

Interview en arbeidsproeven bij allochtone sollicitanten

(4 maart 1999)

Auteurs: P. van der Maesen de Sombreff & P. Abell

Verschenen in: F. van de Vijver & N. Bleichrodt (Eds.), Selectie van allochtone kandidaten, Lisse: Swets & Zeitlinger, 2001.

dr. P.E.A.M. van der Maesen de Sombreff is manager Development van PiMedia in Maarssen.

drs. P. Abell is directeur van EGA Onderzoek en Advies voor personeelsmanagement in Amsterdam.

Inleiding

Dit hoofdstuk heeft betrekking op het selecteren van allochtone kandidaten met behulp van het interview en met arbeidsproeven. Bepaalde typen van interview en arbeidsproeven (waaronder assessment center-opdrachten) kan men zien als predictoren van het "sample"-type. Na een korte uitleg over het verschil tussen "sample"- en "sign"-benadering van predictie, beschrijven we welke factoren de interesse verklaren voor de "sample"-benadering bij het selecteren van allochtone kandidaten. Redenen om te denken dat sample-tests beter dan sign-tests geschikt zijn voor gebruik bij het selecteren van allochtone sollicitanten kunnen zijn dat de sample-tests allochtone sollicitanten minder discrimineren of dat allochtone sollicitanten betere kansen hebben met behulp van die tests toegelaten te worden (reductie van "adverse impact") bij gelijkblijvende predictieve validiteit.

Na een kort overzicht te hebben gegeven van onderzoeksresultaten over discriminatie en adverse impact van sign- en sample-tests volgt een paragraaf waarin wordt nagegaan onder welke condities tests van het sample-type een bijdrage kunnen geven aan het realiseren van twee doelen bij de selectie: verhogen van de predictieve validiteit en vergroten van de aanstellingskansen van allochtone kandidaten.

Vervolgens wordt het meest gebruikte selectie-instrument, het interview, behandeld.

Nadat geconstateerd is dat het interview een potentieel belangrijke rol speelt in de selectie van allochtone sollicitanten, omdat het immers op zeer grote schaal wordt gebruikt, zal worden aangetoond dat een interview met lage predictieve validiteit de oneerlijke beoordeling van allochtone sollicitanten zal versterken. Vervolgens komen ingrediënten van een valide interview aan de orde, en daarmee ook het effect dat verbeteringen in het interview zal hebben op de beoordeling van allochtone sollicitanten.

Veel aandacht wordt besteed aan het criteriumgerichte interview. Wat rest zijn verschillende vormen van 'onbedoelde' discriminatie. Er zal worden ingegaan op verschillende factoren die onbedoelde discriminatie kunnen veroorzaken. Het deel over interviews zal worden afgesloten met een aantal praktische aanwijzingen.

Sample- versus sign-tests

De sample- en de sign-methode kan men zien als de polen van benaderingen die dienen om gedrag van mensen te voorspellen (Roe, 1983). De sign-benadering (ook trait-benadering genoemd) gaat uit van het bestaan van een wetmatige relatie tussen de aanwezigheid van een trek (dispositie) enerzijds en (de frequentie of intensiteit van) het optreden gedrag anderzijds. Uit het bestaan van dat wetmatig verband en kennis van het niveau van een persoon op de trek kan men deduceren dat de persoon met een mate van waarschijnlijkheid het gedrag zal vertonen. De trek dient als een signaal van het gedrag op grond van de empirische relatie tussen die twee. Voorbeelden van tests van het sign-type zijn de intelligentietest, de persoonlijkheidsvragenlijst en diepte-interviews.

De "sample"-benadering gaat uit van generaliseerbaarheid van gedrag in situatie A (test) naar situatie B (functie). De test kan in die benadering worden opgevat als een steekproef uit

het geheel van taken die in de functie van belang zijn. Als de testsituatie en -omstandigheden maar goed genoeg lijken op de situatieve aspecten van taken in de functie, kan op grond van de wetmatigheid van de "behavioral consistency" (Wernimont & Campbell, 1968) voorspeld worden dat het gedrag dat de persoon laat zien in de test (gedragssteekproef) zal terugkomen in de functie (gedragsuniversum). Tests van het sample-type zijn assessment-opdrachten, arbeidsproeven en interviews waarin wordt gevraagd naar reeds getoond gedrag in situaties die ook voorkomen in de functie.

Validering in het kader van selectie is het aandragen van argumenten waarom uit gedragingen in de test gevolgtrekkingen te maken zijn over gedrag in de functie (NIP, 1988). Argumenten zijn gebaseerd op gegevens. Hoe meer ondersteunende gegevens aanwezig zijn en hoe gevarieerder de methoden om die gegevens te verzamelen, hoe beter de validiteit onderbouwd kan worden. De twee polaire benaderingen, sample en sign, hebben een voorkeur voor bepaalde soorten gegevens om hard te maken dat de test voldoende valide is. In de sign-aanpak wordt vooral beroep gedaan op onderzoek naar de empirische relatie tussen test- en criterium-uitslagen. Meestal worden die uitslagen gelijktijdig verzameld, soms ook volgtijdelijk. In het laatste geval is sprake van predictie in de ware betekenis van het woord. Ook wordt bij tests van het sign-type moeite gedaan om aan te tonen dat het gemeten (unidimensionale) begrip is wat het pretendeert te zijn en niet is wat het ook niet pretendeert te zijn; de relaties van het begrip met andere begrippen uit het conceptuele netwerk wordt onderzocht. Dit is ondersteuning van de begripsvaliditeit.

Argumenten voor de validiteit van sample-tests worden vaak ontleend aan oordelen van deskundigen. De methoden om die oordelen te verzamelen en de argumenten die men aan de resultaten van dat onderzoek ontleent noemt men inhoudsvalidering. Die deskundigen kunnen bijvoorbeeld aangeven dat bepaalde situaties zeer belangrijk zijn voor succes in de functie omdat personen in die situaties "de mist kunnen ingaan" (kritieke incidenten). Bovendien moeten deskundigen het met elkaar eens zijn over de mate van belangrijkheid van die kritieke situaties en over de mate van representativiteit van de verzamelde situaties. Een volgende stap is dat de deskundigen een oordeel geven van de mate waarin de voor de kritieke situatie gemaakte test (simulatie) qua belangrijke aspecten zoals de moeilijkheidsgraad lijkt op de situatie in de functie zelf. Ook hier is weer de consensus tussen experts van belang. Tenslotte moeten experts ook nog met elkaar overeenstemmen in hun waardering van antwoorden op de situaties.

Uit het feit dat de testconstructie inhoudsgeoriënteerd tot stand komt volgt nog niet dat de test valide is. Daarvoor is het nog nodig te laten zien wat de correlatie is tussen test en criterium. Toch is wel beweed (bijvoorbeeld door Guion, 1978) dat het gebruik van inhoudsvaliderende methoden bij de testconstructie alleen al voldoende steun geeft aan het gebruik van de test als selectiemiddel en is door rechtbanken in de VS die bewering ook aanvaard. Hierover wordt heden ten dage toch genuanceerder gedacht (zie bijvoorbeeld Hoekstra, 1995).

Sample-tests bij de selectie van allochtone sollicitanten

De interesse voor het gebruik van sample-tests bij de selectie dateert halverwege de 60-er jaren. Kravitz e.a. (1996) geven een overzicht van die ontwikkelingen. De Civil Rights Act van 1964 moest ervoor zorgen dat alle individuen gelijke kansen en behandeling kregen, zonder onderscheid naar ras, kleur, religie, geslacht of nationale afkomst. In de jaren daarna zijn er instituten in het leven geroepen om te zorgen voor naleving van de Act op het gebied van tewerkstelling, en zijn wetten en instructies uitgevaardigd, waaronder de Guidelines on Employee Selection Procedures (1966, 1978). President Johnson verkondigde een Executive Order die bedrijven die aan de overheid diensten en producten leverden verplichtte zich te gedragen ("forced compliance") conform de regels van een "affirmative action"-plan. Waar die bedrijven eerst nog een plan hoefden over te leggen om de "equal opportunity" van kandida-

ten te waarborgen, werden zij in 1971 aangesproken op de behaalde resultaten. In enkele geruchtmakende rechtszaken werden bedrijven verplicht de validiteit aan te tonen van selectie- of personeelsbeoordelingsprocedures die in het nadeel van individuen uit minderheidsgroepen hadden gewerkt. Niet alleen rechtszaken die door individuen waren aangespannen leidden tot de plicht van organisaties om de validiteit van selectiemethoden te bewijzen. Organisaties met selectieprocedures die een schending inhielden van de “80%-regel” (de selectieratio van minderheidsgroepen moet minstens 80% zijn van die van de groep met de hoogste selectieratio) konden erop rekenen aangeklaagd te worden op grond van onwettige discriminatie.

Deze overheidsmaatregelen hebben ervoor gezorgd dat tests bij werkgevers en (dus) selectiepsychologen vanwege hun mogelijk discriminatoir karakter sterk in de belangstelling kwamen te staan. De volgende factoren hebben daarbij de sample-tests de wind in de rug gegeven.

In de eerste plaats is het aantonen van predictieve validiteit van welke test dan ook geen eenvoudige zaak. Met een sign-test was men helemaal op dit soort onderzoek aangewezen. Vaak ontbreekt een steekproef van voldoende omvang (om betrouwbare uitspraken te doen) en als die grote steekproef er wel is, is het uitvoeren van het onderzoek erg kostbaar. Werkgevers willen er meestal niet in investeren.

In de tweede plaats werd het niet getolereerd dat werkgevers beroep deden op validiteitsgegevens van een test die elders verzameld waren, tenzij op grond van een functieanalyse kon worden aangetoond dat de functie waarvoor werd geselecteerd vergelijkbaar was met de functie waarvoor de test valide was gebleken. Er heerste een sterk geloof dat de validiteit van een test situatiespecifiek is. De redenering was dat de validiteit van de test in situatie X heel anders kon uitpakken dan in situatie Y. Onderzoeksresultaten gaven daartoe ook alle aanleiding: eenzelfde test kon een positieve validiteit hebben in de ene, en een negatieve validiteit in een andere functie. Gemiddeld was de validiteit laag. De heersende mening was met andere woorden dat, alvorens een test voor een functie als selectie-instrument te mogen gebruiken, eerst in die groep validiteitsonderzoek met positief resultaat moest zijn verricht.

Een derde oorzaak van de aandacht voor sample-tests was dat ook onder psychologen de geldigheid van de sign-aanpak in twijfel werd getrokken. Van invloed waren publicaties van Mischel (1968) die liet zien dat de coherentie van gedragingen die horen tot eenzelfde trek in verschillende situaties in empirie veel lager is dan in theorie. In de arbeidspsychologie had McClelland (de vader van de competentie-gedachte) invloed met zijn artikel van 1973 waarin hij bepleitte om per functie specifieke vaardigheden te meten in plaats van intelligentie. Met behulp van kritieke incidenten zou zijn op te sporen welke vaardigheden tot succes leidden. Bovendien, waarom zou men ook moeilijk doen met de sign-benadering? Communicatie tussen opdrachtgever en psycholoog met behulp van trekken geeft allerlei misverstanden. Neem de eigenschap ‘flexibiliteit’; welke diversiteit van gedragingen kan men daaronder niet verstaan? Vergeleken daarmee was de sample-aanpak simpel: de psycholoog voorspelt immers zonder omweg van allerlei constructen gedrag in de functie.

Een vierde reden om kritiek te leveren op de sign-tests was dat enkele aspecten waarmee dat type tests behept was maatschappelijk moeilijk lagen. Minderheden scoorden (en scoren) op intelligentietests gemiddeld een standaarddeviatie lager dan blanke kandidaten, hetgeen bij toepassing van eenzelfde selectiestrategie voor beide groepen leidt tot “adverse impact” van de minderheidsgroep. Persoonlijkheidsvragenlijsten hebben te lijden van de sociale wenselijkheidstendens bij kandidaten en zouden teveel indringen in de persoonlijke levenssfeer. Het was moeilijk te accepteren dat opdrachten en vragen die inhoudelijk niet veel te maken leken te hebben met taken uit de functie gebruikt werden om te beslissen of mensen al dan niet geschikt waren voor die functie.

Het testen op signs stond onder vuur. In verband met de selectie van minderheden waren die tests bij voorbaat verdacht. Aan het gebruik van sample-tests, zoals het interview, werden veel minder eisen gesteld; het volstond om experts te laten zeggen dat de test functiegerelateerd was. Dat stelsel van factoren is verantwoordelijk geweest voor de aandacht voor sample-benadering. Het assessment center werd big business. Aspirant-brandweermannen moesten een brand blussen in plaats van een testje maken. Gedurende de jaren 70 nam het gebruik van sign-tests in de VS af (Hale, 1982).

Nieuwe ontwikkelingen m.b.t. sign-tests

Vanaf het eind van de jaren '70 pasten Schmidt, Hunter en medewerkers de methode van meta-analyse toe op de validiteit van selectie-instrumenten. De methode en de resultaten zijn elders in dit boek (hoofdstuk ..) uitvoerig besproken. Daarom volstaan wij hier met de belangrijkste conclusies. In tegenstelling met de heersende doctrine dat de validiteit van tests situatiespecifiek was, is de boodschap van Schmidt, Hunter c.s. dat validiteit van de cognitieve capaciteitentest generaliseerbaar is over vele soorten functies. Een zeer gedetailleerde functieanalyse was volgens de onderzoekers niet nodig om te besluiten dat functies al dan niet lijken op elkaar. De methode van validiteitsgeneralisatie zoals meta-analyse van testvaliditeit werd genoemd, is ook toegepast op allerlei andere tests en predictoren, zoals interview, persoonlijkheidsvragenlijsten, AC, beoordeling van ervaring, proefperiode, opleiding en leeftijd. Op de kritiek van Mischel op de voorspellende waarde van persoonlijkheidstrekken is door andere auteurs gereageerd (bijvoorbeeld Rushton, Brainerd & Preston, 1983; Epstein & O'Brien, 1985). De voorspellende waarde van een trek zit hem in het principe van aggregatie van een groot aantal gedragingen in veel situaties. Mischel werd verweten met één gedraging in één situatie één andere gedraging in een andere situatie te willen voorspellen. Dit is vanwege de onbetrouwbaarheid onmogelijk; van een test met één item is redelijkerwijs geen sterke voorspellende waarde te verwachten.

De validiteit van een goed geconstrueerde intelligentietest (namelijk één die een betrouwbare indicatie geeft van g) is een van de beste, zo werd aangetoond in een aantal grote studies (McHenry e.a., 1991; Ree e.a., 1991, 1994; Linn & Hastings, 1984) bij resp. US Army en Airforce en Law Schools) en in meta-analyses (Hunter & Hunter, 1984). De gecorrigeerde validiteit wordt op .45 à .50 geschat. In het algemeen is de voorspellende waarde van de intelligentietest groter voor trainingsresultaten dan voor succes in functies. Het laatste wordt bijna steeds geoperationaliseerd door oordelen van leidinggevend. Zie ook Van der Maesen de Sombreff (1997a).

Bovendien blijkt differentiële validiteit van de intelligentietest voor verschillende etnische groepen niet aantoonbaar aanwezig (Schmidt, Hunter; Zeidner, 1987; Schmitt, 1987; Te Nijenhuis, 1997). Intercept en steilheid van de lijn die de regressie van criterium op predictor weergeeft zijn voor etnische groepen identiek. Tests discrimineren allochtone sollicitanten niet in de zin dat de test niet een onderpredictie geeft van het succesniveau dat allochtone sollicitanten later in de functie behalen. In die zin is de intelligentietest 'eerlijk' (fair). Voor de productiviteit van het werknemersbestand is de beste selectiestrategie die organisaties kunnen toepassen het rangordenen van kandidaten op grond van hun testscores en het 'afromen' van de beste kandidaten zonder aanzien van etnische afkomst totdat de vacatures zijn vervuld.

Er is echter een tweede vorm van eerlijkheid van selectieprocedures in het spel. Deze heeft betrekking op de relatieve welvaart en welzijn van bevolkingsgroepen in hetzelfde land. Allochtone sollicitanten hebben aantoonbaar een slechtere positie dan autochtonen, ook op de arbeidsmarkt. Door een intelligentietest te gebruiken op de wijze die de meeste utiliteit geeft worden allochtone sollicitanten het meest benadeeld. Allochtone sollicitanten raken in

een vicieuze cirkel van achterstelling, wat kan leiden tot grote maatschappelijke kosten. Als een selectietestprocedure allochtone sollicitanten benadeelt kan dat liggen aan twee factoren: geringe validiteit en een groot verschil tussen gemiddelde scores op de predictor ten nadele van allochtone groep. De intelligentietest is weliswaar de meeste valide test, maar het verschil tussen de gemiddelde score van allochtone sollicitanten en autochtonen is één standaarddeviatie (Hunter & Hunter, 1984; Te Nijenhuis, 1997). Zelfs bij een selectieratio van .90 betekent dit dat de 80% regel (zie boven) niet wordt gehaald. Bij een SR van .10 is het 'adverse impact'-percentage zelfs 14%. Meerderheids- en minderheidsgroep verschillen meer op de test dan op functiesucces. Een meta-analyse van Ford e.a. (1986) wijst uit dat de gemiddelde beoordeling is voor niet-blanken .50 SD lager is dan voor blanken.

Dat allochtone sollicitanten gemiddeld lager scoren op een intelligentietest dan autochtonen mag niet geïnterpreteerd worden in de zin dat de eerste groep minder intelligent is dan de tweede. Hoe minder ervaring allochtone sollicitanten hebben kunnen opdoen met de Nederlandse cultuur -wat ten dele bepaald wordt door de verblijfsduur in Nederland- hoe meer de test voor allochtone sollicitanten een cultuurtest is in plaats van een intelligentietest (Drenth, 1988; Bissessur e.a., 1989; Van der Maesen, 1997b). Cultuur wordt bovendien opgedaan op school. Allochtone sollicitanten hebben minder scholing dan autochtonen (Van der Meer & Wielers, 1997); velen komen reeds op achterstand in hun vroege jeugd (SCP, 1996).

Validiteit en adverse impact van alternatieve tests, in het bijzonder arbeidsproeven

Uit het vorige komt als vraagstuk voor selectiepsychologen naar voren hoe een alternatieve selectiemethode kan worden ontworpen die voldoet aan twee ontwerpeisen: die (1) qua validiteit vergelijkbaar is met de intelligentietest maar waardoor (2) de 'adverse impact' kan worden verkleind.

Eén van die middelen zou de arbeidsproef kunnen zijn. Meta-analytisch onderzoek naar de validiteit van work samples is bij ons weten in geringe mate verricht, behalve naar psychomotorische arbeidsproeven en naar assessment center-oefeningen voor hogere staf- en managementfuncties. In het algemeen is de validiteit van psychomotorische arbeidsproeven hoog (Hunter & Hunter, 1984). Het overzicht van Schmidt, Ones en Hunter (1992) van validiteitsstudies naar assessment centers laat consistent lagere validiteiten zien voor deze methode dan voor de intelligentietest. Veelbelovend lijken arbeidsproeven die trainbaarheid meten. Kandidaten krijgen een nieuwe vaardigheid aangeleerd en moeten die daarna toepassen. Robertson & Downs (1989) deden een meta-analyse naar de validiteit van dit soort tests en vonden een ongecorrigeerde validiteit van .48. Van der Maesen ontwierp een instructie- en oefenprogramma om met een tekstverwerker te leren omgaan en een aantal oefeningen om het geleerde in praktijk te brengen (Van der Maesen & Westen, 1991). De gelijktijdige validiteit van alleen al de tijd om de instructies en oefeningen door te nemen, met als criterium de som van beoordelingen door leidinggevenden, was -.46.

Robertson & Kandola (1982) verzamelden negen studies naar de 'adverse impact' van arbeidsproeven; minderheidsgroepen bleken nauwelijks in gemiddelde score achter te blijven op meerderheidsgroepen. Pukalos & Schmitt (1996) vonden op een *situationele beoordelingstest* een verschil van .41 SD ten voordele van blanke kandidaten in de VS. Schmitt en anderen (1997) bekeken de 'adverse impact' van nog andere instrumenten. De cijfers (effect-groottes: verschil tussen gemiddelden in standaarddeviaties) waren: gestructureerd interview (.09), persoonlijkheidsschaal gewetensvolheid (.09), biodata (.20).

Naspeuringen naar In Nederland uitgevoerd onderzoek naar de differentiële validiteit van arbeidsproeven of assessment-oefeningen hebben niets opgeleverd. Wèl is het zo dat

psychologische adviesbureaus in Nederland hebben geëxperimenteerd met specifieke selectieprocedures voor allochtone kandidaten. In die procedures werden o.a. gebruikt: individueel afgenomen intelligentietests, persoonlijkheidsvragenlijsten voor kinderen, taaltests, interviews naar opleiding en ervaring die werd opgedaan in het land van herkomst, interviews door allochtone interviewers, assessment center-oefeningen. In eerste instantie werd het succes van deze methoden afgelezen aan de hand van de mogelijkheid om de adverse impact te reduceren. Het is niet verwonderlijk dat de validiteit van die methoden niet onderzocht is gegeven de noodzaak, voor het bereiken van betrouwbare conclusies, om in dit soort onderzoek data te verzamelen bij een grote steekproef van allochtone sollicitanten in combinatie met het vaak kortstondig, experimenteel karakter van dit soort methoden.

Over adverse impact van alternatieve methoden in ons land is bij ons weten ook weinig bekend. In onderzoek van Van Leest (1996) en Te Nijenhuis (1997) komt naar voren dat allochtone sollicitanten op veel wenselijk geachte persoonlijkheidsschalen ongunstiger scoren dan autochtonen. Onderzoek naar de LSCP Test voor Sociale Beslissingen (LTSB), een schriftelijke test met elementen van een arbeidsproef voor het meten van moreel inzicht van aspirant politie-agenten (Van der Maesen, 1995; Van Zanten, Van der Maesen & Posthumus, 1996) wees uit dat allochtone sollicitanten niet lager scoren op die test (Klinkenberg, 1997).

We besluiten dit onderdeel met het aangeven van een ontwikkeling op selectiegebied die de instroom van allochtone sollicitanten in sommige functies zou kunnen bevorderen zonder de eis van goede validiteit te verontachtzamen.

Een veelbelovende methode

Zoals reeds eerder is beschreven hebben allochtone kandidaten het moeilijk met de meest valide predictor van succes in de functie, de intelligentietest. Gebrek aan ervaring van allochtone burgers met de autochtone cultuur kan de oorzaak zijn van die achterstand. Voor een deel kan dat gebrek aan ervaring tot stand komen door moeilijk te overbruggen verschillen tussen de autochtone en allochtone culturen. Greenfield (1997) bijvoorbeeld maakt duidelijk dat IQ-tests gebaseerd zijn op sociale conventies die van de cultuur deel uitmaken. Zo zit 'achter' de IQ-test de waardering voor individuele prestaties en competitie; in allochtone culturen is kennis eerder het product van een groepsinspanning. De wijze van communicatie in allochtone culturen is vaker mondeling dan schriftelijk. Binnen de schriftelijke manier van bevragen is het multiple choice-format ongewoon in allochtone culturen.

Het aspect van de schriftelijke manier van informatie-uitwisseling in gangbare tests willen we wat verder uitwerken. Voor een deel kan de lagere score van allochtone sollicitanten op tests worden toegeschreven aan de geringere leesvaardigheid van allochtone sollicitanten. Recent Amerikaans onderzoek van Chan en Schmitt (1997) wijst uit dat de score-achterstand van allochtone sollicitanten op een test in sterke mate gereduceerd wordt, wanneer men de leesvaardigheidscomponent van de test verwijdt en men daarvoor in de plaats beroep doet op de vaardigheid om gesproken taal te begrijpen.

Het is dan de vraag of de verwijdering of reductie van gedrukte informatie ten voordele van mondelinge informatie in de test de validiteit daarvan zou verminderen. Zolang die vraag niet door criterium-gerelateerd validiteitsonderzoek is beantwoord, spelen argumenten een rol die betrekking hebben op de inhoudsvaliditeit van de methode. De vraag is hier of er in functies waarin met name veel mensen van allochtone afkomst instromen eerder mondeling dan schriftelijk wordt gecommuniceerd.

Kijken we naar arbeidsmarktgegevens (bijvoorbeeld Sociaal en Cultureel Planbureau, 1996),

dan valt op dat een belangrijk gedeelte van de allochtone sollicitanten een plaats vindt in de dienstverlening (horeca, detailhandel, groothandel, politie(surveillance), schoonmaakbedrijf, personenvervoer, beveiliging) of (minder dan voorheen) in de industriële productie waar in toenemende mate in teamverband wordt gewerkt. In beide sectoren zijn het ontvangen, verwerken en verzenden van informatie in rechtstreekse mondelinge interacties met klanten en collega's de kritieke taakkenmerken. Dit neemt natuurlijk niet weg dat er ook veel functies zijn waarin allochtone werknemers moeten beschikken over een goed niveau van lees- en schrijfvaardigheid. Een onderzoek van Van Leest en Dutman (1994) bijvoorbeeld bracht aan het licht dat een simpele leesvaardigheidstest een krachtige voorspeller was van succes van allochtone kandidaten in de politieopleiding.

De videotecnologie heeft nieuwe mogelijkheden gebracht. In een videotest zijn opdrachten op te nemen die bestaan uit korte scènes van sociale interacties. Videotests voor selectie-doelinden zijn in Nederland ontwikkeld en toegepast voor het meten van effectiviteit van leidinggeven, van effectiviteit in de dagelijkse communicatie (empathie, assertiviteit), van commerciële klantcontacten en van zakelijk overleg met collega's (Van der Maesen & Van Noort, 1989; Meltzer, 1995; Moonen, 1995). In deze gevallen ging het om open-antwoord-tests. In deze testvorm is het de bedoeling dat, na afloop van een scène, de sollicitant zich verplaatst in de rol van collega, chef of adviseur en spontaan een mondeling antwoord geeft op het probleem waarvan hij zojuist deelgenoot was.

Chan en Schmitt deden onderzoek met een versie van de videotest waarin na de scène, eveneens op video, een aantal antwoordalternatieven wordt vertoond. Voor de videoversie werd een parallelversie in schriftelijke vorm ontwikkeld. Het design was 2 X 2: afkomst (blank-niet blank) X methode (schriftelijk-video). De auteurs konden aantonen dat de twee versies dezelfde dimensies meten. De door de auteurs ontwikkelde videotest heeft niet de *high fidelity* van een assessment center-oefening en ook niet van een open-antwoord videotest, maar heeft duidelijk een groter realiteitsgehalte dan de schriftelijke beoordelingstest en de intelligentietest.

De kandidaat moet elk van de antwoorden beoordelen op een 6-puntsschaal van effectiviteit. De antwoorden zijn indicatief voor vier dimensies: werkbetrokkenheid, kwaliteit van het werk, conflictbeheersing en empathie. Scoring van de antwoorden van kandidaten werd gedaan door ze te vergelijken met de antwoorden van experts. Opvallende resultaten waren dat:

- de adverse impact van zwarte (ten opzichte van blanke) kandidaten veel kleiner was als gevolg van de videoversie dan als gevolg van de schriftelijke versie (-.28 ten opzichte van -.1.19). De donkere groep profiteerde met andere woorden in veel grotere mate dan de blanke groep van het verschil in methode (video versus schriftelijk);
- de leesvaardigheidscomponent van de schriftelijke versie voor een groot deel dit verschil in adverse impact bepaalde;
- zwarte kandidaten de videoversie veel acceptabeler (face valide) vinden dan de schriftelijke versie. Bij donkere kandidaten is het verschil in acceptatie .84 SD, bij blanke kandidaten .13 SD.

Het onderzoek van Chan en Schmitt is uniek omdat het toelaat te concluderen dat het effect (verkleining van adverse impact van allochtone sollicitanten) alleen toeschrijfbaar is aan de methode (video in plaats van schriftelijke test) en niet aan de inhoud van de test.

De vraag is interessant welke processen maken dat allochtone sollicitanten een grote scorewinst behalen. Natuurlijk heeft het feit dat er door de videotest geen beroep wordt gedaan op de leesvaardigheid ermee te maken. Aan de andere kant is begrip van taal ook in de videotest nodig, met het verschil dat die hier gesproken wordt. Het is mogelijk dat de context waarin de boodschap wordt overgebracht het begrip van de boodschap begunstigt. De context levert niet alleen de gesproken woorden en de syntactische structuur

waaruit de betekenis van de boodschap moet worden afgeleid, maar ook de paralinguale informatie (toonhoogte, spreek-snelheid, pauzes) en de visuele informatie (nonverbale aspecten zoals gelaatsuitdrukkingen, informatie over de werkomgeving). Eerder dan dat die extra cues uit andere kanalen leiden tot *overload* door informatie (en verminderde prestatie) zou het kunnen dat die cues gecombineerd de kandidaat beter in de bedoelde interpretatierichting wijzen.

De predictieve validiteit van de nieuwe test werd in dit onderzoek niet bestudeerd. Wel is een andere belangrijke parameter van de gebruikswaarde van een test in het onderzoek meegenomen, namelijk de mate waarin kandidaten de methode face-valide vinden. Er is de laatste tijd een groeiende belangstelling voor onderzoek naar percepties van face validiteit en eerlijkheid van selectieprocedures en -instrumenten en naar de relatie tussen de tevredenheid van kandidaten met de selectieprocedure enerzijds en de testmotivatie en -prestaties van kandidaten en de validiteit van die prestaties anderzijds (Chan, Schmitt, DeShon, Clause, Delbridge, 1997; Ployhart & Ryan, 1998). Er is enige evidentie uit onderzoek dat een lage tevredenheid met een test -en bij allochtone sollicitanten betreft die vooral de schriftelijke test- de motivatie om die test te maken verlaagt, waardoor die test een minder accuraat beeld geeft van de prestaties van die kandidaten. Geringe motivatie zou dus invloed hebben op de validiteit (die wordt lager) en op de adverse impact (die wordt groter). Die geringe motivatie kan te maken hebben met de indruk van de face validiteit van de test maar natuurlijk ook met een geringe collectieve *self-efficacy*. Allochtone kandidaten weten hoe moeilijk andere mensen van allochtone afkomst het hebben om de horde van een op schrift afgenomen test te passeren.

De uitslag van Chan en Schmitt is veelbelovend. Een onderzoek met een videotest van twee andere auteurs (Weekley & Jones, 1997) vertelt echter een ander verhaal. Videotests werden ontwikkeld, afgenomen en gevalideerd in twee functies, medewerkers van een warenhuis en van een verzorgingstehuis. De adverse impact in het nadeel van niet-blanke personen van de videotest was ongeveer even groot als die van een cognitieve capaciteitentest (ongeveer .50 SD). De effectgrootte voor prestatiebeoordelingen in de functie was kleiner (.15 SD in de groep van medewerkers van het verzorgingstehuis). De validiteit van de videotest (zowel van een empirisch gesleutelde versie als van een versie waar de scoring van de antwoorden was gebaseerd op oordelen van klanten) was redelijk hoog (in de buurt van de .30). Het is moeilijk te gissen waaraan het verschil in uitkomst tussen de studies van Chan & Schmitt en Weekley & Jones te wijten is. Wel is het opvallend dat in de videotest van Weekley & Jones de antwoordalternatieven niet geacteerd worden maar worden gedrukt op het scherm en voorgelezen. Bovendien moet niet elk alternatief afzonderlijk worden beoordeeld, maar moeten de kandidaten aangeven welk antwoord zij uit vier alternatieven zouden kiezen.

Samenvattend oordeel over arbeidsproeven bij de selectie van allochtone sollicitanten

Er is weinig informatie voorhanden over de differentiële validiteit van arbeidsproeven, niet in de VS en helemaal niet in Nederland. Er zijn wel enkele gunstige berichten over de verminderde adverse impact van arbeidsproeven. Mogelijkerwijs speelt de verbeterde acceptatie van arbeidsproeven door allochtone kandidaten door toedoen van verbeterde testmotivatie een rol in het verhogen van de validiteit (en het verlagen van de differentiële validiteit) van dat soort tests. Er is een opleving van onderzoek naar de invloed van kandidaattevredenheid op predictieve validiteit, en het is interessant om te zien wat dat onderzoek zal opleveren. Voor functies waarin allochtone werknemers verantwoordelijk zijn voor de mondelinge interacties van de organisatie met de buitenwereld (*boundary spanners*) of waarin zij met collega's overleggen om teamprestaties te leveren, kan een videotest van het type Chan & Schmitt en Weekley & Jones een nuttig en acceptabel medium zijn. Een afdoend antwoord op de vraag naar de gebruikswaarde van videotests bij allochtone sollicitanten is echter nog niet gegeven. Verdere resultaten van onderzoek naar de predictieve validiteit en de adverse impact van dit

soort tests zullen ongetwijfeld binnenkort door Amerikaanse psychologen worden gepubliceerd.

Het laatste deel van dit hoofdstuk zal besteed worden aan het meestgebruikte selectiemiddel: het interview, en de waarde ervan voor de selectie van allochtone sollicitanten.

Het gebruik van het interview bij de selectie van allochtone sollicitanten

De allochtone beroepsbevolking is (nog steeds) oververtegenwoordigd in de lagere functieniveaus. Van de vier te onderscheiden beroepsniveaus - elementair, lager, middelbaar en hoger/wetenschappelijk - is het merendeel van de Turken en Marokkanen werkzaam in de beroepen op elementair en lager niveau; ook het percentage Surinamers en Antillianen in de beroepen op elementair en lager niveau is nog 10% hoger dan het percentage van de autochtone beroepsbevolking in vergelijkbare beroepen (Tesser e.a. 1998). De selectie van allochtone sollicitanten zal zich voor een relatief groot deel dus afspelen in de elementaire en lagere functies. Het meest gebruikte selectiemiddel voor deze functies is het interview: andere selectiemiddelen als het psychologische onderzoek en het assessment center worden, in tegenstelling tot de hogere functieniveaus, bij de selectie voor de lagere en elementaire functies veel minder gebruikt. Eventuele oneerlijke beoordeling van allochtone sollicitanten zal zich dus het meest doen gelden bij gebruik van het interview. De vraag is dan, of en in hoeverre het interview kan leiden tot een oneerlijke beoordeling van allochtone sollicitanten.

Een aantal onderzoekers is van mening dat er tot nog toe geen duidelijke aanwijzingen zijn dat het interview allochtone sollicitanten zou benadelen. Zo stellen Campion en Arvey (1989) dat '...there is insufficient evidence to draw firm conclusions regarding racial bias in interviewer judgement'. Uit de resultaten van een omvangrijk onderzoekprogramma van het ILO (International labour Office, zie ook Bihning en Beijl, 1995), dat begin jaren negentig is gestart en nog steeds loopt, blijkt echter dat er wel degelijk sprake is van oneerlijke beoordeling van allochtone sollicitanten bij gebruik van het selectieinterview. In dit onderzoeksprogramma, dat inmiddels al in een groot aantal landen is uitgevoerd (waaronder in Nederland, zie Gras e.a. 1996), wordt gebruik gemaakt van zogenaamde 'situatietests', waarbij twee gelijkwaardige kandidaten, één van allochtone en één van autochtone afkomst, reageren op één en dezelfde advertentie voor een vacature. De kandidaten zijn gematched op een aantal belangrijke criteria voor de betreffende vacature, zoals kwalificaties, werkervaring en leeftijd. Uit de situatietests blijkt dat de allochtone kandidaat vaak niet verder komt dan een eerste telefoongesprek. Als het al tot een interview komt, blijkt de beslissing vaker negatief uit te vallen voor de allochtone kandidaat dan voor de autochtone kandidaat. In de Verenigde Staten bijvoorbeeld bleek dat witte kandidaten (whites) aan het einde van het interview vier keer zo vaak een baan werd aangeboden als zwarte kandidaten (blacks), terwijl de kwalificaties van de kandidaten gelijk waren (Bendick, 1996). Soortgelijke situatietests in andere landen laten vergelijkbare resultaten zien (vgl. Goldberg e.a. 1996, Colectivo o.e., 1995 en Arriijn e.a. 1998). De discriminatie blijkt overigens altijd impliciet te zijn (zie ook Essed, 1991).

Indien het interview als selectiemiddel wordt ingezet, betreft het doorgaans een traditioneel interview. Daarmee wordt bedoeld op een nauwelijks gestructureerde vorm van interviews, waarbij aspecten als gespreksonderwerpen, te stellen vragen en de wijze van beoordeling niet vast liggen. Meestal is niet duidelijk aan welke eisen een sollicitant moet voldoen, of zijn de functie-eisen zodanig ruim en daardoor multi-interpreteerbaar omschreven (vgl. conceptuele en interpersoonlijke effectiviteit), dat elke selecteur er in principe een andere invulling aan kan geven. Ook is het interview vaak in handen van ongeoefende gelegenheidsselecteurs, die ieder op hun eigen wijze de geschiktheid van sollicitanten

trachten vast te stellen. Dat de predictieve validiteit van dergelijke selectie interviews daardoor gering is (zie ook van der Maesen de Sombreff, 1992) en het traditionele interview een instrument is dat relatief weinig garantie biedt voor een eerlijke beoordeling van welke categorie sollicitanten dan ook, zal duidelijk zijn. Daarnaast krijgen - door onder andere gebrek aan structuur en voorschriften ten aanzien van de beoordeling - persoonlijke voorkeuren, en daarmee ook eventuele vooroordelen, alle ruimte, met benadeling van allochtone sollicitanten als gevolg.

Het traditionele interview: mogelijke oorzaken van een oneerlijke beoordeling

In het hierna volgende wordt wat uitgebreider ingegaan op enkele specifieke aspecten van het traditionele, ongestructureerde interview die een oneerlijke beoordeling van allochtone sollicitanten tot gevolg kunnen hebben.

Een van deze aspecten betreft de functie eisen. Allochtone werkzoekenden zullen - méér dan autochtone werkzoekenden - door het gemiddeld lagere opleidingsniveau aangewezen zijn op de elementaire en lagere functieniveaus. Al geruime tijd is er sprake van een verdringing van laag- door hoger geschoolden die vooral in de categorie 'elementaire beroepen' ernstige vormen blijkt aan te nemen. Zo nam in 1995 de werkgelegenheid in deze categorie toe met 40.000 banen, waarbij meer dan 50 procent van deze banen echter vervuld bleek te zijn door personen met een HAVO/VWO opleiding, terwijl het vereiste opleidingsniveau niet meer dan maximaal basisonderwijs bedroeg (Minderhedenbeleid 1997,1996). In een traditioneel interview ligt meestal niet goed vast hoe bepaald kan worden of aan de functie-eisen is voldaan. Het kiezen van de 'beste' sollicitant zal er toe leiden dat minder gekwalificeerde, maar overigens wel geschikte sollicitanten worden afgewezen. Een oneerlijke beoordeling van - gemiddeld lagere opgeleide - allochtone sollicitanten kan het gevolg zijn.

Een ander punt betreft het gebruik van formele en marginale criteria. De formele criteria hebben betrekking op de wijze waarop de functie wordt vervuld (bijvoorbeeld de snelheid waarmee gewerkt moet worden), de marginale criteria corresponderen met algemene vereisten en hebben niet direct betrekking op de functie (Roe, 1983). De formele criteria worden ook wel technisch- instrumentele criteria genoemd, de marginale criteria staan bekend als sociaal-normatieve criteria. Sociaal-normatieve criteria, zoals passen in de bedrijfscultuur, motivatie, creativiteit, flexibiliteit en dergelijke blijken in de loop der jaren meer aan belang te winnen. Uit onderzoek blijkt dat in 35 jaar het percentage vacatures, waarin sociaal-normatieve functie-eisen worden gesteld verviervoudigd was. Waren in 1955 functie-eisen als kwaliteitsbewustzijn, zelfstandigheid, leiding geven en intelligentie nog de enige sociaal-normatieve functie-eisen, in 1990 werden er al 19 geteld, waaronder functie-eisen als ondernemingsgeest, creativiteit, ambitie en aanpassingsvermogen. (Moelker,1992). Of sollicitanten voldoen aan sociaal-normatieve functie-eisen is moeilijk objectief vast te stellen. De (subjectieve) mening van de selecteur is meestal doorslaggevend, waardoor er alle ruimte is voor persoonlijke interpretaties, en eventuele vooroordelen vrij spel kunnen krijgen.

Ook voorkeuren van selecteurs kunnen in het traditionele interview doorslaggevend zijn. Een onderzoek naar voorkeuren van werkgevers voor sollicitanten naar functies op een elementair (lower skilled) niveau laat zien dat de factoren leeftijd, sekse, etnische afkomst en gezondheid 70 % van de voorkeur van werkgevers bepalen. Met andere woorden: de voorkeur van werkgevers voor sollicitanten gaat uit naar jonge, gezonde, autochtone mannen (Van Beek 1993). Van Beek concludeert dat '...if selection criteria used by employers do not change, the low employment rate on the Dutch labour market for older workers, ethnic minorities, women and workers with a slightly increased probability of sick leave, may prove to be lasting'. Dus, als de voorkeur van selecteurs voor bepaalde categorieën sollicitanten de vrije hand wordt gelaten - zoals in het traditionele interview - is de kans op een oneerlijke beoordeling van allochtone sollicitanten duidelijk aanwezig.

Selecteurs blijken zich eerder te laten leiden door negatieve dan door positieve informatie (Rowe, 1989). Het aannemen van een ongeschikte sollicitant wordt zoveel mogelijk vermeden. Dat de kans daarmee toeneemt dat een geschikte kandidaat ten onrechte wordt afgewezen, neemt men op de koop toe. De strategie van veel selecteurs is om te gaan zoeken naar negatieve informatie. Een buitenlandse opleiding waarvan de waarde niet geheel duidelijk is, een accent, een andere dan gebruikelijke presentatie kan als negatieve informatie worden opgevat. Als de relevante selectiecriteria niet goed zijn vastgelegd, wordt de kans op een oneerlijke behandeling van allochtone sollicitanten vergroot.

Het hanteren van de sign-benadering - waar eerder in dit hoofdstuk al uitgebreid op ingegaan werd - en het extrapoleren van de interviewsituatie naar de praktijk is in het traditionele interview een populaire aanpak. In het kort komt een dergelijke aanpak er op neer, dat op basis van meestal niet meer dan impressies van de kandidaat, via '... indrukwekkende en soms gewaagde psychologische theorieën over ontwikkeling, gedrag en capaciteiten' (Drenth, 1988) geëxtrapoleerd wordt naar de functie waarnaar wordt gesolliciteerd. Terwijl het extrapoleren van de interviewsituatie naar de praktijk alleen in een zeer beperkt aantal gevallen zinvol is, kunnen allochtone sollicitanten extra benadeeld worden; in de eigen cultuur kunnen andere reacties en een ander optreden gewenst zijn, of kan solliciteren op de wijze als hier gebeurt ongebruikelijk zijn.

In de meeste selectieprocedures is de selectie in handen van ongeoefende gelegenheidsselecteurs. Gezien het gebrek aan voorschriften en procedures zullen deze selecteurs aangewezen zijn op hun eigen methoden om de geschiktheid van sollicitanten vast te stellen. De kans is groot dat dubieuze methodes als stress-interviews en het stellen van strikvragen gehanteerd worden. Juist dit soort technieken doet - nog al eens ten onrechte - een groot beroep op de taalvaardigheid van sollicitanten, waardoor de kansen van sollicitanten met een eventuele taalachterstand kleiner worden.

Mogelijkheden voor een valide interview: het criteriumgericht interview

In het hierna volgende worden de mogelijkheden besproken van een meer valide interview, dat ook meer recht zal doen aan de kwaliteiten van allochtone sollicitanten. De ingrediënten voor een dergelijk interview zijn verenigd in het criteriumgericht interview. Betoogd zal worden dat behalve de voorspellende waarde, ook de kans op een eerlijke beoordeling toeneemt. Tot slot zal ingegaan worden op de verschillende factoren die onbedoelde discriminatie kunnen veroorzaken.

Eén van de bezwaren van het traditionele interview - het hanteren van niet heldere, voor meerdere uitleg vatbare functie-eisen - wordt in het criteriumgericht interview ondervangen doordat gebruik wordt gemaakt van een functieanalytische methode zoals de kritieke incidentenmethode (Flanagan, 1955). Explicitering van de criteria door het formuleren van concrete, duidelijk aan de functie gerelateerde en toetsbare functie-eisen levert niet alleen een hogere validiteit op (zie Mc Daniel e.a. 1986), maar voorkomt ook het gebruik van sociaal-normatieve functie eisen. De (subjectieve) mening van de selecteur is minder doorslaggevend, waardoor persoonlijke interpretaties en eventuele vooroordelen minder vrij spel krijgen.

In het criteriumgericht interview wordt de eerder in dit hoofdstuk beschreven sample benadering gehanteerd. Tijdens het interview worden gegevens verzameld die een steekproef vormen van voor de functie relevant gedrag. Ingewikkelde extrapolaties van impressies tijdens het interview - die vooral bij sollicitanten met een andere etnisch-culturele achtergrond vertekend kunnen zijn - naar de functie in kwestie hoeven niet te worden gemaakt. De sample benadering zal niet alleen de validiteit doen toenemen, maar ook een

eerlijker beoordeling van allochtone sollicitanten ten goede kunnen komen.

Structuur en een gestandaardiseerde aanpak zijn kenmerkend voor het criteriumgericht interview. Onderzoek toont duidelijk aan dat deze twee ingrediënten de validiteit van het interview aanmerkelijk verhogen (Latham e.a. 1980; Janz 1989; Van der Maesen de Sombreff, 1992, Campion e.a. ,1997). Dat niet alleen de kans op een eerlijker beoordeling van alle categorieën sollicitanten toeneemt zal duidelijk zijn. Daarnaast zullen door het aanbrengen van structuur en standaardisatie persoonlijke voorkeuren, en daarmee ook eventuele vooroordelen, minder ruimte krijgen.

Ook een beoordelingssysteem - bijvoorbeeld een systeem met gedragsverankerde beoordelingsschalen - is kenmerkend voor het criteriumgericht interview. Een aantal kenmerken van het traditionele interview die een oneerlijke beoordeling kunnen veroorzaken, zoals de nadruk op negatieve informatie en het optreden van contrasteffecten zullen bij gebruik van een beoordelingssysteem een minder prominente rol spelen.

Het inzetten van gelegenheidsselecteurs - zoals vaak de praktijk is - die ieder op hun eigen, vaak subjectieve en onbetrouwbare wijze de geschiktheid van sollicitanten trachten vast te stellen, is uiteraard af te raden. In principe zou elke selecteur principe een (korte) training selectietechnieken moeten volgen. In een dergelijke training kunnen selecteurs vertrouwd worden gemaakt met de methode van criteriumgericht interviewen en vooral het systematisch beoordelen met behulp van een beoordelingssysteem. Daarnaast kan selecteurs geleerd worden de juiste interviewvaardigheden toe te passen, in plaats van dubieuze methodes als stress-interviews en het stellen van strikvragen die een te groot beroep doen op de taalvaardigheid van sollicitanten. Ook het inzetten van meerdere selecteurs kan leiden tot betere resultaten: het gezamenlijke oordeel van meerdere selecteurs is betrouwbaarder dan het oordeel van één enkele selecteur.

Onbedoelde discriminatie

De hierboven genoemde voorwaarden voor een valide interview zijn uiteraard geen garantie dat directe discriminatie - waarbij etnische afkomst doelbewust en op een ontoelaatbare wijze als selectiecriteria wordt gehanteerd - wordt voorkomen. Een bevooroordeelde selecteur die koste wat kost geen allochtone sollicitant wil aannemen, zal altijd wel ergens aanleiding vinden om deze af te wijzen. Wel wordt doelbewuste afwijzing op basis van etnische afkomst en andere oneigenlijke gronden moeilijker gemaakt.

Een oneerlijke beoordeling hoeft niet per se het gevolg te zijn van doelbewuste, directe discriminatie. Er kan ook onbedoelde discriminatie optreden: door bijvoorbeeld verschillen in culturele achtergronden van selecteur en sollicitant kan miscommunicatie ontstaan, met een oneerlijke beoordeling als gevolg. Op enkele van deze factoren zal hieronder dieper worden ingegaan. Het hierna volgende is deels ontleend aan Abell (1998).

Taalgebruik

Als sollicitanten niet met het Nederlands zijn opgegroeid - wat veelal het geval zal zijn waar het de eerste generatie(s) allochtone sollicitanten betreft - en de selecteur wél, kunnen zich onvermoede communicatieproblemen voordoen. Hoewel er in algemene zin sprake kan zijn van een goede taalbeheersing - geen hoorbaar accent, grammaticaal goede zinsconstructies - moet bedacht worden dat elke taal gebruik maakt van idioom, metaforen en vaste uitdrukkingen die voor een 'tweede taalverwerver' vaak moeilijk te doorgronden zijn. De ware bedoeling van een idiomatische uitdrukking als 'dat is het hele eiereneten' is uit de letterlijke betekenis van de afzonderlijke woorden niet af te leiden. Hetzelfde geldt voor een metafoor als 'door een ringetje te halen'. Of vergelijk een Engels spreekwoord als 'It's an ill wind that blows nobody good' waarvan de betekenis de meeste Nederlanders - die over het algemeen

het Engels goed menen te beheersen - zal ontgaan.
In de spreektaal wordt veel gebruikt gemaakt van idiomatische en vaste uitdrukkingen. Timman (1984) vermeldt een onderzoek waaruit bleek dat een emotionele discussie voor 20% bestaat uit idioom. Idiomatische uitdrukkingen worden ook vaak gebruikt als men iets op een kernachtige wijze duidelijk wil maken (verg. 'achter het net vissen'). Het gebruik maken van idioom, metaforen en vaste uitdrukkingen - met het oprechte doel iets duidelijk te maken - kan averechts werken: de selecteur meent iets duidelijk te hebben gemaakt, terwijl het voor de sollicitant nog onduidelijker is wat er bedoeld wordt.

Verschillen in betekenistoekenning kunnen een andere bron van miscommunicatie vormen. Het Nederlands dat veel Surinamers en Antillianen hanteren zou omschreven kunnen worden als een vorm van dialect, of 'ethnolect' en kent tal van verschillen in betekenistoekenning. Zo heeft het woord 'liegen' voor Surinamers een zwaardere lading dan voor autochtone Nederlanders. Eerder zou een Surinamer 'jokken' gebruiken, wat in Nederland weer kinderachtig klinkt. Een Surinamer betichten van liegen kan dus meer beschuldigend overkomen dan in feite bedoeld.

Veel mensen die het Nederlands als tweede taal hebben geleerd - of thuis nog de moedertaal spreken - gebruiken het Nederlands deels volgens conventies uit de eigen moedertaal. Hierdoor kunnen woorden op een ongebruikelijke wijze worden benadrukt, is het stemvolume luider dan gebruikelijk of is de grammaticale opbouw afwijkend. Een voorbeeld is het directieve taalgebruik dat door sommige Surinamers en Antillianen wordt gehanteerd. De kans is groot dat een verzoek als 'Ga me die postzegels geven' aan een medewerker van het postkantoor - onbedoeld - irritatie zal opwekken.

Ook indirect taalgebruik - in de zin van niet direct 'to the point' komen, iets met veel omhaal van woorden duidelijk maken kan een in de moedertaal gebruikelijke wijze van uitdrukken zijn, die bijvoorbeeld bij meer gevoelige onderwerpen, of als uiting van beleefdheid wordt gehanteerd. Selecteurs zullen daarentegen eerder geneigd zijn indirect taalgebruik op te vatten als een poging om iets te verbergen, bijvoorbeeld leemtes in opleiding of arbeidsverleden: impliciet wordt verondersteld dat een sollicitant snel naar voren moeten komen met de relevante informatie.

Een laatste punt dat de communicatie betreft is het verschil in nadruk op het inhoudsniveau of het betrekkningsniveau. Het inhoudsniveau heeft betrekking op de concrete, letterlijke inhoud van de boodschap. Het betrekkningsniveau betreft de wijze waarop de informatie moet worden opgevat en hoe de één de relatie met de ander ziet (Watzlawick e.a., 1974). In Nederland geldt daarbij in zekere zin dat het betrekkningsniveau ondergeschikt is aan het inhoudsniveau. Met andere woorden: wàt er gezegd wordt is belangrijker dan hóe het gezegd wordt. In een aantal andere culturen kan juist de relatie ofwel het betrekkningsniveau de basis vormen. In dat geval geldt eerder 'c'est le ton qui fait la musique'. De aard van de relatie is dan van grote invloed op de aard en hoeveelheid informatie die de sollicitant verstrekt. In concreto houdt dit in dat dan veel meer aandacht wordt besteed aan de relationele aspecten van de communicatie waardoor vooral bij het begin van een gesprek veel beleefdheidsformuleringen worden uitgesproken en geïnformeerd zal worden naar allerlei zaken die niet met het gespreksonderwerp te maken hebben. Een dergelijke wijze van communiceren kan makkelijk irritaties bij selecteurs opwekken.

Normen en waarden

Naast factoren die met taalgebruik te maken hebben, kunnen ook de verschillende - onuitgesproken - opvattingen van sollicitant en selecteur over de wijze waarop een selectiegesprek gevoerd zou moeten worden onbedoelde discriminatie teweeg brengen. Van sollicitanten wordt verwacht dat zij zichzelf 'verkopten'. Dit houdt in, dat de sollicitanten geacht

worden zich zo positief mogelijk te presenteren en vooral hun kwaliteiten te benadrukken. In veel culturen kan dit ongebruikelijk zijn, zodat een vraag als 'kunt u mij uw positieve eigenschappen noemen' of 'wat zouden wij missen als wij u niet zouden aannemen' de sollicitant in ernstige verlegenheid kan brengen.

Een ander punt in dit verband is het statusverschil tussen selecteur en sollicitant. Sommige sollicitanten uit de etnische groepen ervaren in selectiegesprekken grote statusverschillen, en benaderen daarom de selecteurs met gepast respect. Dit kan betekenen, dat zij hen niet in de rede vallen, zeker niet tegen zullen spreken en daardoor een - in de ogen van een autochtone selecteur - passieve, afwachtende houding zullen aannemen.

Wantrouwen

Een laatste factor die in het selectiegesprek een rol kan spelen is de opstelling van de sollicitant. Selecteurs moeten zich realiseren dat de kans groot is dat allochtone sollicitanten veel negatieve sollicitatie ervaringen achter de rug kunnen hebben. Eerder in dit hoofdstuk werd het onderzoek met behulp van situatietests beschreven, waaruit bleek dat de beslissing vaker negatief uitvalt voor de allochtone kandidaat dan voor de autochtone kandidaat. Het gevolg kan zijn dat allochtone sollicitanten met veel minder zelfvertrouwen een sollicitatiegesprek aangaan dan autochtone kandidaten, met alle gevolgen van dien.

Samenvattend oordeel over het gebruik van het interview bij de selectie van allochtone sollicitanten

In het voorgaande werd betoogd dat gestructureerd, criteriumgericht interviewen directe discriminatie weliswaar niet kan voorkomen, maar wel moeilijker zal maken. Wat betreft het tegengaan van onbedoelde discriminatie is geen 'panklare' oplossing te geven.

Miscommunicatie kan ontstaan door onder andere het taalgebruik, normen en waarden en de opstelling van de sollicitant. Er a priori van uitgaan dat al deze factoren een rol zullen spelen zou veel allochtone kandidaten te kort doen: beslist niet elke allochtone kandidaat zal het Nederlands hanteren zoals hiervoor beschreven of heeft andere opvattingen over het voeren van een sollicitatiegesprek. Wat selecteurs wél kunnen doen is zich realiseren dat vragen, antwoorden en waargenomen gedrag voor beide partijen niet dezelfde betekenis hoeven te hebben, en dat met het uitspreken van een oordeel de nodige voorzichtigheid moet worden betracht.

Literatuur

Abell, J.P. (1998) Uitsluiting door selectie. In: Haffmans, M en Sijses, B. (red.) *Verborgen Drempels*. Utrecht: Forum

Arriijn. P.; Feld, S. en Nayer, A. (1998). *Discrimination in access to employment on grounds of foreign origin: the case of Belgium*. Geneva: Employment Department, International Labour Office

Beek, K. van (1993). To be hired or not to be hired, the employer decides. Relative chances of unemployed job-seekers on the Dutch labor-market. *Academisch Proefschrift*. Amsterdam: Universiteit van Amsterdam.

Bendick, M. (1996). *Discrimination against racial/ethnic minorities in access to employment in the United States: Empirical findings from situation testing*. Geneva: Employment Department, International Labour Office

- Bissessur, T. , Hockx, M. & Maesen de Sombreff, P.E.A.M. van der (1989). Psychologische selectie van allochtonen, *Gids voor Personeelsmanagement*, 4, 11-17
- Böning, W.R. en Zegers de Beijl, R. (1995). *The integration of migrant workers in the labour market: Policies and their impact*. Geneva: Employment Department, International Labour Office
- Campion, J. E. en Arvey, R.D. (1989) Unfair discrimination in the employment interview. In: Eder, R.W. en Ferris, G.R.(eds.). *The Employment Interview. Theory, Research and Practice*. Newbury Park: Sage.
- Campion, M.A., Palmer, D.K., & Campion, J.E. (1997). A review of structure in the selection interview, *Personnel Psychology*, 50, 655-702.
- Chan, D. & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: subgroup differences in test performance and face validity perceptions, *Journal of Applied Psychology*, 82, 143-159
- Chan, D., Schmitt, N., DeShon, R.P., Clause, C.S., & Delbridge, K. (1997). Reactions to cognitive ability tests: the relationships between race, test performance, face validity perceptions, and test-taking motivation, *Journal of Applied Psychology*, 82, 300-310
- Drenth, P. J. D. (1988). De waarde van het selectie-interview. *Gedrag en Organisatie*, 1, 18-25.
- Drenth, P.J.D. (1988). Psychologische selectie en discriminatie. *Gedrag en Organisatie*, 1, 13-21
- Epstein, S. & O'Brien, E.J. (1985). The person-situation debate in historical and current perspective, *Psychological Bulletin*, 98, 513-537
- Essed., Ph. (1991). *Understanding Everyday Racism*. Wijk bij Duurstede: Huntenaar.
- Flanagan, J.C. (1955). The critical incident technique. *Psychological Bulletin*, 51(4), 327-358.
- Ford, J.K., Kraiger, K. & Schlechtman, S.L. (1986). Study of race effects in objective indices and subjective evaluations of performance: A meta-analysis of performance criteria, *Psychological Bulletin*, 99, 330-337
- Gras, M.J.I.; F. Bovenkerk; K. Gorter; P. Kruiswijk & D. Ramssoedh (1996). *Een schijn van kans: Twee empirische onderzoeken naar discriminatie op grond van handicap en etnische afkomst*. Deventer: Gouda Quint
- Greenfield, P.M. (1997). You can't take it with you. Why ability assessments don't cross cultures, *American Psychologist*, 52, 1115-1124
- Guion, R.M. (1978). Scoring of content domain samples: the problem of fairness. *Journal of Applied Psychology*, 63, 499-506
- Hale, M. (1982). History of employment testing. In A. Wigdor & W. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (Part II). Washington: National Academy Press

- Herrnstein, R.J., & Murray, Ch. (1994). *The bell curve. Intelligence and class structure in American life*, New York: The Free Press
- Hoekstra, H.A. (1995). Managementselectie via simulaties: de methodologie van het assessment center. In: F.J.R.C. Dochy & T.R. de Rijke (Eds.), *Assessment centers. Nieuwe toepassingen in opleiding, onderwijs en HRM* (pp. 53-72), Utrecht: Lemma
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance, *Psychological Bulletin*, 96, 72-98
- Hunter, J.E., & Schmidt, F.L. (1982). Fitting people to jobs: The impact of personnel selection on national productivity. In: M.D. Dunnette & E.A. Fleishman (Eds.), *Human capability assessment*, Hillsdale, N.J.: Lawrence Erlbaum
- Janz, T. (1989). In : Eder, R.W. en Ferris, G.R. (eds.) *The Employment Interview: theory, research and practice*. Newbury Park: Sage.
- Klinkenberg, E. (1997). *Rapportage onderzoek: betrouwbaarheid en interne structuur LTSB*, Woerden: Meurs Personeelsadvies
- Kravitz, D.A., Harrison, D.A., Turner, M.E. et al. (1996). Affirmative action: a review of psychological and behavioral research, *Report of the Society for Industrial and Organizational Psychology*
- Latham, G.P., Saari., L.M. Pursell, E.D. en Campion, M.A. (1980). The situational interview. *Journal of Applied Psychology*, 61, 193-198.
- Leest, P. van (1997). Persoonlijkheidsmeting bij allochtonen. De gebruikswaarde van persoonlijkheidsschalen bij selectie van allochtonen. *Academisch Proefschrift Vrije Universiteit van Amsterdam*, Lisse: Swets & Zeitlinger, 1997
- Leest, P. van & Dutman, A. (1994). De voorspellende waarde van de psychologische selectie voor de Turken- en Marokkanenpool van de politie, *Paper gepresenteerd op Nederlands Psychologencongres*, Nijmegen
- Linn, R.L. & Hastings, C.N. (1984). Validity of predictors of performance in law school. *Journal of Educational Measurement*, 21, 245-259
- Maesen de Sombreff, P.E.A.M. van der (1992). Het rendement van personeelsselectie, *Academisch Proefschrift RUG*
- Maesen de Sombreff, P.E.A.M. van der (1995). *Selectie van adspirant-agenten op morele integriteit*, Interne publicatie, Hilversum: LSCP
- Maesen de Sombreff, P.E.A.M. van der (1997a). Intelligentietests en beroepssucces, in: W. Tomic & H.T. van der Molen (Eds.), *Intelligentie en sociale competentie*, (hoofdstuk 12, bldz. 227-242). Lisse: Swets & Zeitlinger, Open Universiteit Heerlen
- Maesen de Sombreff, P.E.A.M. van der (1997b). De positie van allochtone burgers op de arbeidsmarkt, *Gids voor Personeelsmanagement*, 76, 11, 54-62
- Maesen de Sombreff, P. van der & Noort, G.J. van (1989). Videotest voor communicatieve vaardigheden, *Personeelbeleid*, 25, 71-76

- Maesen de Sombreff, P.E.A.M. van der & Westen, G. (1991). Praktijkopdracht voor dataverwerking bij de Belastingdienst, *Personeelbeleid*, 27, 593-596
- McClelland, D.C. (1973). Testing for competence rather than for intelligence, *American Psychologist*, 28, 1-14
- McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., Hunter, J.E., Mauer, S. en Russell, J. (1986). *The validity of employment interviews: a review and meta-analysis*. Defense Personnel Security Research and Education Centre, Monterey Cal . USA.
- McHenry, J.J., Hough, L.M., Toquam, J.L., Hanson, M.A., & Ashworth, S. (1990). Project A validity results: the relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354
- Meer, P. van der & Wielers, R. (1997). Hoe belangrijk is opleiding voor de arbeidsmarktpositie van allochtonen? *Mens & Maatschappij*, 72, 40-47
- Meltzer, P.H. (1995). Videotest voor communicatieve vaardigheden, in: F.J.R.C. Dochy & T.R. de Rijke (Eds.), *Assessment centers. Nieuwe toepassingen in opleiding, onderwijs en HRM*, Utrecht: Lemma, 109-122
- Minderhedenbeleid 1997 (1996). *Tweede kamer, vergaderjaar 1996-1997*, 25001, nrs. 1-2. 's-Gravenhage: SDU
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley
- Moelker, R. (1992). *Zou hij onze nieuwe werknemer kunnen zijn? Veranderingen in gevraagde kwalifikaties als weerspiegeling van veranderingen in de betekenis van prestaties. Een inhoudsanalyse van personeelsadvertenties*. De Lier: Academisch Boeken Centrum.
- Moonen, H. (1995). Sociale intelligentie niet te filmen! Of toch.... Ontwikkeling en evaluatie van de videotests VASE (Video Assessment Sociale Effectiviteit), *Doctoraalscriptie Vakgroep Personeelwetenschappen KUB*, Utrecht: B/CPO WS&L Belastingdienst
- Nederlands Instituut van Psychologen (1988). *Richtlijnen voor ontwikkeling en gebruik van psychologische tests en toetsen*, Amsterdam: NIP
- Nijenhuis, J. te (1997). Comparability of test scores for immigrants and majority group members in the Netherlands, *Academisch Proefschrift VU*, Amsterdam: Vrije Universiteit
- Ployhart, R.E., & Ryan, A.M. (1998). Applicants reactions to the fairness of selection procedures: the effects of positive rule violations and time of measurement, *Journal of Applied Psychology*, 83, 3-16
- Pukalos, E.D., & Schmitt, N. (1996). An evaluation of two strategies for reducing adverse impact and their effects on criterion-related validity, *Human Performance*, 9, 241-258
- Ree, M.J. & Earles, J.A. (1991). Predicting training success: not much more than g. *Personnel Psychology*, 44, 321-332
- Ree, M.J., Earles, J.A. & Teachout, M.S. (1994). Predicting job performance: not much more

than g. *Journal of Applied Psychology*, 79, 518-524

Robertson I.T., & Kandela, R.S. (1982). Work sample tests: validity, adverse impact and applicant reaction. *Journal of Occupational Psychology*, 55, 171-183

Robertson, I.T., & Downs, S. (1989). Work-sample tests of trainability: a meta-analysis. *Journal of Applied Psychology*, 74, 402-410

Roe, R.A. (1983). *Grondslagen der personeelsselectie*. Assen: Van Gorcum

Rowe, P.M. (1989). Unfavorable Information and Interview Decisions. In: Eder, R.W. en Ferris, G.R.(eds.). *The Employment Interview. Theory, Research and Practice*. Newbury Park: Sage

Rushton, J.P., Brainerd, C.J., & Pressley, M. (1983). Behavioral development and construct validity: the principle of aggregation, *Psychological Bulletin*, 94, 18-38

Schmitt, N. & Ostroff, C. (1986). Operationalizing the "behavioral consistency" approach: selection test development based on a content oriented strategy. *Personnel Psychology*, 39, 91-108

Schmitt, N. (1987). Fairness in employment selection, *Paper presented on the UMIST International Conference on Advances in Selection and Assessment*, Buxton, Engeland

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, 719-730

Sociaal en Cultureel Planbureau (1996). *Rapportage Minderheden*, Rijswijk

Tesser, P.T. M., van Dugteren, F.A. en Merens, J.G.F. (1998). *Rapportage minderheden 1998*. Rijswijk: SCP.

Timman, Y. (1989). *De sleutels van de taal*. Amsterdam: Universiteit van Amsterdam, Instituut voor Algemene Taalwetenschap

Watzlawick, P; Helmick Beavin, J en Jackson, D.D. (1974). *De pragmatische aspecten van de menselijke communicatie*. Deventer: Van Loghem Slaterus.

Weekley, J.A., & Jones, C. (1997). Video-based situational testing. *Personnel Psychology*, 50, 25-49

Wernimont, P.F. & Campbell, J. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52, 372-376

Zanten, D. van, Maesen de Sombreff, P.E.A.M. van der, & Posthumus, H. (1996). *Handleiding LTSB. Interne publicatie*, Hilversum: LSOP

Zeidner, M. (1987). Test of the cultural bias hypothesis: some Israeli findings, *Journal of Applied Psychology*, 72, 38-48