele relatie vertoonde met werkprestatie ( $r = .00$ ,  $ns$ ), terwijl het Consciëntieusheidsfacet Prestatiemotivatie ( $r = .35$ ,  $p < .01$ ) een significante en positieve relatie vertoonde met werkprestatie. Op het gebied van werkgedrag bleek verder bijvoorbeeld dat de persoonlijkheidsdimensie Openheid niet significant gerelateerd was aan contraproductief werkgedrag. Daarentegen kwam naar voren dat het Openheidsfacet Initiatief significant en negatief gerelateerd was aan contraproductief werkgedrag, terwijl het Openheidsfacet Spanningsbehoefte een significante maar positieve correlatie vertoonde met contraproductief werkgedrag.



Deze resultaten maken duidelijk dat een uitsluitende gerichtheid op de brede eigenschappen zal leiden tot een verlies van relevante informatie: de smalle persoonlijkheidskenmerken geven een meer specifiek inzicht in iemands persoonlijkheid en de relatie met zijn of haar gedrag en prestaties.

#### *Etnische scoreverschillen*

Tot nu toe heeft een kleine groep van onderzoekers in Nederland etnische scoreverschillen op persoonlijkheidstests onderzocht. Van Leest (1997) heeft bijvoorbeeld aangetoond dat Turkse allochtonen significant lager scoorden op assertiviteit en significant hoger scoorden op het vermijden van onzekerheid (gerelateerd aan Consciëntieusheid) dan de Nederlandse autochtonen. De Meijer, Born, Terlouw en Van der Molen (2006) vonden verder dat etnische minderheden, waaronder Turkse allochtonen, significant hoger scoorden op Consciëntieusheid dan de Nederlandse autochtonen. Echter, vooral op het gebied van persoonlijkheidstests blijven de meeste bevindingen op het gebied van etnische scoreverschillen onverklaard. Hoewel eerdere onderzoekers hebben geconcludeerd dat de verschillen tussen culturele groepen op persoonlijkheidsvragenlijsten klein zijn (bijv. Ployhart & Holtz, 2008), is er nog nauwelijks onderzoek gedaan naar mogelijke groepsverschillen op de specifieke onderliggende schalen. Aangezien een uitsluitende gerichtheid op de brede persoonlijkheidsfactoren belangrijke relaties tussen de persoonlijkheidsfacetten en criteria kan maskeren, is een interessante vraag of het combineren van facetten tot een brede persoonlijkheidsfactor ook tot gevolg kan hebben dat etnische scoreverschillen op facetniveau ten onrechte gemaskeerd worden. Voor zover er daadwerkelijk groepsverschillen zijn op specifieke schalen, kan dit gevolgen hebben voor de inzet van persoonlijkheidsinstrumenten in selectiesituaties.

Een studie beschreven in het proefschrift van De Vries (2012) heeft etnische scoreverschillen op een persoonlijkheidstest onderzocht tussen autochtone ( $N = 211$ ,  $M_{leeftijd} = 34.8$ ,  $SD = 11.3$ , 55.5% man, 69.7% hbo/wo-niveau) en Turkse allochtone ( $N = 81$ ,  $M_{leeftijd} = 31.4$ ,  $SD = 8.6$ , 48.1% man, 77.7% hbo/wo-niveau) werknemers. De autochtone werknemers maakten deel uit van de eerdergenoemde studie naar de relatie tussen persoonlijkheid en werkprestatie. Om deze steekproef aan te vullen met Turkse allochtone werknemers hebben we gebruikgemaakt van een sneeuwbaltechniek om Turkse allochtonen te benaderen via e-mail, telefoon en individuele contacten. Om mee te mogen doen aan het onderzoek moesten de proefpersonen, naast het hebben van een Turkse achtergrond en het beheersen van de Nederlandse taal, minimaal één jaar werkzaam zijn bij hun huidige werkgever. Analyses om de autochtone en de Turkse allochtone werknemers te vergelijken in termen van geslacht, leeftijd en opleidingsniveau, toonden aan dat de autochtone werknemers significant ouder waren ( $t = 2.79$ ,  $p < .01$ ).

Met betrekking tot etnische scoreverschillen op de brede persoonlijkheidsfactoren bleek dat Turkse allochtone werknemers significant hoger scoorden op Consciëntieusheid dan de Nederlandse autochtone werknemers. Dit is in overeenstemming met eerder onderzoek van Van Leest (1997) en De Meijer et al. (2006). Met betrekking tot de smallere facetschalen kwam naar voren dat Turkse alloch-

tone werknemers significant hoger scoorden dan autochtone werknemers op de Consciëntieusheidsfacetten Ordelijkheid en Behoeftte aan regels en zekerheid, op het Vriendelijkheidsfacet Aandachtsbehoefte, op het Openheidfacet Leergierigheid en op het Integriteitsfacet Eerlijkheid. Tot slot bleek dat Turkse werknemers vergeleken met autochtone werknemers significant lager scoorden op het Emotionele Stabiliteitsfacet Emotionele beheersing, het Vriendelijkheidsfacet Vertrouwen in anderen, het Openheidfacet Spanningsbehoefte en het Integriteitsfacet Hebzuchtvermijding.

Bovenstaande resultaten geven aan dat de brede factoren etnische scoreverschillen lijken te maskeren. Vooral in het geval van Vriendelijkheid, Openheid en Integriteit scoorden allochtone werknemers significant lager op één facet en significant hoger op een ander facet dan autochtone werknemers, terwijl beide facetten behoren tot dezelfde brede factor. Aangezien de factoren Vriendelijkheid, Openheid en Integriteit zelf geen etnische scoreverschillen lieten zien, suggereert dit dat het combineren van facetten tot een brede persoonlijkheidsfactor tot gevolg kan hebben dat etnische scoreverschillen op facetniveau ten onrechte gemaskeerd worden. Dit betekent dat voor een compleet beeld van de persoonlijkheid van autochtone en allochtone werknemers niet alleen gekeken zou moeten worden naar de scores op de persoonlijkheidsfactoren, maar ook naar de scores op facetniveau.

Er is echter enige voorzichtigheid geboden bij de interpretatie van de gevonden persoonlijkheidsverschillen: de Turkse allochtone groep in dit onderzoek, bestaande uit *werknemers* met een relatief hoog opleidingsniveau en een goede beheersing van de Nederlandse taal, is niet representatief voor de gehele Turkse allochtone groep die momenteel woonachtig is in Nederland. Vervolgonderzoek zou kunnen nagaan of de gevonden etnische scoreverschillen echte verschillen in persoonlijkheid weergeven en wat de consequenties van deze eventuele scoreverschillen zijn indien persoonlijkheid wordt meegenomen in selectiesituaties. Daarnaast dient de vraag beantwoord te worden of er mogelijk andere factoren van invloed zijn geweest op het invullen van de persoonlijkheidstest, zoals sociale wenselijkheid of het reference-groep effect (zie §5).

#### 4 Frame-of-reference effect

Een tweede manier om de voorspellende waarde van persoonlijkheid te verbeteren, naast het gebruik van smalle persoonlijkheidskenmerken, is het toevoegen van een specifieke context aan elk persoonlijkheidsitem, zoals 'ik ben ordelijk op het werk' (frame-of-reference). Volgens Wright en Mischel (1987) gedraagt een individu zich consistent en voorspelbaar binnen *vergelijkbare* situaties, maar gedraagt hij of zij zich niet altijd consistent in *verschillende* situaties. In het geval van een algemene (d.w.z. niet-gecontextualiseerde) persoonlijkheidstest zou het kunnen dat een respondent denkt aan een werksituatie bij het beantwoorden van één item, maar aan een thuissituatie bij de beantwoording van het volgende item. Aangezien iemand niet altijd dezelfde gedragingen laat zien in verschillende situaties, kan dit resulteren in inconsistente antwoorden (*within-person inconsistency*;

Lievens et al., 2008). Daarnaast is beargumenteerd dat een persoonlijkheidstest zonder context tot gevolg kan hebben dat de ene respondent denkt aan een thuis-situatie tijdens het beantwoorden van een bepaalde vraag, terwijl een andere respondent denkt aan een werksituatie bij het beantwoorden van dezelfde vraag (*between-person variability*; bijv. Bing, Whanger, Davison & VanHook, 2004). Dit leidt vervolgens tot irrelevante scoreverschillen tussen respondenten. Lievens en collega's (2008) hebben aangetoond dat het toevoegen van een relevante context aan persoonlijkheidsitems, zoals de context 'op school' als men studieprestaties wil voorspellen, leidt tot meer consistente antwoorden (*within-* en *between-*respondenten) tijdens het invullen van een persoonlijkheidstest. Hierdoor gaan de betrouwbaarheid en de voorspellende waarde van de persoonlijkheidstest omhoog. Daarnaast hebben Lievens en collega's aangetoond dat het gebruik van een context die niet relevant is voor het criterium (bijvoorbeeld de context 'op het werk' als men het niveau van studieprestaties wil voorspellen), leidt tot een lagere voorspellende waarde van persoonlijkheid.

Met betrekking tot de voorspelling van *werkprestaties* is het frame-of-reference effect nog nauwelijks onderzocht. Daarom richt dit onderzoek zich op de vraag of het toevoegen van een werkcontext aan persoonlijkheidsitems een positieve invloed heeft op de voorspellende waarde van persoonlijkheid bij het voorspellen met werkprestaties en -gedrag. De volgende twee hypothesen zijn opgesteld:

*Hypothese 3a:* Werkspecifieke persoonlijkheidsschalen zijn sterker gerelateerd aan prestaties op het werk en contraproductief werkgedrag dan niet-gecontextualiseerde en (irrelevante) thuispecifieke persoonlijkheidsschalen.

*Hypothese 3b:* Thuispecifieke persoonlijkheidsschalen laten een zwakkere relatie zien met prestaties op het werk en contraproductief werkgedrag dan de overeenkomstige niet-gecontextualiseerde persoonlijkheidsschalen.

Het huidige onderzoek heeft, vergelijkbaar met eerdere studies (bijv. Lievens et al., 2008), gebruikgemaakt van een within-subject design. In totaal beantwoordden 289 werknemers ( $M_{leeftijd} = 37.9$ ,  $SD = 14.1$ , 77.9% vrouw) online 168 persoonlijkheidsvragen behorende tot de persoonlijkheidsdimensies Consciëntieusheid en Integriteit: een derde niet-gecontextualiseerde items, een derde werkspecifieke items en een derde thuispecifieke items. Om de lengte van de persoonlijkheidstest te beperken is ervoor gekozen alleen de persoonlijkheidsschalen mee te nemen die relevant waren voor de hypothesen. Met betrekking tot de niet-gecontextualiseerde items ontvingen deelnemers de standaardinstructies en werd gevraagd aan te geven in hoeverre ze het eens dan wel oneens waren met iedere stelling. Eveneens in overeenstemming met eerdere studies (Bing et al., 2004; Lievens et al., 2008; Robie, Schmit, Ryan & Zickar, 2000; Schmit, Ryan, Stierwalt & Powell, 1995) zijn de contextspecifieke persoonlijkheidsschalen ontworpen door het toevoegen van een tag, in dit geval 'op het werk' of 'thuis', aan elk item. Het item 'ik wil afmaken waar ik aan begonnen ben' werd bijvoorbeeld veranderd in 'ik wil afmaken waar ik aan begonnen ben *op het werk*' en 'ik wil afmaken waar ik aan begonnen ben *thuis*'. Als meetinstrument is gekozen voor de Multiculturele Persoonlijkheidstest – Big Six, omdat na het toevoegen van de

tags, de items nog steeds logisch geformuleerd bleken te zijn waardoor het niet nodig was om deze te herschrijven. Een week nadat de deelnemers de persoonlijkheidsvragen online hadden beantwoord, kregen ze toegang tot een tweede vragenlijst, bestaande uit achtergrondvariabelen en vragen over prestaties op het werk en contraproductief werkgedrag. Het opleidingsniveau van de werknemers varieerde van basisonderwijs (0.7%), middelbare school (6.6%), lbo (1.4%), mbo (19.7%), hbo (34.6%) tot wo (37.0%) niveau. De deelnemers werkten in allerlei beroepssectoren, waaronder de gezondheidszorg en de zakelijke dienstverlening. Uit de resultaten blijkt dat Consciëntieusheid verschillende correlaties vertoonde met werkprestatie in de werkspecifieke ( $r = .44, p < .01$ ), de niet-gecontextualiseerde ( $r = .25, p < .01$ ) en de thuisspecifieke ( $r = .07, ns$ ) conditie. Verdere analyses toonden aan dat de werkspecifieke Consciëntieusheidschaal significant sterker gerelateerd was aan werkprestatie dan de niet-gecontextualiseerde en (irrelevante) thuisspecifieke Consciëntieusheidschalen (zie voor deze en onderstaande resultaten De Vries, 2012). Daarnaast kwam naar voren dat de thuisspecifieke Consciëntieusheidschaal significant een zwakkere relatie met werkprestatie liet zien dan de overeenkomstige niet-gecontextualiseerde schaal. Integriteit liet geen significante relatie zien met werkprestatie in de niet-gecontextualiseerde ( $r = .09, ns$ ) en thuisspecifieke ( $r = .06, ns$ ) conditie. Hoewel de werkspecifieke Integriteitschaal een significante relatie vertoonde met werkprestatie ( $r = .12, p < .05$ ), was het verschil met de andere condities te klein waardoor deze relatie niet *significant* sterker was.

Met betrekking tot contraproductief werkgedrag lieten de resultaten zien dat Consciëntieusheid een significante en negatieve correlatie vertoonde in de werkspecifieke ( $r = -.45, p < .01$ ), de niet-gecontextualiseerde ( $r = -.32, p < .01$ ) en de thuisspecifieke ( $r = -.16, p < .01$ ) conditie. Zoals verwacht was de correlatie tussen Consciëntieusheid en contraproductief werkgedrag significant sterker in de werkspecifieke conditie dan in de twee andere condities. Daarnaast was de correlatie tussen Consciëntieusheid en contraproductief werkgedrag significant minder sterk in de thuisspecifieke conditie dan in de niet-gecontextualiseerde conditie. Integriteit was tevens significant en negatief gerelateerd aan contraproductief werkgedrag in de werkspecifieke ( $r = -.40, p < .01$ ), de niet-gecontextualiseerde ( $r = -.30, p < .01$ ) en de thuisspecifieke ( $r = -.29, p < .01$ ) conditie. Terwijl de niet-gecontextualiseerde en de thuisspecifieke schalen vergelijkbare correlaties vertoonden, liet de werkspecifieke schaal significant een sterkere relatie zien met contraproductief werkgedrag.

Er kan geconcludeerd worden dat, in het bijzonder voor Consciëntieusheid, een werkspecifieke persoonlijkheidsschaal sterker gerelateerd is aan werkprestatie en aan contraproductief werkgedrag dan een niet-gecontextualiseerde en (irrelevante) thuisspecifieke persoonlijkheidsschaal (hypothese 3a deels aangenomen). Daarnaast blijkt dat een thuisspecifieke persoonlijkheidsschaal een zwakkere relatie met deze werkcriteria laat zien dan de overeenkomstige niet-gecontextualiseerde persoonlijkheidsschaal (hypothese 3b deels aangenomen). Dit bevestigt het idee dat door het toevoegen van een relevante context aan persoonlijkheidsitems de voorspellende waarde van persoonlijkheid verhoogd zou kunnen worden. Met andere woorden, als men het niveau van werkprestaties wil voorspellen,

zal de voorspellingskracht van persoonlijkheid verbeterd kunnen worden door aan de items een relevante context, in dit geval 'op het werk', toe te voegen. Het toevoegen van een context die niet relevant is voor het criterium (bijv. 'op school' als men werkprestaties wil voorspellen) lijkt te resulteren in een lagere voorspellende waarde van persoonlijkheid.

Het huidige onderzoek heeft geen aandacht besteed aan etniciteit met betrekking tot het frame-of-reference effect. Interessant is echter dat cross-cultureel onderzoek laat zien dat westerse versus niet-westerse groepen op een andere manier reageren op zelfrapportagevragenlijsten. Mensen met een westerse achtergrond hebben over het algemeen minder moeite zichzelf te beschrijven zonder enige specifieke context, terwijl mensen met een niet-westerse achtergrond zichzelf van nature meer beschrijven in de context van hun groep, het 'wij'-principe (bijv. Cousins, 1989). Het plaatsen van persoonlijkheidsitems in een context (op school, op het werk, thuis, in mijn vrije tijd) zal daarom kunnen overkomen als een meer realistische en natuurlijke manier van zelfrapportage voor mensen met een niet-westerse achtergrond. Een interessant onderwerp voor toekomstig onderzoek zal zijn of – in het bijzonder voor niet-westerse kandidaten – een gecontextualiseerde persoonlijkheidstest zal leiden tot een hogere betrouwbaarheid en voorspellende waarde van de test.

## 5 Reference-group effect

Ten slotte is onderzocht of de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest, effect heeft op de antwoorden die deze persoon geeft (reference-group effect; Heine et al., 2002). De sociale vergelijkingstheorie (social comparison theory; Festinger, 1954) gaat ervan uit dat mensen de neiging hebben hun eigen gedrag te vergelijken met dat van vergelijkbare anderen uit hun eigen omgeving. Dit kan betekenen dat wanneer een kandidaat een persoonlijkheidstest invult, de antwoorden op de items niet gebaseerd zijn op het absolute niveau van een persoonlijkheidskenmerk maar op hoe de respondent zichzelf beoordeelt ten opzichte van een bepaalde vergelijkingsgroep (Credé, Bashshur & Niehorster, 2010).

In het geval van lengte bijvoorbeeld: als Can uit Turkije zich qua lengte moet beoordelen op een 5-punts Likertschaal (met ankers variërend van zeer lang tot zeer kort), zal hij zichzelf met een lengte van 178 centimeter beoordelen als relatief lang in vergelijking met andere Turkse mannen (gemiddelde lengte 174 centimeter). In Nederland ligt de gemiddelde mannelijke lengte hoger (184 centimeter), dus Sem uit Nederland met een lengte van 180 centimeter, die in feite langer is dan Can, zal zichzelf als relatief kort beoordelen vergeleken met andere mannen in Nederland. Indien deze beoordelingen van Can en Sem direct worden vergeleken, zou men ten onrechte kunnen concluderen dat Can langer is dan Sem.

Ook in het geval van crossculturele persoonlijkheidsvergelijkingen kunnen ware verschillen gemaskeerd worden als iedereen zich vergelijkt met soortgelijke anderen, dat wil zeggen, met mensen uit de eigen culturele groep (in-group). Meestal worden ordinale of (quasi-)intervalschalen, zoals de hierboven genoemde Likert-

schaal, gebruikt voor crossculturele persoonlijkheidsvergelijkingen, terwijl deze schalen bijzonder gevoelig zijn voor het reference-group effect. Mensen met een Chinese achtergrond bijvoorbeeld zullen zichzelf waarschijnlijk vergelijken met de norm van hun eigen Chinese culturele groep, die altijd is gelegen rond het midden van de schaal. Aan de andere kant zullen mensen met een Amerikaanse achtergrond zichzelf vergelijken met de norm van hun culturele groep, die ook is gelegen rond het midden van de schaal (Peng, Nisbett & Wong, 1997). Als gevolg hiervan kunnen ware persoonlijkheidsverschillen tussen de culturele of etnische groepen worden gemaskeerd.

Crede en collega's (2010) waren onlangs de eersten die het reference-group effect hebben onderzocht op basis van zelfgerapporteerde *persoonlijkheidskenmerken*. Deelnemers moesten zichzelf tijdens het invullen van de persoonlijkheidstest vergelijken met verschillende referentiegroepen, variërend in leeftijd, geslacht en vorm van de relatie (zoals familie, vrienden en mensen in het algemeen). Uit de resultaten bleek dat de specifieke vergelijkingsgroep die iemand in gedachten had tijdens het beantwoorden van persoonlijkheidsvragen, effect had op *welke* antwoorden die persoon gaf. Er zijn echter geen studies uitgevoerd om het reference-group effect te onderzoeken bij het invullen van een persoonlijkheidstest waarbij verschillende etnische referentiegroepen gebruikt worden. Aangezien in Nederland nog altijd sprake is van een toenemende diversiteit aan mensen met verschillende culturele achtergronden, hebben wij onderzocht of mensen met een Turkse achtergrond die woonachtig zijn in Nederland, andere antwoorden geven op een persoonlijkheidstest wanneer zij zichzelf vergelijken met mensen uit hun eigen Turkse cultuur (in-group) dan wanneer zij zichzelf vergelijken met Nederlandse autochtonen (out-group). Uitgaande van het reference-group effect was de verwachting (zie voor specifieke hypothesen het proefschrift van De Vries, 2012) dat wanneer Turkse allochtonen zichzelf vergelijken met vergelijkbare anderen, in dit geval met mensen uit hun eigen etnische groep (in-group vergelijking), er geen persoonlijkheidsverschillen tussen de autochtone en Turkse allochtone groep gevonden worden. We verwachtten echter wel scoreverschillen tussen autochtonen en Turkse allochtonen wanneer laatstgenoemden zichzelf vergelijken met de autochtone groep (out-group vergelijking).

Om het reference-group effect te onderzoeken hebben we een e-mail gestuurd naar een Turks-Nederlandse netwerkorganisatie en naar psychologiestudenten van de Vrije Universiteit Amsterdam. Vervolgens is aan de Turkse allochtone groep gevraagd om online een algemene persoonlijkheidstest met de standaardinstructies in te vullen ( $N = 95$ ,  $M_{\text{leeftijd}} = 25.2$ ,  $SD = 5.1$ , 25.3% man, 45% hbo- en 51% wo-niveau). Verder hebben 74 Nederlandse autochtone studenten ( $M_{\text{leeftijd}} = 29.6$ ,  $SD = 10.2$ , 26.4% man, 89% wo-niveau) dezelfde persoonlijkheidstest ingevuld met alleen de standaardinstructies. Een week later hebben we de Turkse allochtonen gevraagd om weer dezelfde persoonlijkheidsvragenlijst in te vullen, maar nu met de specifieke instructie om zichzelf te vergelijken met autochtone Nederlanders (out-group vergelijking). Nog een week later werd aan de deelnemers gevraagd de persoonlijkheidstest in te vullen met de instructie om zichzelf te vergelijken met mensen uit hun eigen Turkse allochtone groep (in-group vergelijking). Uiteindelijk hebben 34 Turkse allochtonen ( $M_{\text{leeftijd}} = 24.8$ ,  $SD = 4.3$ ,

17.6% man) alle drie de versies ingevuld. Analyses om de autochtone en de Turkse allochtone deelnemers te vergelijken in termen van geslacht, leeftijd en opleidingsniveau, toonden aan dat de autochtone Nederlanders ouder waren ( $t = 3.32$ ,  $p < .01$ ) en een hoger opleidingsniveau hadden ( $t = 4.12$ ,  $p < .01$ ). Vanwege de verschillen tussen de twee groepen en de relatief kleine Turkse allochtone groep is enige voorzichtigheid geboden als het gaat om de interpretatie van onderstaande resultaten.

De resultaten laten geen scoreverschillen zien op de persoonlijkheidstest tussen autochtonen en Turkse allochtonen wanneer laatstgenoemden zich vergeleken met mensen uit hun eigen Turkse groep. Dit was inderdaad de verwachting: wanneer mensen zichzelf vergelijken met vergelijkbare anderen, in dit geval met mensen uit hun eigen etnische groep (in-group vergelijking), kunnen eventuele persoonlijkheidsverschillen tussen etnische groepen worden gemaskeerd. Toen de Turkse allochtonen echter moesten nadenken over hun persoonlijkheid en gedrag in vergelijking met de autochtone groep (out-group vergelijking), beoordeelden ze zichzelf als minder integer dan de autochtone groep. Ten slotte, wanneer de Turkse allochtonen zichzelf vergeleken met de autochtone groep, beoordeelden zij zichzelf als meer emotioneel en minder vriendelijk en minder open dan wanneer ze zichzelf vergeleken met hun eigen Turkse allochtone groep.

Deze bevindingen suggereren dat de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest, effect heeft op het soort antwoorden dat deze persoon geeft. Vervolgonderzoek zou moeten uitwijzen wat voor invloed dit heeft op de betrouwbaarheid en voorspellende waarde van de test.

## 6 Algemene conclusie en discussie

Het doel van deze studies was tweeledig. Het eerste doel was de relatie tussen de relatief nieuwe persoonlijkheidsdimensie Integriteit en werkprestatie – ongeacht het beroep – te onderzoeken. Uit de resultaten blijkt dat Consciëntieusheid een positieve relatie vertoont met werkprestatie, maar geconcludeerd kan worden dat Integriteit vooral een relatie vertoont met contraproductief werkgedrag en *niet* met werkprestatie. Dit laatste is tegen de verwachting in, omdat recent onderzoek heeft aangetoond dat hoe meer integer een werknemer is, hoe hoger zijn op haar prestaties op het werk zijn (Johnson et al., 2011). Het onderzoek van Johnson en collega's richtte zich echter op werknemers uit een specifieke beroepssector, namelijk de zorgverlening. In het huidige onderzoek is gekeken naar de relatie tussen persoonlijkheid en werkprestatie onder werknemers uit allerlei verschillende beroepssectoren: twee uitgevoerde studies hebben *geen* relatie tussen Integriteit als persoonlijkheidsdimensie en werkprestatie aan kunnen tonen. Men dient dus voorzichtig te zijn met de conclusie dat Integriteit een *algemene* voorspeller is van prestatie op het werk. Het zou kunnen zijn dat Integriteit vooral van belang is voor het functioneren van werknemers in organisaties die zorg verlenen aan andere mensen, maar dat werknemers met minder integriteit wellicht beter presteren in organisaties waar zelfpromotie belangrijk is, zoals in

organisaties waar activiteiten gericht op de verkoop centraal staan. Vervolgonderzoek zou kunnen nagaan of de relatie van Integriteit met werkprestaties afhankelijk is van de specifieke kenmerken van een beroep.

Het tweede doel was aan te tonen dat door meer specifieke persoonlijkheidsmetingen de voorspellende waarde van werkprestatie en contraproductief werkgedrag verbeterd kan worden. Ten eerste tonen onze bevindingen aan dat persoonlijkheid bepaalde werkcriteria, zoals werkprestatie en contraproductief werkgedrag, beter kan voorspellen als gekeken wordt naar de specifieke onderliggende schalen van de brede dimensies. Deze resultaten sluiten aan bij Christiansen en Robie (2011), die het belang van studies benadrukken die zich concentreren op zowel de brede als de smalle persoonlijkheidskenmerken, aangezien tot nu toe relatief weinig studies verschenen zijn die dit in combinatie hebben gedaan. Verder komt uit de resultaten naar voren dat het combineren van facetten tot een brede persoonlijkheidsfactor tevens tot gevolg kan hebben dat etnische scoreverschillen op facetniveau ten onrechte gemaskeerd worden. Deze resultaten suggereren dat het belangrijk is om rekening te houden met de scores op de persoonlijkheidsfacetten om zodoende een compleet beeld van iemands persoonlijkheid te krijgen en de relatie met prestatie en (werk)gedrag. Ten tweede werd onderzocht of door het toevoegen van een specifieke context aan persoonlijkheidsitems de voorspellingskracht van persoonlijkheid verhoogd kan worden. Met betrekking tot de voorspelling van werkprestatie is dit nog nauwelijks onderzocht. Interessant is dat het toevoegen van een specifieke en relevante context aan de persoonlijkheidsitems, in ons geval een werkcontext, leidt tot een hogere voorspellende waarde van de persoonlijkheidstest als het gaat om het voorspellen van werkprestatie en -gedrag. Het toevoegen van een context die niet relevant is voor werkprestatie (bijv. 'thuis'), lijkt te resulteren in een lagere voorspellende waarde van persoonlijkheid. Ten derde is nagegaan of de specifieke vergelijkingsgroep die iemand in gedachten heeft tijdens het invullen van een persoonlijkheidstest, effect heeft op het soort antwoorden dat iemand geeft. Tot nu toe is hier nog geen onderzoek naar gedaan, maar de bevindingen suggereren dat het gebruik van specifieke vergelijkingsgroepen de wijze waarop allochtone kandidaten een persoonlijkheidstest invullen, beïnvloedt. Er is echter meer onderzoek nodig om aan te kunnen geven welke gevolgen dit kan hebben voor de inzet van persoonlijkheidsinstrumenten in selectiesituaties.

Naar aanleiding van het huidige onderzoek kan de suggestie gewekt worden dat iemand die niet integer is, bij het invullen van een persoonlijkheidstest de vragen hierover minder eerlijk (dus sociaal wenselijk) invult. Ten eerste heeft integriteit niet alleen te maken met eerlijkheid. Binnen het HEXACO-model heeft Integriteit betrekking op de mate waarin personen hebzucht vermijden, en oprecht, rechtvaardig en bescheiden zijn. Een integer iemand is relatief ongeïnteresseerd in luxe, sociale status en privileges. Mensen met een gebrek aan integriteit daarentegen zetten zichzelf op de eerste plaats, zijn gehecht aan materiële zaken, kunnen de verleiding moeilijker weerstaan om regels te overtreden als ze er zelf beter van worden, en zullen anderen eerder vleien als dit helpt om persoonlijke doelen te realiseren. Ten tweede blijkt uit onderzoek dat anderbeoordelingen van integriteit in hoge mate overeenkomen met zelfbeoordelingen en dat sociale wenselijk-



heid niet het bovenstaande verwachte effect heeft op integriteitsscores (De Vries, Zettler & Hilbig, 2014). Ten derde speelt Integriteit vooral een rol als mensen er voordeel mee kunnen behalen. In veel situaties – zoals ontwikkel- en loopbaantrajecten – valt er weinig voordeel te halen als de vragenlijst ‘niet integer’ wordt ingevuld. Alleen in selectiesituaties zijn vooral schalen zoals Consciëntieusheid en Integriteit vatbaar voor sociaalwenselijke antwoorden, maar onderzoek toont aan dat dit effect niet sterk genoeg is om de voorspellende kracht van beide schalen te ondermijnen (Ones, Viswesvaran & Reiss, 1996).

Een belangrijke beperking van onze studies is dat contraproductief gedrag door de werknemers zelf is gerapporteerd. Studies hebben echter aangetoond dat zelfgerapporteerde contraproductieve gedragingen gerelateerd zijn aan objectieve metingen van contraproductief gedrag (Nicol & Paunonen, 2002). Daarom worden zelfgerapporteerde contraproductieve gedragingen vaak gebruikt als indicatie van contraproductief gedrag. Aangezien werknemers zowel persoonlijkheid als contraproductief gedrag zelf gerapporteerd hebben, is de kans aanwezig dat de wens om consistent te antwoorden een rol heeft gespeeld (*common method variance*). Spector (2006) stelt echter dat dit effect niet groot is. Ondanks deze geringe impact is het aan te raden om objectieve metingen te verzamelen. Een ander alternatief voor vervolgonderzoek zou kunnen zijn om naast zelfbeoordelingen, ook beoordelingen door directe collega's te verzamelen.

## 7 Implicaties voor de praktijk

In organisaties kan het nuttig zijn om sollicitanten – naast een intelligentietest – ook een persoonlijkheidstest in te laten vullen in het kader van selectie. Onze bevindingen suggereren dat het daarbij belangrijk is om rekening te houden met zowel de scores op de brede dimensies als de scores op de smalle onderliggende schalen. Een uitsluitende gerichtheid op de brede eigenschappen zal leiden tot een verlies van relevante informatie: de smalle persoonlijkheidskenmerken geven een meer specifiek inzicht in de persoonlijkheid van de kandidaat en de relatie met zijn of haar gedrag en uiteindelijke prestaties.

Voor het voorspellen van toekomstige werkprestaties is het belangrijk te kijken naar relevante Consciëntieusheidsfacetten: de resultaten van dit onderzoek tonen aan dat consciëntieuze werknemers die hard werken, die oog hebben voor detail, die afmaken waaraan ze begonnen zijn en die een zo goed mogelijke prestatie willen leveren, over het algemeen beter presteren op de werkvloer dan werknemers die deze eigenschappen in mindere mate bezitten.

Voor het voorspellen van contraproductief werkgedrag kan onder andere gekeken worden naar relevante Consciëntieusheids- en Integriteitsfacetten: werknemers die een voorkeur hebben voor regels, die een gestructureerde stijl van werken aannemen, die hun impulsen weten te remmen en die eerlijk zijn tegenover anderen, zullen over het algemeen minder contraproductieve werkgedragingen laten zien dan werknemers die minder consciëntieus en integer zijn. Daarnaast zullen werknemers die op zoek zijn naar spanning, die creatief zijn, graag in de schijnwerpers staan en die openstaan voor schijnbaar vreemde of radicale ideeën (facet-

ten van Openheid en Extraversie), over het algemeen *meer* kans hebben om contraproductieve werkgedragingen te vertonen dan andere werknemers.

De resultaten van het huidige onderzoek benadrukken verder dat door het toevoegen van een relevante context aan persoonlijkheidsvragen de voorspellende waarde van persoonlijkheidstests verhoogd kan worden. Met andere woorden, als men de mate van werkprestatie wil voorspellen, zal de voorspellingskracht van persoonlijkheidstests verbeterd worden als aan de items een relevante context, in dit geval 'op werk', wordt toegevoegd. Het toevoegen van een context die niet relevant is voor het criterium (bijvoorbeeld 'thuis' als men de mate van werkprestatie wil voorspellen), lijkt te resulteren in een lagere voorspellende waarde van persoonlijkheid. Deze bevindingen wijzen erop dat niet-gecontextualiseerde (of irrelevante gecontextualiseerde) persoonlijkheidstests minder informatie zullen opleveren dan relevante gecontextualiseerde persoonlijkheidsschalen als het gaat om het voorspellen van gedrag en prestatie op de werkvloer. Deze conclusie kan testontwerpers helpen om persoonlijkheidstests te ontwikkelen die niet alleen betrouwbaar, maar ook valide zijn in de situatie waarin de tests gebruikt gaan worden.

### Praktijkbox

Wat betekenen de resultaten van dit proefschrift voor de praktijk?

- Zet bij selectie- en intakeassessment – naast een capaciteitentest – ook een persoonlijkheidstest in.
- Kijk naar zowel de scores op de brede dimensies als naar de scores op de smalle onderliggende schalen voor een compleet overzicht van iemands persoonlijkheid.
- Kijk naar de scores op relevante Consciëntieusheidsfacetten voor het voorspellen van werkprestaties.
- Kijk naar de scores op – onder andere – relevante Consciëntieusheids- en Integriteitsfacetten bij het voorspellen van contraproductief werkgedrag.
- Verhoog de voorspellende waarde van persoonlijkheidstests door de persoonlijkheidsvragen in een context te plaatsen die zo goed mogelijk aansluit bij de situatie.

### Literatuur

- Ashton, M.C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior*, 19, 289-303.
- Ashton, M.C. & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11, 150-166.
- Barrick, M.R. & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.

- Barrick, M.R., Mount, M.K. & Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9-30.
- Bing, M.N., Whanger, J.C., Davison, H.K. & VanHook, J.B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology*, 89, 150-157.
- Borman, W.C. & Motowidlo, S. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, 10, 99-109.
- Butler, A.B. (2007). Job characteristics and college performance and attitudes: A model of work-school conflict and facilitation. *Journal of Applied Psychology*, 92, 500-510.
- Christiansen, N.D. & Robie, C. (2011). Further consideration of the use of narrow trait scales. *Canadian Journal of Behavioural Science*, 43, 183-194.
- Conard, M.A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, 40, 339-346.
- Cousins, S.D. (1989). Culture and self-perception in Japan and the United States. *Journal of Personality and Social Psychology*, 56, 124-131.
- Crédé, M., Bashshur, M. & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment*, 92, 390-399.
- Cronbach, L.J. & Gleser, C.C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana: University of Illinois Press.
- De Meijer, L.A.L., Born, M.Ph., 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, 219-251.
- De Vries, A. (2012). *Specificity in personality measurement*. Proefschrift: VU Universiteit Amsterdam.
- De Vries, A., De Vries, R.E. & Born, M.Ph. (2011). Broad versus narrow traits: Conscientiousness and Honesty-Humility as predictors of academic criteria. *European Journal of Personality*, 25, 336-348.
- De Vries, A., Van den Berg, R.H., Born, M.Ph. & De Vries, R.E. (2011). Consciëntieusheid en Integriteit als voorspellers van studieprestaties en contraproductief studiegedrag. *Tijdschrift voor Hoger Onderwijs*, 29, 20-30.
- De Vries, R.E., Ashton, M.C. & Lee, K. (2009). De zes belangrijkste persoonlijkheidsdimensies en de HEXACO Persoonlijkheidsvragenlijst. *Gedrag & Organisatie*, 22, 232-274.
- De Vries, R.E., De Vries, A., De Hoogh, A. & Feij, J. (2009). More than the Big Five: Egoism and the HEXACO model of personality. *European Journal of Personality*, 23, 635-654.
- De Vries, R.E., Zettler, I. & Hilbig, B.E. (2014). Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment*, 21, 286-299.
- Festinger, L. (1954). A theory of social comparisons. *Human Relations*, 7, 117-140.
- Furnham, A. & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and Individual Differences*, 37, 943-955.
- Goldberg, L.R. (1990). An alternative 'description of personality': The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Gruys, M.L. & Sackett, P.R. (2003). Investigating the dimensionality of counterproductive work behavior. *International Journal of Selection and Assessment*, 11, 30-42.
- Heine, S.J., Lehman, D.R., Peng, K. & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, 82, 903-918.
- Hurtz, G.M. & Donovan, J.J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85, 869-879.

- Johnson, M.K., Rowatt, W.C. & Petrini, L. (2011). A new trait on the market: Honesty-Humility as a unique predictor of job performance. *Personality and Individual Differences*, 50, 857-862.
- Lee, K., Ashton, M.C. & De Vries, R.E. (2005). Predicting workplace delinquency and integrity with the HEXACO and Five-Factor models of personality structure. *Human Performance*, 18, 179-197.
- Lee, K., Ashton, M.C. & Shin, K.H. (2005). Personality correlates of workplace anti-social behavior. *Applied Psychology: An International Review*, 54, 81-98.
- Lievens, F., De Corte, W. & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, 93, 268-279.
- Nicol, A.A.M. & Paunonen, S.V. (2002). Validity evidence for the different item styles of overt honesty measures. *Journal of Business and Psychology*, 16, 431-445.
- NOA. (2009). *Handleiding Multiculturele Persoonlijkheids Test – Big Six*. Amsterdam: NOA.
- O'Neill, T.A. & Allen, N.J. (2011). Personality and the prediction of team performance. *European Journal of Personality*, 25, 31-42.
- Ones, D.S. & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, 17, 609-626.
- Ones, D.S. & Viswesvaran, C. (2001). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment*, 9, 31-39.
- Ones, D.S., Viswesvaran, C. & Reiss, A.D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.
- Paunonen, S.V. & Ashton, M.C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality*, 35, 78-90.
- Paunonen, S.V., Haddock, G., Forsterling, F. & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European Journal of Personality*, 17, 413-433.
- Paunonen, S.V., Rothstein, M.G. & Jackson, D.N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior*, 20, 389-405.
- Peng, K., Nisbett, R.E. & Wong, N. (1997). Validity problems of cross-cultural value comparison and possible solutions. *Psychological Methods*, 2, 329-341.
- Ployhart, R.E. & Holtz, B.C. (2008). The diversity-validity dilemma: Strategies for reducing race/ethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology*, 61, 153-172.
- Poropat, A. E. (2009). A meta-analysis of the Five-Factor Model of personality and academic performance. *Psychological Bulletin*, 135, 322-338.
- Robie, C., Schmit, M.J., Ryan, A.M. & Zickar, M.J. (2000). Effects of item context specificity on the measurement equivalence of a personality inventory. *Organizational Research Methods*, 3, 348-365.
- Rotundo, M. & Sackett, P.R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy capturing approach. *Journal of Applied Psychology*, 87, 66-80.
- Sackett, P.R. & DeVore, C.J. (2001). Counterproductive behaviors at work. In N. Anderson, D.S. Ones, H. Sinangil & C. Viswesvaran (Eds.), *Handbook of Industrial, Work, and Organizational Psychology* (Vol. 1). London: Sage.
- Salgado, J.F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, 10, 117-125.

- 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.
- Schmit, M.J., Ryan, A.M., Stierwalt, S.L. & Powell, S.L. (1995). Frame-of-reference effects on personality scores and criterion-related validity. *Journal of Applied Psychology*, 80, 607-620.
- Spector, P.E. (2006). Method variance in organizational research: Truth or urban legend? *Organizational Research Methods*, 9, 221-232.
- Tett, R.P., Steele, J.R. & Beaugregard, R.S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior*, 24, 335-356.
- Van Leest, P. (1997). *Persoonlijkheidsmeting bij allochtonen [Personality measurement of ethnic minorities]*. Unpublished Doctoral Dissertation, Vrije Universiteit, Amsterdam.
- Wright, J.C. & Mischel, W. (1987). A conditional approach to dispositional constructs: the local predictability of social behavior. *Journal of Personality and Social Psychology*, 55, 454-469.

### **Personality as a predictor of job performance and counterproductive work behavior: the importance of specificity in personality measurement**

*Anita de Vries, Reinout E. de Vries, Marise Ph. Born & Remko H. van den Berg, Gedrag & Organisatie, volume 27, November 2014, nr. 4, pp. 407-427.*

The main purpose of this study was to improve the prediction of job performance and counterproductive work behavior using personality measurements. Three characteristics were examined: (1) trait specificity, i.e., using narrow rather than broad personality traits; (2) contextual specificity, i.e., the specificity of the situational context to which respondents refer when completing a personality inventory, such as the home or work context; and (3) referent specificity, i.e., the specificity of the comparison other in self-reported personality. The results show that narrow traits are able to explain more variance than broad traits in the prediction of work outcomes, and that adding a relevant context to personality items leads to higher predictive validities. Furthermore, the findings suggest that using reference groups influences the way people respond to a personality test. Taken together, this article accentuates the importance of specificity in personality measurement for improving the prediction of work outcomes.

*Key words:* personality, HEXACO, job performance, counterproductive work behavior, predictive validity