



Faculteit Psychologie en Pedagogische Wetenschappen

Academiejaar 2009-2010

Welk effect hebben gestandaardiseerde oefeninginstructies op
Assessment Center Oefeningen?

Masterproef neergelegd tot het behalen van de graad van
Master in de Psychologie optie Bedrijfspsychologie en Personeelsbeleid.

Door Saartje Cromheecke

Saartje.Cromheecke@Ugent.be

Promotor: Prof. Dr. F. Lievens

Begeleiding: Lic. Eveline Schollaert

Saartje Cromheecke

20055181

Ondergetekende, Saartje Cromheecke geeft toelating tot het raadplegen van de scriptie door derden.

Saartje Cromheecke

ABSTRACT

In de huidige bedrijfswereld is het belang van assessment centers (ACs) niet meer weg te denken (Thornton, 2006). Ondanks het feit dat ze vrij goede voorspellers zijn van toekomstig gedrag blijkt er een probleem te zijn met de constructvaliditeit. Verschillende dimensies correleren onderling hoger in dezelfde oefening dan dat verschillende dimensies dat doen over verschillende oefeningen heen (Lievens, 2002). Dit probleem werd reeds benaderd vanuit de trek activatie theorie (TAT) (Tett & Guterman, 2000); deze studie gaat hierop verder. De TAT werd in deze scriptie, in tegenstelling tot vorige onderzoeken, als prescriptief middel gebruikt. Dit door middel van gestandaardiseerde oefeninginstructies, die gebruikt worden in AC oefeningen om trekrelevant gedrag uit te lokken. Er werd een positief effect van deze gestandaardiseerde oefeninginstructies verwacht op de constructvaliditeit en de interbeoordelaarsbetrouwbaarheid. Dit zou kunnen verklaard worden aan de hand van het trekactivatiepotentieel (trait activation potential, TAP; Tett & Burnett, 2003).

De data voor deze scriptie werden ingezameld tijdens een sollicitatietraining. De resultaten toonden een positief effect van het gebruik van gestandaardiseerde oefeninginstructies op de constructvaliditeit, zowel de convergente als de discriminante validiteit verbeterden. Het verwachte positief effect van het gebruik van gestandaardiseerde oefeninginstructies op de interbeoordelaarsbetrouwbaarheid werd in deze scriptie niet gevonden.

Theoretische en praktische implicaties van het onderzoek worden besproken en suggesties voor toekomstig onderzoek worden aangereikt.

INHOUD

ABSTRACT	i
INHOUD.....	ii
LIJST MET TABELLEN	v
VOORWOORD	vi
INLEIDING	1
ASSESSMENT CENTER.....	4
Historisch	4
Eigenschappen.....	5
Soorten Assessment Center.....	6
Assessment Center voor Selectie.....	6
Assessment Center voor Ontwikkeling	7
Diagnostisch Assessment Center.....	7
Mini-Assessment Center.....	7
Locatie van het Assessment Center	7
Praktijksimulaties.....	8
Postbak Oefening.....	9
Mondelinge Presentatie	9
Groepsdiscussie	9
Rollenspel	9
Rollenspeler.....	10
Assessoren.....	10
Psychometrische Kenmerken	12
Interbeoordelaarsbetrouwbaarheid	12
Face Validiteit.....	13
Inhoudsvaliditeit	13
Criteriumvaliditeit.....	13
Constructvaliditeit.....	14
Paradox	14
TREK ACTIVATIE THEORIE	17
Interactionistische Visie	17
Trek Activatie Theorie en Assessment Center.....	19

Trek Activatie Theorie en Oefeninginstructies	21
METHODE	24
Steekproef	24
Procedure.....	24
Opzet.....	24
Dimensies	25
Assessorentraining.....	26
AC Oefeningen.....	27
Werkgesprek	27
Presentatie.....	27
Gestandaardiseerde Oefeninginstructies	28
Experimenteel Design.....	28
Opstellen van Gestandaardiseerde Oefeninginstructies.....	28
Manipulation Check.....	29
Design Overwegingen	29
Data-analyse	30
RESULTATEN	32
Descriptieve resultaten	32
Werkgesprek	32
Presentatie.....	32
Normaalverdelingen	32
Constructvaliditeit	33
Convergente Validiteit.....	34
Discriminante Validiteit.....	35
Interbeoordelaarsbetrouwbaarheid	36
DISCUSSIE	37
Bespreking resultaten	37
Constructvaliditeit.....	37
Convergente validiteit.....	37
Discriminante validiteit	39
Interbeoordelaarsbetrouwbaarheid	40
Methodologische reflecties	41

Steekproef	41
Procedure	42
Data-analyse	43
Theoretische en Praktische Implicaties	44
Toekomstig Onderzoek	45
Inhoudelijk	45
Methodologie	46
Conclusie	48
REFERENTIES	49
BIJLAGEN	54
Bijlage 1	54
Bijlage 2	55
Bijlage 3	56
Bijlage 4	57
Bijlage 5	58
Bijlage 6	61
Bijlage 7	65
Bijlage 8	66

LIJST MET TABELLEN

Tabel 1	Beschrijvende statistieken en correlaties tussen de variabelen van deze studie.....	31
Tabel 2	Generaliseerbaarheidsanalyse	34

VOORWOORD

Voor mij vormt deze scriptie een prachtige afsluiter van een vijf jaar durende leerrijke studie. Deze scriptie en mijn studies waren nooit mogelijk geweest zonder de hulp van een aantal mensen.

Vooreerst wil ik graag mijn promotor Prof. Dr. Filip Lievens bedanken. Dankzij hem heb ik een scriptie kunnen schrijven rond een uitermate boeiend en praktijkrelevant onderwerp.

Mijn begeleidster Lic. Eveline Schollaert wil ik eveneens bedanken. Zij stond altijd klaar voor mij, met tips en aanwijzingen, en hielp mij om deze scriptie tot een goed einde te brengen.

Daarnaast is dit het ideale moment om mijn ouders en vriend, die mij gedurende het schrijven van deze scriptie en tijdens mijn volledige studies gesteund hebben, te bedanken. Ik heb veel kracht gehaald uit jullie steun en aanmoedigingen.

Ten slotte wil ik iedereen die rechtstreeks en onrechtstreeks bij deze scriptie was betrokken, van de data-inzameling tot het herlezen, oprecht bedanken. Hopelijk genieten jullie evenzeer van het lezen van deze scriptie als ik heb gedaan bij het schrijven ervan.

Saartje Cromheecke

Machelen aan de Leie, april 2010

INLEIDING

Assessment centers (ACs) maken steeds meer opgang in de wereld van het personeelsmanagement. Vooral in het kader van selectieprocedures worden ACs veel toegepast, zowel in de private als in de publieke sector (Thornton, 2006). Het AC proces wordt vandaag in een verscheidenheid van settings gebruikt zoals overheid, onderwijs industrie, ordediensten, militaire settings, en op organisatieniveau (Howard, 1997).

Het AC proces heeft als doel het evalueren en ontwikkelen van personeel. Via ACs bekomt men een gestandaardiseerde evaluatie van gedrag. Er wordt gebruik gemaakt van praktijksimulaties waarbij meerdere getrainde assessoren observeren (International Task Force, 2009).

ACs zijn zo populair omdat er enkele interessante voordelen aan verbonden zijn. Dankzij deze methode is het mogelijk om vaak complex gedrag op een adequate manier te meten en te voorspellen. Deze metingen en voorspellingen gebeuren op basis van attributen en mogelijkheden die relevant zijn voor de effectiviteit van de organisatie. Het adequaat meten van dit complex gedrag is vooral mogelijk door de praktijksimulaties die in het AC proces gebruikt worden, welke een weerspiegeling vormen van de functie. Zo bieden ze een representatieve steekproef van de taken die een persoon zal moeten uitvoeren (Jansen & de Jongh, 1997). De kandidaten worden verwacht te reageren zoals ze dat in de praktijk zouden doen en zo het gewenste gedrag openlijk te vertonen. De simulaties hebben een groot interactief karakter, waardoor ACs zich onderscheiden van andere technieken die in selectieprocedures worden gebruikt. Deelnemers zien deze methode als fair en ze wordt dan ook positief onthaald (Fletcher & Kerslake, 1992; Thornton, 1992, 2006). Door de goede predictieve waarde, de jobgerelateerdheid, het positieve imago, en het hoge prijskaartje worden ACs wel eens de Rolls Royce onder de selectietechnieken genoemd (Lievens, 2006).

Desondanks de voordelen die het biedt, blijkt er een probleem te zijn met de constructvaliditeit. Het voornaamste probleem is dat de beoordelingen van verschillende dimensies te sterk samenhangen met elkaar binnen eenzelfde oefening. Daardoor zouden de scores afhankelijk zijn van het type oefening en niet van de te scoren dimensies (Henderson, Anderson, & Rick, 1995; Sackett & Dreher, 1982). Er werd reeds een aanzet gegeven tot enkele aanpassingen van het design en de implementatie

van ACs (Lievens, 2002; Woehr & Arthur, 2003). De auteurs rapporteerden bij sommige van deze aanpassingen een lichte stijging van de constructvaliditeit, maar het probleem blijft bestaan. Deze scriptie zal zich baseren op de trek activatie theorie (TAT) van Tett en Guterman (2000) en zo proberen om een verbetering van de constructvaliditeit te verkrijgen.

De TAT stelt dat een persoonlijkheidstrek of dimensie pas tot uiting kan komen wanneer de situatie bepaalde cues bevat, die ervoor zorgen dat die trek tot uiting kan komen. Opdat de assessor de trek kan observeren is het dus nodig dat deze trek tot uiting komt. In dit kader kan deze theorie mogelijkheden bieden voor het verbeteren van de constructvaliditeit (Lievens, 2009). Toegepast op het AC proces zal men proberen om het relevante gedrag van de trek op te roepen. Op basis van de TAT zullen voor de AC oefeningen gestandaardiseerde oefeninginstructies ontwikkeld worden om het trekrelevante gedrag tot uiting te laten komen. Op deze manier wordt de TAT prescriptief angewend, wat vernieuwend is in de wetenschappelijke literatuur.

Indien de gestandaardiseerde oefeninginstructies het trekrelevante gedrag beter oproepen, zouden ze een rol kunnen spelen in het verbeteren van de constructvaliditeit (Thornton & Mueller-Hanson, 2004). Verder zou door deze cues de overeenkomst tussen de beoordelingen van de verschillende assessoren verhoogd kunnen worden.

Naast deze theoretische bijdrage kan deze toepassing van de TAT ook van praktisch belang zijn. Het opstellen van gestandaardiseerde oefeninginstructies is een kleine aanpassing van het design, die geen grote kost met zich meebrengt. Deze kleine aanpassing zou een verbetering kunnen betekenen wat betreft de kwaliteit van de beoordeling in het AC. De gestandaardiseerde oefeninginstructies zullen in dit onderzoek worden toegepast in het kader van twee praktijksimulaties: Het rollenspel en de mondelinge presentatie. Twee essentiële onderdelen van het AC proces (Kok & De Jongh, 2004).

Concreet zal de invloed worden nagegaan van de gestandaardiseerde oefeninginstructies in de AC oefeningen op de constructvaliditeit en de interbeoordelaarsbetrouwbaarheid.

Eerst zullen enkele belangrijke concepten voor dit onderzoek verduidelijkt worden. Zo wordt een gemeenschappelijk begrippenkader gecreëerd. Daarna worden, met betrekking tot het huidige onderzoek, enkele hypothesen geformuleerd. Vervolgens worden de algemene resultaten van deze studie weergegeven. Ten slotte volgt een kritische bespreking van deze resultaten en worden enkele implicaties voor de praktijk meegegeven.

ASSESSMENT CENTER

ACs zijn niet meer weg te denken uit het bedrijfsleven. Het is een gestandaardiseerde evaluatie van gedrag die gebaseerd is op verschillende informatiebronnen met meerdere getrainde observatoren en verschillende technieken (International Task Force, 2009). Veel mensen denken dat ACs specifieke plaatsen zijn waar mensen naartoe kunnen gaan om testen te maken. Deze veronderstelling klopt echter niet: Het is een methode om persoonlijke kwaliteiten en dimensies vast te stellen. Daarom wordt er vaak gesproken van de assessment methode of het AC proces (Lievens & Van Keer, 1999; Van Beirendonck 1998).

Historisch

Het leger speelde een zeer belangrijke rol in de ontwikkeling van de methode voor personeelsevaluatie die vandaag wordt toegepast in ACs. Tijdens de Tweede Wereldoorlog wou het Duitse leger goede manschappen rekruteren. Er heerste grote werkloosheid waardoor men methodes moest vinden om de beste manschappen te selecteren. Zo kwam het dat men, naast de klassieke test, ook enkele functie-relevante praktische proeven ging gebruiken. Deze kunnen de voorloper van ACs genoemd worden (Van Beirendonck, 1998).

In 1948 werd door het Office of Strategic Services het boek *Assessment of Men* uitgebracht. Hierin werden voor de functie-relevante praktische proeven zowel de methodes als de resultaten uitgewerkt. Dit bureau stond onder militaire controle, toch werkten hier veel burgers die een psychologische opleiding hadden genoten. Door de aanwezigheid van deze burgers werd een transfer mogelijk van een militaire naar een niet-militaire context. In 1956 werden ACs op grote schaal geïmplementeerd door Douglas Bray in The American Telephone and Telegraph Company (AT&T). Het AC proces zoals het hier werd gebruikt was toonaangevend voor de volgende jaren (Riggio & Mayes, 1997).

Door de stijging van de populariteit van de assessment methode, ontwikkelden er zich toepassingen die niet volledig in overeenstemming waren met de oorspronkelijke richtlijnen. Toch vielen deze toepassingen ook onder de noemer van ACs, waardoor ze

voor een blaam op het imago zorgden (Thornton, 2006). Zo kwam er nood aan ethische richtlijnen. De belangrijkste afspraken werden neergeschreven in *Guidelines and Ethical Considerations for Assessment Center Operations* of kortweg *Guidelines* (International Task Force, 2009). Indien er aan de fundamentele kenmerken van het AC proces niet werd geraakt, lieten deze Guidelines innovatie toe. In de loop der jaren werden ze dan ook al enkele malen herzien (Howard, 1997).

In het begin van de jaren '70 werden veel consultingbureaus opgericht die ACs aanboden aan bedrijven. Hierdoor kwam de grote doorbraak voor de assessment methode. Deze methode kende een explosieve groei en werd vervolgens veelvuldig toegepast (Howard, 1997; Riggio & Mayes, 1997). In het Belgische bedrijfsleven maakte deze methode in de jaren '80 opgang en werd ze frequent toegepast (Lievens & Van Keer, 1999).

Eigenschappen

De *Guidelines* (International Task Force, 2009) vat het unieke van het AC proces samen in vijf primordiale kenmerken. De combinatie van deze kenmerken, om mensen te beoordelen op hun persoonlijke kwaliteiten, vindt men enkel terug bij ACs.

Ten eerste is het gebruikelijk om voorafgaand een *jobanalyse* uit te voeren. Een jobanalyse is het proces dat gebruikt wordt voor het bepalen van de dimensies, welke geassocieerd worden met slagen of falen in een job (International Task Force, 2009). Hierop worden alle verdere stappen van het AC proces gebaseerd.

Ten tweede bestaan ACs uit een *diversiteit aan informatiebronnen*: Persoonlijkheidsvragenlijsten, intelligentietesten, interviews, en praktijksimulaties. Door de diversiteit aan informatiebronnen slaagt men erin om het complexe gedrag van de assessee te bestuderen (International Task Force, 2009; Lievens, 2006). De assessee zijn de personen waarvan het gedrag wordt geëvalueerd gedurende het AC proces (Thornton, 2006).

Ten derde is het nodig dat de *assessoren ook voldoende getraind* worden om de persoonlijke kwaliteiten goed te evalueren. De assessoren zijn de getrainde observatoren die het gedrag van de assessee evalueren. Zij moeten in staat zijn om verschillende taken te vervullen, zowel tijdens als na ACs. Ze moeten kunnen observeren, noteren, het

gedrag classificeren in dimensies, het prestatieniveau beoordelen, de eigen observatie meedelen, en de informatie afkomstig van andere observatoren integreren. Met deze acties maken ze een evaluatie van het geheel en geven ze suggesties voor ontwikkelingen in de toekomst (Thornton & Mueller-Hanson, 2004).

Ten vierde is het noodzakelijk dat de assessee door *verschillende assessoren* wordt beoordeeld. Om eventuele vertekeningen te voorkomen is een rotatie van de assessoren eveneens aangewezen (Lievens, 2006; Thornton, 2006).

Ten vijfde is het ook van belang dat de *observaties op correcte wijze worden geïntegreerd*. Na afloop van het AC proces zullen de assessoren de informatie over de assessee's verwerken. Indien er verschillen zijn in de scores worden deze in groep besproken tot ze tot een overeenkomst komen (Lievens 2006; Thornton, 2006).

Soorten Assessment Center

Onder de noemer Human Resources (HR) kunnen diverse functies worden ondergebracht: Rekrutering, selectie, plaatsing, training en ontwikkeling, prestatiebeoordeling, organisatieontwikkeling, HR planning, promotie en overplaatsing, en ontslag (Thornton, 2006). Het AC proces kan voor elk van deze functies een nuttige bijdrage leveren. Om een goede toepassing te krijgen is het belangrijk dat de kenmerken van ACs overeenstemmen met het doel ervan (Thornton, 2006). Er kan een opdeling gemaakt worden in drie categorieën naargelang het doel: Selectie, ontwikkeling, en diagnose. Deze drie zullen achtereenvolgens worden besproken. Daarna zal er ook nog dieper worden ingegaan op het mini-assessment en de locatie van ACs (Kok & De Jongh, 2004).

Assessment Center voor Selectie

Deze ACs vinden plaats bij personeelsselectie. Het doel hier is om te achterhalen wat de kans is dat een persoon succesvol zal functioneren in de functie (Kok & De Jongh, 2004).

Assessment Center voor Ontwikkeling

Een assessment gericht op ontwikkeling wordt ook wel een development center genoemd. De vraag is hier niet meer of een persoon geschikt is voor een functie, maar op welke punten een persoon sterk is, en welke punten hij nog verder moet ontwikkelen. Er wordt hier als het ware een sterkte-zwakteanalyse uitgevoerd (Lievens, 2006; Thornton, 2006). Er bestaan twee soorten development center: Het potentieelassessment en het functioneringsassessment. Bij het eerste wordt er gekeken welk potentieel werknemers hebben en in welke mate ze dat kunnen ontwikkelen naar de toekomst toe. Vaak wordt het potentieelassessment toegepast bij jongere werknemers. Bij het functioneringsassessment zoekt men een oorzaak voor de problemen die een werknemer ondervindt in zijn dagelijks functioneren (Kok & De Jongh, 2004).

Diagnostisch Assessment Center

Hier ligt de focus op de wens om van functie te veranderen. Er zijn twee vormen te onderscheiden op basis van wie de wens uitdrukt: Het loopbaan assessment en outplacement. Bij een loopbaanassessment heeft de werknemer de wens om van functie te veranderen. Als de wens uitgaat van het bedrijf waar de werknemer tewerkgesteld is, wordt er gesproken van outplacement. Bij deze vormen van assessment wordt er gepeild naar de interesses van de persoon en worden deze gediagnosticeerd. Vaak wordt vooraf een interview afgenomen om de interesses te leren kennen en daar wordt dan het assessment op gebaseerd (Kok & De Jongh, 2004).

Mini-Assessment Center

Het mini-assessment center bestaat slechts uit één tot twee opdrachten. Dit soort assessment wordt vaak gebruikt wanneer er tijdsdruk is of als er weinig financiële middelen voorhanden zijn (Van Minden, 2006). Door de economische crisis en de vaak noodzakelijke besparingen wordt deze vorm van ACs steeds vaker toegepast.

Locatie van het Assessment Center

Zoals eerder al vermeld zijn ACs geen plaatsen. Het AC proces kan dus plaatsvinden op verschillende locaties. Er zijn twee mogelijkheden: Een in-company groepsassessment en een individueel programma bij een bureau. De eerste komt vooral

voor bij grote organisaties; zij organiseren ACs *in house*. Dit kan op hun bedrijf, in een hotel of op een conferentieplaats. Verschillende deelnemers, die niet noodzakelijk concurrenten zijn, nemen deel aan deze ACs. Recent komt het individueel programma bij een assessment bureau frequenter voor. Bedrijven huren een assessment bureau in om alles te organiseren (Kok & De Jongh, 2004).

Praktijksimulaties

Praktijksimulaties zijn waarschijnlijk het meest kenmerkend aan ACs. Het zijn oefeningen of technieken die ontworpen zijn om, via situationele stimuli, gedrag uit te lokken. Dit gedrag is gerelateerd aan de dimensies die verband houden met de latere jobprestaties (International Task Force, 2009). De assessee krijgt enkele complexe stimuli aangeboden; hij/zij moet aantonen met deze complexe set van stimuli overweg te kunnen. Om alles zo realistisch mogelijk te maken, gebeurt het aanbieden van deze stimuli in een situatie die gelijkend is op de targetfunctie. In de praktijksimulatie moet de assessee het complex gedrag dat vereist is voor de targetfunctie openlijk vertonen.

Dat complex gedrag heeft twee voordelen. Ten eerste blijkt dit een goede voorspeller te zijn van het toekomstige gedrag. Ten tweede is het een goede bron van feedback naar de kandidaat toe. De assessee is zich namelijk meer bewust van zijn gestelde gedrag dan van dieperliggende eigenschappen (Kok & De Jongh, 2004).

Het gaat hier dus niet om covert gedrag, zoals bijvoorbeeld bij een bevraging door Likertschalen, maar om openlijk vertoond gedrag met een interactief karakter.

Uiteraard is de afwezigheid van dergelijk complex gedrag ook belangrijk, in sommige gevallen misschien zelfs belangrijker (Kok & De Jongh, 2004; Thornton, 2006).

Het gedrag kan worden nagegaan aan de hand van verschillende types van AC oefeningen: Een postbak oefening, een mondelinge presentatie, een groepsdiscussie, en een rollenspel. Voor elk van deze oefeningen is reeds aangetoond dat ze waardevol kunnen zijn (Thornton & Byham, 1982; Thornton & Mueller-Hanson, 2004).

Postbak Oefening

De postbak oefening of in-basket is een simulatieoefening die lange tijd zeer frequent gebruikt is, maar nu iets minder wordt toegepast. De assessee krijgt bij deze simulatie een verzameling van e-mails en memo's van een manager die een tijdje niet op zijn werk is geweest. Er wordt gevraagd aan de assessee om al deze informatie binnen een bepaalde tijdspanne te verwerken. Tegenwoordig zijn ook elektronische versies van deze oefening beschikbaar (Kok & De Jongh, 2004).

Mondelinge Presentatie

Voor deze oefening is de presentatievaardigheid cruciaal: Enthousiasme, verbale helderheid, overzichtelijkheid, uitdrukkingsvaardigheid, overtuigingskracht, en levendigheid zijn factoren die van belang zijn. De assessee zal dan ook eerder op de vorm dan op de inhoud worden beoordeeld. Voorafgaand krijgt men de tijd om alles voor te bereiden. De bedoeling is dat de assessee concreet kan uitleggen wat het probleem is en hoe het moet aangepakt worden (Kok & De Jongh, 2004).

Groepsdiscussie

Bij een groepsdiscussie moeten de assessees door overleg en bespreking een bepaald probleem oplossen. Het doel van het AC proces bepaalt het onderwerp van het probleem dat moet worden aangepakt (Kok & De Jongh, 2004).

Rollenspel

Het doel van een rollenspel is om de assessee zoveel mogelijk de kans te geven om zijn vaardigheden te tonen. Het gaat hier vooral om interpersoonlijke vaardigheden. Aan de hand van zeer realistische simulaties van jobsituaties worden deze ingeschat. Het is een moment waarop kandidaat en rollenspeler met elkaar in interactie treden (Kok & De Jongh, 2004).

Rollenspeler

De rollenspeler neemt de rol aan van iemand anders, terwijl de assessee zichzelf moet zijn. De assessee moet reageren zoals hij zou reageren indien de gestelde situatie zich zou voordoen in de realiteit. Zo verkrijgt men informatie over de capaciteiten van de kandidaat. De constructen die worden bekeken hangen af van de doelstellingen van het rollenspel (Jansen & De Jongh, 1997).

Er is geen vast script voor de rollenspeler want er moet ruimte gelaten worden voor de verschillende reacties van de kandidaten. Het is wel nodig dat ze enige consistentie vertonen in hun reacties, zodat de kandidaten gelijke mogelijkheden krijgen (Kok & De Jongh, 2004).

Het kan nuttig zijn om de rollenspelers een training te geven. Zij moeten namelijk het dimensiegerelateerde gedrag uitlokken en de onbedoelde variantie in de oefening beperken. Als zij hierin worden getraind, zal het relevante gedrag makkelijker zichtbaar zijn (Lievens, 1998).

Assessoren

Zoals reeds vermeld zijn de assessoren, de observatoren of beoordelaars die bij een praktijksimulatie aanwezig zijn. Zij zullen meestal een vast patroon hanteren bij de beoordeling. Sommige auteurs beschrijven het beoordelingsproces zoals terug te vinden in Bijlage 1. De verschillende fases van dit model worden hier overlopen (Lievens & Van Keer, 1999).

In een eerste fase gaan de assessoren de kandidaten observeren tijdens de praktijkopdrachten. Er moet gelet worden op zowel verbaal als non-verbaal gedrag.

In de tweede fase moeten de assessoren noteren wat ze waarnemen. Een gedragsmatige checklist kan hier een hulp zijn om ervoor te zorgen dat iedereen dezelfde belangrijke gedragingen observeert (Lievens, 1998).

In de derde fase worden de gedragingen geclassificeerd volgens vooropgestelde gedragskenmerken. Het is van belang dat de gedragskenmerken en hun definities representatief zijn voor de cultuur en het referentiekader van de organisatie. Daarom is

het gebruikelijk dat deze worden bepaald door mensen die vertrouwd zijn met de organisatie. Dit kunnen bijvoorbeeld mensen zijn van de personeelsafdeling of de lijnmanagers. De kandidaten moeten dan een score krijgen op die gedragskenmerken. Er bestaan hiervoor vier scoringsmethodes (Bijlage 2): De gedragsrapportering methode, de oefening-dimensie methode, de binnen-oefening methode, en de cross-oefening methode (Robie, Osburn, Morris, Etchegaray, & Adams, 2000).

Tijdens de vierde fase zullen de assessoren discussiëren over de kandidaten. Hier komen de gemaakte aantekeningen van pas. De discussie is nodig omdat er zo tot een gefundeerde en meer genuanceerde beoordeling kan worden gekomen over het profiel van iedere assessee (Lievens, 2006; Thornton, 2006).

In de laatste fase worden de bevindingen van de assessoren in een rapport gegoten. In de loop van de week volgend op het AC proces wordt zowel aan de assessee als aan het bedrijf feedback gegeven. Het rapport is een weerspiegeling van hoe de assessee scoort op de verschillende vooraf opgestelde dimensies (Thornton, 2006). Wat er concreet met het rapport wordt gedaan, hangt af van het soort, en het doel van het AC proces. Zo kan men een oordeel vellen over de geschiktheid van de kandidaat en kunnen er verbeteringspunten aangereikt worden of kan er een ontwikkelingsplan worden opgesteld.

Een assessor kan een psycholoog zijn, maar vaak wordt deze taak ook door andere personen uitgevoerd. Lievens (1998) stelt wel voor om psychologen een sleutelrol te laten spelen. Uit verschillende onderzoeken blijkt dat zij deze rol beter zouden vervullen. Psychologen blijken meer dimensies te herkennen dan managers (Sagie en Mangezy, 1997). Verder kunnen zij ook beter interpersoonlijke vaardigheden in jobprestaties voorspellen, dan mensen die reeds ervaring hebben met de functie (Damitz, Manzey, Kleinmann, & Severin, 2003).

Het wordt aangeraden om assessoren een training te laten volgen om eventuele vertekeningen in de beoordeling te voorkomen. Er kunnen vier soorten trainingen worden onderscheiden die elk een andere vorm van vertekening trachten te reduceren (Woehr & Hufcutt, 1994). De eerste training is de 'rater-error-training'; deze dient om beoordelingseffecten zoals het halo-effect te reduceren. De tweede training is de

‘gedragsobservatie-training’; hier worden technieken aangeleerd om correcter te beoordelen. Daardoor tracht men de observatiefouten tot een minimum te herleiden. Een derde training noemt de ‘prestatie-dimensie-training’; in deze training leert men de verschillende prestatie-dimensies te onderscheiden. Men legt hier de nadruk op de definitie en de operationalisering van de dimensies. De vierde en laatste training wordt de ‘referentiekadertraining’ of ‘frame-of-reference-training’ genoemd; hier wordt verder gebouwd op de prestatie-dimensie-training. Bovenop het leren kennen van de dimensies, zorgt deze training ervoor dat de assessoren dezelfde standaarden meekrijgen om de prestaties van de kandidaten te beoordelen (Lievens, 1998).

Psychometrische Kenmerken

Interbeoordelaarsbetrouwbaarheid

Dit is de mate waarin twee of meer beoordelaars het eens zijn over hun beoordeling van gedrag. Om de beoordeling als valide te beschouwen, is het nodig dat beoordelaars het met elkaar eens kunnen zijn. Dit is zeker nodig voor ACs, omdat de data hier sterk afhankelijk zijn van elkaar (Reber, 2002). Er is nog niet veel onderzoek gedaan naar de interbeoordelaarsbetrouwbaarheid bij ACs, maar uit de schaarse onderzoeken blijkt wel dat er met dit psychometrisch kenmerk weinig problemen zijn (Lievens & Van Keer, 1999). Er zijn reeds enkele factoren bekend die de interbeoordelaarsbetrouwbaarheid kunnen verhogen in ACs.

Vooreerst blijkt dat hoe gemakkelijker de observeerbaarheid van de dimensies, hoe hoger de interbeoordelaarsbetrouwbaarheid (Lievens & Van Keer, 1999).

Ten tweede blijkt dat wanneer de assessoren een referentiekadertraining krijgen, de interbeoordelaarsbetrouwbaarheid significant stijgt (Jackson, Atkins, Fletcher, & Stillman, 2005).

Ten derde is de interbeoordelaarsbetrouwbaarheid ook hoger na de discussiefase van de assessoren (Lievens, 2001; Schmitt, 1977).

Ten slotte is de interbeoordelaarsbetrouwbaarheid hoger bij gestructureerde dan bij ongestructureerde interviews en naarmate ACs een hogere mate van standaardisatie hebben. (Conway, Jako, Goodman, 1995; Lievens, 2001).

Face Validiteit

Face validiteit gaat om de indruk die mensen hebben over een bepaalde methode of meetinstrument dat wordt gebruikt. In vergelijking met andere testen blijkt het AC proces een goede face validiteit te hebben. Assessee's gaan ervan uit dat het AC proces complexe attributen op een adequate manier kan meten. De assessee's aanzien deze methode als fair en daarom wordt ze met een positieve attitude onthaald (Fletcher & Kerslake, 1992; Thornton, 1992, 2006).

Inhoudsvaliditeit

De inhoudsvaliditeit wordt omschreven als de relatie tussen de test en het te meten construct. Dit betekent dat de gesimuleerde situaties overeenstemmen met de werkelijkheid en dat de te meten constructen in overeenstemming zijn met de gedragskenmerken die nodig zijn in de realiteit (Woehr & Arthur, 2003). Zoals vermeld is het aangewezen om voorafgaand aan het AC proces een jobanalyse uit te voeren. Een goed uitgevoerde jobanalyse is dan ook een bron van inhoudsvaliditeit. Verder zijn een goede training van de assessor, de vorm en uniformiteit van de instructies, de antwoordmogelijkheden, en het scoringssysteem ook facetten die voor een goede inhoudsvaliditeit kunnen zorgen (Sackett, 1987).

Criteriumvaliditeit

Criteriumvaliditeit of predictieve validiteit gaat over de voorspellende kracht van ACs. ACs blijken een vrij goede voorspeller te zijn (Damitz et al., 2003; Woehr & Arthur, 2003). Om deze vorm van validiteit te bekomen, neemt men de gemiddelde scores van wat men gemeten heeft in ACs en deze worden dan de predictor. De mate van succes in de job is het criterium. Deze kunnen gemeten worden aan de hand van verschillende indicatoren zoals de effectiviteit in een managementfunctie, de snelheid van promoties, het managementpotentieel, en de stijging van het salaris. Men berekent tussen predictor en criterium de correlatie en zo bekomt men de criteriumvaliditeit (Jansen & De Jongh, 1997). In een meta-analyse omtrent de criteriumvaliditeit bij ACs werd een validiteitscoëfficiënt gevonden van .37 (Gaugler, Rosenthal, Thornton, & Bentson, 1987). In een recentere meta-analyse van Arthur, Day, McNelly en Edens

(2003) werd een validiteitscoëfficiënt gevonden van .25 tot .39 voor de criteriumvaliditeit van ACs.

Constructvaliditeit

Constructvaliditeit of begripsvaliditeit wordt omschreven als: De studie van de relatie van de metingen met andere metingen van gelijke en verschillende constructen. Met andere woorden: De vraag of het instrument de constructen meet die het zou willen meten. De constructen zijn de gedragskenmerken die worden gemeten (Lievens & Van Keer, 1999). Twee belangrijke facetten van constructvaliditeit zijn convergente validiteit en discriminante validiteit. Convergente validiteit houdt in dat verschillende metingen van eenzelfde construct met elkaar overeenstemmen. Discriminante validiteit wil zeggen dat metingen van verschillende constructen door één instrument helemaal niet met elkaar in overeenstemming zijn (Lievens & Van Keer, 1999).

Paradox

Ondanks het feit dat er positieve resultaten gevonden werden voor de inhoudsvaliditeit en de criteriumvaliditeit zijn de resultaten niet altijd eenduidig met betrekking tot de constructvaliditeit. Dit noemt men de paradox (Woehr & Arthur, 2003). Aangezien het ontbreken van constructvaliditeit een probleem vormt voor ACs (Bycio, Alvares, & Hahn, 1987; Lievens, 2006; Sackett & Dreher, 1982; Schneider & Schmitt, 1992) wordt hieronder dieper ingegaan op de paradox.

In oefeningen van ACs zouden gelijkaardige constructen gerelateerd moeten zijn en niet-gelijkaardige constructen niet gerelateerd. Er wordt dus zowel een goede convergente als een goede discriminante validiteit verwacht. De convergente validiteit wordt in sommige studies aangetoond, al is het wel in beperkt mate. Voor de discriminante validiteit wordt er echter minder evidentie gevonden. Als er discriminante validiteit ontbreekt, gaan dimensies binnen éénzelfde oefening te hoog met elkaar correleren. Onderzoek toont aan dat er meer gelijkenissen bestaan onder dimensies binnen één oefening, dan onder die dimensies gescoord over de oefeningen heen (Henderson et al., 1995; Joyce, Thayer, & Pond, 1994; Lievens, 2006; Schneider & Schmitt, 1992). Dit zou dus willen zeggen dat de oefening, in plaats van de dimensies,

meer gewicht in de schaal legt. De assessee kan dan, los van zijn vaardigheden, beter scoren naargelang het soort oefening (Lievens, 1998; Schneider & Schmidt, 1992). Zoals besproken bij de eigenschappen van het AC proces zijn deze verschillende oefeningen wel nodig. Vooreerst zorgt deze verscheidenheid aan oefeningen voor meervoudige input over het gedrag van de assessee. Daarnaast vraagt een accurate testing van de jobinhoud om een verscheidenheid aan oefeningen waarbij de assessee geconfronteerd wordt met verschillende aspecten van de functie (International Task Force, 2009).

Naast de impact van de oefeningen kan ook de samenstelling van de constructen een invloed hebben op de discriminante validiteit. Zo zitten in één praktijksimulatie vaak verscheidene gelijkaardige dimensies die behoren tot hetzelfde construct (Howard, 1997).

Er werden reeds enkele aanpassingen van het design en de implementatie van het AC proces uitgetoetst. Deze aanpassingen hadden een beperkte doch positieve invloed op de constructvaliditeit (Woehr & Arthur, 2003).

De eerste aanpassing van het design heeft betrekking op het aantal dimensies die worden gemeten. Uit een meta-analyse van Woehr en Arthur (2003) blijkt dat de cognitieve capaciteiten van assessoren maar een beperkt aantal dimensies aankunnen. Verder stelt men ook vast dat, als het de bedoeling is om een groot aantal dimensies te meten, die worden opgesplitst in drie tot vijf subcategorieën. Daarom zou het beter zijn om het aantal dimensies te beperken tot drie (Lievens, 1998). Een daling van het aantal dimensies blijkt een gunstig effect te hebben op de convergente validiteit (Gaugler & Thornton, 1989; Lievens & Conway, 2001; Woehr & Arthur, 2003).

Ten tweede blijkt dat wanneer de assessoren de assessee's moeten beoordelen op conceptueel verschillende dimensies, dit de discriminante validiteit ten goede komt (Lievens, 1993; Lievens, Chasteen, Day, & Christiansen, 2006).

Ten derde kan de scoringsmethode ook een invloed hebben op de constructvaliditeit. Het gebruik van de cross-oefening methode, in vergelijking met de binnen-oefening methode, heeft een gunstig effect op zowel de convergente als op de discriminante validiteit (Woehr & Arthur, 2003).

Ten vierde kan ook het type assessor een invloed hebben op de constructvaliditeit (Lievens, 2002; Lievens & Conway, 2001; Woehr & Arthur, 2003). Psychologen zouden alle dimensies kunnen onderscheiden, terwijl managers de dimensies opsplitsen in twee subgroepen. Het gebruik van psychologen als assessoren kan dus een invloed hebben op de constructvaliditeit (Sagie & Mangezy, 1997). Uit onderzoek blijkt dat wanneer men psychologen inzet als assessor, dit een gunstig effect kan hebben op de constructvaliditeit (Woehr & Arthur, 2003).

Ten vijfde zou het gebruik van een gedragsmatige checklist, een gunstig effect kunnen hebben op de constructvaliditeit, al zijn de resultaten hieromtrent niet eenduidig (Donahue, Truxillo, Cornwell, & Gerrity, 1997; Reilly, Henry, & Smither, 1990). Het verschil in resultaten zou te wijten kunnen zijn aan de manier waarop de gedragsmatige checklists worden opgesteld. Deze verschillen systematisch over reeds uitgevoerde studies (Lievens, 1998).

TREK ACTIVATIE THEORIE

Interactionistische Visie

De trek activatie theorie (TAT) is een recente theorie met als focus de interactie tussen persoon en situatie (Tett & Burnett, 2003). Een persoon wordt gekenmerkt door persoonlijkheidstrekken, deze zijn relatief consistent binnen het individu. Het zijn de verschillende neigingen die individuen hebben om zich op een bepaalde manier te gedragen. Zo kunnen mensen een hoge mate van agressie, empathie of vreugde hebben. Deze neiging komt echter niet tot uiting in elke situatie. In welke situatie zal deze dan wel tot uiting komen? Hier tracht de TAT een verklaring voor te vinden. Deze theorie stelt dat: De uiting van een persoonlijkheidstrek een activering vereist door situationele cues die relevant zijn voor die trek (Tett & Guterman, 2000).

De activering van een persoonlijkheidstrek zal afhangen van drie variabelen: De trekrelevantie, de sterkte van de situatie, en de persoonlijkheid van de persoon. De trekrelevantie en de sterkte van de situatie representeren de twee kenmerken van situaties voor het concept trekactivatiepotentieel (trait activation potential, TAP; Tett & Burnett, 2003). Aan de hand van de analogie met een radio worden de twee concepten van elkaar onderscheiden. De trekrelevantie wordt dan gelijkgesteld met het kanaal waarop de radio staat. De sterkte van de situatie wordt vergeleken met het volume van de radio (Tett & Burnett, 2003).

De trekrelevantie is het kwalitatieve luik van de situatie. Het betreft de informatie over welke cues aanwezig zijn om het gedrag van de latente trek uit te lokken. De trekken van het vijf factoren model (VFM) worden het meest gebruikt (Bijlage 3). De reden hiervoor is dat deze duidelijke gedragsdomeinen hebben en dat ze categorieën weergeven waarin individuen sociaal gedrag classificeren (Lievens, De Fruyt, & Van Dam; 2001). Een situatie wordt als relevant gezien voor een trek indien het cues voorziet, voor de uitdrukking van gedrag, die relevant zijn voor die trek (Tett & Guterman, 2000). Als voorbeeld wordt een persoon met een hoge mate van vreugde genomen. Tijdens een begrafenis zijn er weinig relevante cues aanwezig voor vreugde.

Indien de mate van vreugde nagegaan wordt, doet men dat beter op een feestje dan bij een begrafenis. Op een feestje zullen wel cues aanwezig zijn om dit gedrag uit te lokken; de situatie is met andere woorden trekrelevant.

De sterkte van de situatie is het kwantitatieve luik. Het is een continuüm dat refereert naar de duidelijkheid omtrent de interpretatie van de situatie. Een sterke situatie zal individuele verschillen onderdrukken, omdat deze vraagt om een duidelijk gedrag met een duidelijk resultaat. Bij een zwakke situatie zullen individuele verschillen wel tot uiting komen, net omdat de interpretatie van de situatie niet duidelijk is. Met andere woorden een sterke situatie is niet ambigue, een zwakke situatie is ambigue. Een voorbeeld van een sterke situatie is een begrafenis, iedereen zal hier stil en ingetogen zijn. Een feestje is dan weer een voorbeeld van een zwakke situatie, hier worden individuele verschillen sneller waargenomen. Een situatie kan ook te sterk zijn, dan lukt ze verhoogde reacties uit bij iedereen en worden zo onderliggende trekken niet getoond. In het heetst van de strijd zullen de verschillen tussen soldaten in hun persoonlijkheidstrekken niet detecteerbaar zijn (Mischel, 1973).

Situaties zijn dan ook niet altijd sterk of zwak, dit zal enkel zo zijn voor bepaalde trekken. Hoe stipt een bepaald persoon ook is, hij zal wel eens 'stijlvol-te-laet' zijn op een feestje (Thornton & Mueller-Hanson, 2004).

De laatste variabele die een rol speelt bij de activering van een persoonlijkheidstrekk, is de persoonlijkheid van een persoon. Bepaalde personen zullen consistent hoger of lager scoren op bepaalde persoonlijkheidstrekken. Personen met een hoge score op vreugde zullen sneller een verhoogde respons of sensitiviteit vertonen dan personen met een gemiddelde tot lage score op vreugde. Bij de mensen die hoog scoren zal deze trek sneller tot uiting komen. Dit zelfs in situaties waar slechts zwakke cues worden aangeboden (Tett & Guterman, 2000).

Trek Activatie Theorie en Assessment Center

Tett en Guterman (2000) geven in hun onderzoek enkele toepassingen mee van de TAT, de toepassing op ACs is voor het huidige onderzoek van belang. Zoals eerder vermeld is een probleem met ACs het ontbreken van constructvaliditeit. De convergente en discriminante validiteit zijn voor verbetering vatbaar. De verschillende dimensies correleren onderling hoger in dezelfde oefening dan dat ze dat doen op verschillende oefeningen heen. De TAT kan hier een verklaring bieden (Lievens, 2009).

Tett en Gutermann (2000) stellen dat diverse oefeningen verschillende trekrelevante cues vertonen, waardoor die bepaalde trek niet in dezelfde mate tot uiting komt in de verschillende oefeningen. Dit strookt ook met de situationele specificiteit hypothese: Deze stelt dat elke oefening een bepaalde situatie is. Hierdoor worden verschillende aspecten van een bepaald gedrag losgemaakt, waardoor men zich telkens anders gedraagt in een bepaalde oefening. Convergente validiteit kan dan ook alleen maar verwacht worden ten opzichte van een trek, wanneer in verschillende oefeningen gelijkende trekrelevante cues aanwezig zijn (Schneider & Schmitt, 1992).

Haaland en Christiansen (2002) vinden in hun onderzoek bevestiging voor de belangrijke rol die TAP speelt voor de convergente validiteit. Zij brengen twee hypothesen naar voor: (1) Beoordelingen van oefeningen die een hoog TAP hebben, zullen sterkere relaties vertonen met scores op de relevante trekken dan beoordelingen van oefeningen met een laag TAP. (2) De convergentie van beoordelingen van oefeningen met een hoog TAP zal sterker zijn dan die van de beoordelingen van de oefeningen met een laag TAP. Beide hypothesen zijn bevestigd. Hieruit kan men besluiten dat de convergente validiteit hoger is voor oefeningen van een bepaalde trek, wanneer deze oefeningen trekrelevante cues vertonen.

Het cross-situationeel gedrag zal dus enkel consistent overkomen, indien het gedrag gerelateerd aan dezelfde persoonlijkheidstrekk, geobserveerd kan worden in elke oefening. Sterker zelfs nog, het geobserveerde gedrag hoeft niet identiek te zijn, zolang het maar gerelateerd is aan dezelfde persoonlijkheidstrekk (Lievens et al., 2006).

In de studie van Lievens et al. (2006) is de invloed van de TAT nagegaan op zowel de convergente als discriminante validiteit. De volgende hypothesen zijn naar voor gebracht: (1) De convergentie van beoordelingen van een gegeven dimensie gelinkt aan een specifieke VFM trek, zal sterker zijn bij oefeningen die deze specifieke VFM trek activeren, dan bij oefeningen die deze specifieke VFM trek niet activeren. (2) De discriminatie tussen beoordelingen binnen de oefening zal beter zijn voor dimensies waar niet één onderliggende VFM trek aan de basis ligt, dan wanneer dat wel het geval is. Voor beide hypothesen is bevestiging gevonden.

Verder vonden zij ook dat de discriminante validiteit vooral hoger ligt in oefeningen die enige vorm van sociale interactie vereisten. Daarnaast werd ook een duidelijke effect gevonden van de activatie van de trek voor de trekken extraversie en consciëntieusheid. Dit kan men verklaren door de observeerbaarheid van deze trekken. Vorig onderzoek toonde aan dat de trekken extraversie en consciëntieusheid makkelijker observeerbaar zijn dan emotionele stabiliteit, altruïsme en openheid (Lievens et al., 2006).

Het descriptief gebruiken van de TAT kan dus in verband worden gebracht met de convergente en discriminante validiteit van scores op dimensies in ACs. Oefeningen van ACs zijn ontworpen om een brede waaier van gedragingen te observeren en zo te gaan differentiëren in de prestaties van de assessee (Lievens et al., 2006). Als basis van de scores op de verschillende dimensies worden de observaties van de gedragingen gebruikt (Tett & Burnett, 2003). Door het feit dat er verschillende trekken geobserveerd worden, is het van belang dat de oefeningen verschillen in de cues die worden gepresenteerd. Dit hangt eveneens af van welke trekken relevant zijn voor een bepaalde oefening. Afhankelijk van de oefening en de cues zal er dus ook een verschil optreden in TAP. De kans om de verschillen in trekgerelateerd gedrag te observeren zal dus afhankelijk zijn van zowel de relevantie als de sterkte van de situatie. Wat op zijn beurt samenhangt met de convergente en discriminante validiteit (Lievens et al., 2006).

Trek Activatie Theorie en Oefeninginstructies

De TAT heeft directe implicaties voor het design van ACs (Haaland & Christiansen, 2002; Lievens, 2009; Lievens et al., 2006). Om de assessee de mogelijkheid te geven een gedrag te stellen dat relevant is voor een bepaalde trek, moet de AC oefening voldoen aan twee voorwaarden: (1) De situatie mag niet te sterk zijn en (2) De situatie moet toelaten om de onderliggende trek te tonen. Met andere woorden de mogelijkheid om trekrelevante gedrag uit te lokken moet verhogen, de TAP dus. De instructies die de kandidaten aan het begin van de oefeningen krijgen, zouden hier een rol in kunnen spelen. Deze oefeninginstructies kunnen cues bevatten om gedrag uit te lokken dat relevant is voor de vooropgestelde dimensies (Thornton & Mueller-Hanson, 2004).

Kleinmann, Kuptsch en Köller (1996) bestudeerden reeds de impact van het vooraf informeren van de assessee over welke gedragingen relevant zijn voor de verschillende dimensies. Zij gingen ervan uit dat de assessee op deze manier hun gedrag meer zouden richten naar de dimensies en zo duidelijker en consistentere dimensiegerelateerd gedrag zouden vertonen. Dit kon er op zijn beurt voor zorgen dat de assessoren beter konden differentiëren tussen de dimensies en een betere beoordeling konden maken. Voor beide veronderstellingen werd empirische evidentie gevonden en was er een invloed op zowel de convergente als discriminante validiteit.

Het nadeel van deze studie was echter dat de dimensies zeer duidelijk werden weergegeven, wat voor een verhoogde transparantie zorgde. Daardoor kon het zijn dat de gedragingen die de assessee stelden niet representatief waren, dat ze zich als het ware conform de dimensie gingen gedragen. Dit komt de criteriumvaliditeit van ACs niet ten goede (Lievens, 1993). Deze studie zal met het negatieve effect van de transparantie trachten rekening te houden door het opstellen van gestandaardiseerde oefeninginstructies.

In de huidige praktijk krijgen assessee eerder vage instructies zoals: Hieronder wordt een probleem geschetst, u krijgt een half uur de tijd om dit probleem op te lossen. Deze instructies zijn niet bedoeld om specifiek trekrelevant gedrag uit te lokken. Dit

onderzoek zal zich dan ook focussen op het geven van gestandaardiseerde oefeninginstructies die trekrelevant gedrag gaan uitlokken. Op deze manier wordt de TAT prescriptief toegepast. Een voorbeeld rond probleemanalyse en -oplossing ter verduidelijking: Hieronder wordt een probleem geschetst, u krijgt een half uur de tijd om een presentatie over dit probleem voor te bereiden. In tegenstelling tot de vorige oefeninginstructie wordt hier gepoogd om een situatie te creëren die het gedrag probleemanalyse en -oplossing uitlokt. Het feit dat de persoon een presentatie moet maken over het probleem kan een trekrelevante cue zijn voor de dimensie probleemanalyse en -oplossing. Of het gedrag effectief gesteld wordt, zal wel nog steeds van de assessee afhangen.

Gestandaardiseerde oefeninginstructies worden gedefinieerd als:

Gepredetermineerde instructies die voorafgaand aan de oefening worden gegeven en over alle kandidaten consistent worden gehouden met als doel het dimensierelevante gedrag uit te lokken. Het is de bedoeling om aan de hand van gestandaardiseerde oefeninginstructies de AC oefeningen te voorzien van trekrelevante cues. Er wordt verondersteld dat de AC oefeningen met gestandaardiseerde oefeninginstructies hoger zullen correleren met de relevante trekken dan de AC oefeningen zonder de gestandaardiseerde oefeninginstructies. Hierdoor zou het TAP in een AC oefening met gestandaardiseerde oefeninginstructies hoger kunnen zijn dan in een AC oefening zonder gestandaardiseerde oefeninginstructies.

In het huidige onderzoek zullen de dimensies verschillende onderliggende trekken hebben, zodat er meerder gedragingen zijn waar de assessoren hun oordeel op kunnen baseren (Brannick, 2008). In de conditie met gestandaardiseerde oefeninginstructies wordt eveneens een hogere TAP verwacht, wat voor een betere differentiatie tussen de verschillende dimensies zou kunnen zorgen. Dit zou een invloed kunnen hebben op de constructvaliditeit. Op basis van deze veronderstelling worden met betrekking tot de convergente en discriminante validiteit de volgende hypothesen voorop gesteld:

Hypothese 1a: De AC oefening met gestandaardiseerde oefeninginstructies zal een hogere convergente validiteit hebben dan een AC oefening zonder gestandaardiseerde oefeninginstructies.

Hypothese 1b: De AC oefening met gestandaardiseerde oefeninginstructies zal een hogere discriminante validiteit hebben dan een AC oefening zonder gestandaardiseerde oefeninginstructies.

Vervolgens zal ook de interbeoordelaarsbetrouwbaarheid getest worden.

Er is al aangegeven dat hoe gemakkelijker het is om de trekken te observeren, hoe hoger de interbeoordelaarsbetrouwbaarheid (Lievens & Van Keer, 1999). Als er in de AC oefening gebruik gemaakt wordt van gestandaardiseerde oefeninginstructies kan er mogelijk een hogere interbeoordelaarsbetrouwbaarheid geconstateerd worden. Dit zou verklaard kunnen worden doordat een stijging van de TAP een betere observeerbaarheid van de trekken kan veroorzaken. Indien de assessoren meer gedragingen hebben die ze kunnen observeren kan, het mogelijk zijn dat ze beter overeenkomen in hun oordeel. Zo is de tweede en laatste hypothese als volgt:

Hypothese 2: In een AC oefening met gestandaardiseerde oefeninginstructies zal de interbeoordelaarsbetrouwbaarheid hoger liggen dan in een AC oefening zonder gestandaardiseerde oefeninginstructies.

METHODE

Steekproef

De data voor deze studie werd verzameld via een AC proces, georganiseerd door de Vakgroep Personeelsbeleid, Arbeids - en Organisationspsychologie aan de Universiteit Gent.

In totaal namen 236 kandidaten deel aan dit AC proces, allen laatstejaarsstudenten uit verschillende afstudeerrichtingen van de Associatie Universiteit Gent.

Kandidaten werden uitgenodigd via e-mail voor dit AC proces dat breder gekaderd werd in een sollicitatietraining. De sollicitatietraining omvatte twee luiken: Een voorbereidingssessie op psychologische testen enerzijds en een voorbereidingssessie op ACs anderzijds. Studenten konden op vrijwillige basis deelnemen aan deze sessies en werden ad random toegewezen aan een conditie.

Voor deze studie werden in totaal 177 studenten Bedrijfspsychologie en Personeelsbeleid (72.9% vrouwen; gemiddelde leeftijd = 20.4, $SD = 1.6$) ingeschakeld als assessoren voor de verschillende AC oefeningen. Er werden van 51 (62.2% vrouwen, gemiddelde leeftijd = 23 jaar, $SD = 1.5$ jaar) AC kandidaten video-opnames gebruikt om aan deze assessoren te tonen.

Procedure

Opzet

De volledige sollicitatietraining vond plaats tussen juni 2007 en februari 2008. Eén volledige sessie duurde ongeveer 6 uur en iedere dag kende hetzelfde verloop. Eerst was er een verwelkoming met een korte uitleg over wat de kandidaten te wachten stond. Om de simulatie zo reëel mogelijk te laten verlopen werd de kandidaten gevraagd om een standaard cv-formulier en enkele vragenlijsten in te vullen. Ten slotte namen ze deel aan twee AC oefeningen.

De twee AC oefeningen waren gericht op een managementfunctie voor starters. Er werden twee specifieke praktijksimulaties ontwikkeld die zich richten op kandidaten

die een startersmanagementfunctie ambieerden. Concreet ging het hier om een werkgesprek tussen een medewerker versus zijn chef en een presentatieoefening. Deze praktijksimulaties worden straks verder besproken.

De sollicitatietraining vond plaats in de Faculteit Psychologie en Pedagogische Wetenschappen aan de Universiteit Gent. De twee AC oefeningen werden gesimuleerd in een skillslab met twee one-way-mirrors. Om de rollenspelers tijdens de praktijksimulatie uit verschillende hoeken te kunnen observeren werden er in het skillslab camera's geplaatst op drie verschillende plaatsen. Hiervoor tekenden de kandidaten vooraf een formulier waarin ze verklaarden akkoord te gaan met het op video vastleggen van hun prestaties.

Na afloop van de sollicitatietraining ontving de kandidaat per mail een feedbackrapport. Verder kregen ze ook een boekje met algemene uitleg, sollicitatietips, en een uitnodiging tot cv-screening toegestuurd.

Dimensies

De dimensies die tijdens de oefeningen werden gemeten zijn '*interpersoonlijke sensitiviteit/empatisch handelen*', gedefinieerd als de mate waarin een individu handelt rekening houdend met de gevoelens en behoeften van andere mensen; '*plannen en organiseren*' gedefinieerd als de mate waarin een individu systematisch zijn/haar werk en middelen regelt alsook dat van anderen voor een efficiënte taakverwezenlijking en de mate waarin een individu anticipeert op de toekomst en er zich op voorbereidt; '*probleemanalyse en -oplossing*' gedefinieerd als de mate waarin een individu informatie verzamelt, relevante technische en professionele informatie begrijpt, effectief data en informatie analyseert, valabele mogelijkheden, ideeën en oplossingen genereert, verdedigbare plannen van actie voor problemen en situaties selecteert, beschikbare middelen op nieuwe manieren gebruikt, en verbeeldingrijke oplossingen genereert en herkent; '*stresstolerantie/omgaan met stress*' gedefinieerd als de mate waarin een individu effectiviteit behoudt in verschillende situaties onder gevarieerde mate van druk, weerstand, en teleurstelling.

Deze vier dimensies werden geselecteerd op basis van de meta-analyse van Arthur et al. (2003) en bleken relevant te zijn voor een startersmanagementfunctie.

Assessorentraining

De assessoren kregen een halve dag training door een gecertificeerd assessor met een master diploma in de Bedrijfspsychologie en Personeelsbeleid.

Het trainingsprogramma bestond uit 3 delen, (a) een introductie over de basis van ACs, (b) informatie over de inhoud van de oefeningen en de vier dimensies, en (c) een workshop over het observatie- en beoordelingsproces dat bestond uit theorie, een oefening, en feedback.

Vooreerst was er een introductie (2.5 uur) waarin ACs gedefinieerd en geplaatst werden binnen de context van Human Resources Management en personeelsselectie. De trainer gaf ook meer informatie over de verschillende componenten van het AC proces, het doel, de geschiedenis, en de huidige toepassingen van ACs.

Tijdens het tweede deel van de training (.5 uur) werd de assessoren de kennis en het begrip van de oefeningen en de vier dimensies bijgebracht. Hier kregen de assessoren achtergrondinformatie over de oefeningen en de definities van de dimensies. Verder kregen ze uitleg over hun concrete taak: Het observeren van de vier dimensies bij kandidaten tijdens een praktijksimulatie op basis van de video-opnames die werden gemaakt van de kandidaten.

In het laatste deel van de training (1 uur) werd de assessoren aangeleerd om het gedrag van kandidaten te observeren. De trainer leerde de assessoren om tijdens het observeren gedragsgerelateerde beschrijvingen te maken in plaats van niet-gedragsgerelateerde interpretaties. Zo werd gepoogd om de assessoren goede observaties neer te leren schrijven (Gaugler & Thornton, 1989). Daarna werd ze aangeleerd om hun gemaakte beschrijvingen te classificeren onder de juiste dimensie. Na de theorie konden de assessoren aan de hand van de eerder opgenomen video's van de assessee hun observatievaardigheden inoefenen. Nadat enkele video-opnames bekeken en beoordeeld waren, bracht de trainer een discussie op gang over de observaties. Discrepancies werden besproken en opgelost en de trainer gaf feedback over de prestaties van de assessoren als observatoren.

AC Oefeningen

Werkgesprek

In dit rollenspel neemt de kandidaat de rol op zich van chef en de rollenspeler doet zich voor als de medewerker. De chef roept de medewerker bij zich omdat er de laatste tijd enkele problemen zijn gemeld. Er zijn klachten gekomen van een belangrijke klant en van collega's. Deze medewerker wil dit jaar graag promotie maken en heeft zich ingeschreven om een cursus te volgen. Het is al het tweede jaar op rij dat deze medewerker wil deelnemen aan deze cursus.

In de ideale situatie analyseert de kandidaat alle klachten zorgvuldig om zo te weten te komen wat er werkelijk aan de hand is. Daarnaast stelt hij een oplossing en/of toekomstgericht plan voor. Verder zou het gesprek zo stressloos mogelijk moeten verlopen en is het goed als de kandidaat de zaken op een interpersoonlijk sensitieve manier aanpakt.

De kandidaat krijgt een voorbereidingstijd van 15 minuten en daarna 15 minuten om het gesprek te voeren.

Presentatie

In de presentatieoefening is de kandidaat een consultant en de rollenspeler is het kaderlid van een park, Mosaic Park. Het park heeft geld gekregen van een overleden baron en wil dit gebruiken om het park te vernieuwen. Maar Mosaic Park kan enkel gebruik maken van het geld als ze rekening houden met de wensen van de baron. Eén wens was, om samen met een consultant, de vernieuwing van het park uit te werken en op te starten. De kandidaat is één van de potentiële consultants en moet zijn/haar ideeën over de vernieuwing presenteren aan het kaderlid van Mosaic Park.

De kandidaat krijgt opnieuw een voorbereidingstijd van 15 minuten en daarna 15 minuten om het gesprek te voeren.

Midden februari werd de naam Mosaic Park veranderd in Flamingo Yards. Dit om eventuele leereffecten tegen te gaan.

Gestandaardiseerde Oefeninginstructies

Experimenteel Design

Voor deze studie werden twee condities opgezet. In de eerst conditie kregen de kandidaten voorafgaand aan de AC oefening gewone oefeninginstructies. In de tweede conditie kregen de kandidaten voorafgaand aan de AC oefening gestandaardiseerde oefeninginstructies.

Opstellen van de Gestandaardiseerde Oefeninginstructies

In de conditie zonder gestandaardiseerde oefeninginstructies kreeg de kandidaat de AC oefening zoals ze oorspronkelijk was opgesteld. De oefeninginstructies waren algemeen van aard. Eerst kreeg de kandidaat informatie over de achtergrond van zijn rol. Daarna werd kort het concrete probleem uit de doeken gedaan. Ten slotte kreeg de kandidaat zijn opdracht: Een presentatie geven of een werkgesprek voeren.

In de conditie met gestandaardiseerde oefeninginstructies kreeg de kandidaat dezelfde instructies maar werden er vier gedetailleerde gestandaardiseerde oefeninginstructies aan toegevoegd. Er werd voor iedere oefening per dimensie een gestandaardiseerde oefeninginstructie ontwikkeld. Bij het opstellen van deze gestandaardiseerde oefeninginstructies werd de TAT als uitgangspunt gebruikt. Deze theorie stelt dat een uiting van een persoonlijkheidstrekk een activering vereist door een situationele cue die relevant is voor die trek (Tett & Guterman, 2000). Met deze gestandaardiseerde oefeninginstructies werd dus getracht om over verschillende oefeningen heen, gelijkaardige trekrelevante cues te geven aan de kandidaten. Om deze acht gestandaardiseerde oefeninginstructies op te stellen werd gesteund op de definities van Arthur et al. (2003). Voor de dimensie probleemanalyse en –oplossing in de presentatie werd de volgende gestandaardiseerde oefeninginstructie ontwikkeld: *Maak een analyse van deze tekst en haal er de kern uit. Tracht ideeën te bedenken en maak een overzicht van uw voorstel, dat u nadien zal presenteren.*

Een volledig overzicht van de gestandaardiseerde oefeninginstructies is terug te vinden in Bijlage 4.

Manipulation Check

Om na te gaan of de gestandaardiseerde oefeninginstructies relevant waren voor de respectievelijke dimensies werden 15 master studenten Bedrijfspsychologie en Personeelsbeleid (53% vrouwen; gemiddelde leeftijd = 25.4, $SD = 3.7$) gevraagd om onafhankelijk de gestandaardiseerde oefeninginstructies te hervers talen, dit door ze toe te wijzen aan één van de vier dimensies. De studenten wisten niets af van het doel van de studie, maar hadden wel de nodige achtergrond om deze oefening uit te voeren. Om deze oefening te maken kregen ze de definities van de dimensies.

Naast de vier gestandaardiseerde oefeninginstructies die werden ontwikkeld voor deze studie, werden ook vier extra instructies voor andere dimensies toegevoegd, die dienst deden als afleiders (Bijlage 5). Voor iedere oefeninginstructie was er een overeenkomst van minstens 90%. Deze gestandaardiseerde oefeninginstructies werden dan ook als relevant gezien voor de vier dimensies.

Design Overwegingen

Om de constructvaliditeit van ACs te verhogen zijn in het verleden al enkele aanbevelingen gedaan. In het opzet van deze studie werd met enkele van deze aanbevelingen rekening gehouden.

Ten eerste, kregen de assessoren een training om zo variaties in scores te minimaliseren (Woehr & Hufcutt, 1994).

Ten tweede, beoordeelden telkens ten minste twee afzonderlijke assessoren, onafhankelijk van elkaar, de assessee tijdens de AC oefeningen. Dit om te vermijden dat er een verband kon zijn tussen de assessoren en de oefeningen, het zogenaamde assessoreneffect (Lievens, 2006; Thornton, 2006).

Data-analyse

De gemiddelden, standaardafwijkingen, en correlaties van de variabelen uit deze studie werden berekend aan de hand van Pearson correlaties.

Voor de constructvaliditeit en de interbeoordelaarsbetrouwbaarheid werd een generaliseerbaarheidsanalyse uitgevoerd. Dit in tegenstelling tot voorgaande studies waar men meestal de multitrek – multimethode (MTMM) matrix opstelde (Lievens, 1993, 2009).

Deze methode houdt echter enkel rekening met twee bronnen van variantie: Oefeningen en dimensies. Al kunnen in ACs nog andere zaken een invloed hebben op het meetproces. Naast de oefeningen en de dimensies, kunnen bijvoorbeeld ook de assessoren het meetproces beïnvloeden. Recent wordt er dan ook aangeraden om gebruik te maken van een generaliseerbaarheidsanalyse als alternatieve benadering. Om zo de verschillende facetten die het meetproces beïnvloeden verder te analyseren (Lievens, 2009).

Een generaliseerbaarheidsanalyse zet een geobserveerde variantie om in twee variantie componenten (VCn). Een variantie component die betrekking heeft op de daadwerkelijke verschillen in score (werkelijke variantie) en een error component die het meetproces beïnvloedt (error variantie) (Kane, 1982).

In tegenstelling tot de klassieke test theorie beschouwt de generaliseerbaarheidsanalyse deze variantie dus als bestaande uit meerdere factoren. Zo laat ze toe om tegelijkertijd verschillende bronnen van variantie in metingen te bekijken. De generaliseerbaarheidsanalyse werd hier dus gebruikt om een genuanceerder beeld te krijgen van de verschillende bronnen van variantie binnen het meetproces van ACs.

Tabel 1.

Beschrijvende statistieken en correlaties tussen de variabelen van deze studie (N=51)

Variabelen	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1.Geslacht	.623	.489											
2.Leeftijd	23	1.569	-.025										
3.Gestandaardiseerde oefeninginstructies	.509	.505	-.063	-.049									
4.Probleemanalyse en-oplossing W	3.278	.743	.130	-.015	.065								
5.Plannen en organiseren W	3.183	.825	.056	.079	-.002	.540**							
6.Interpersoonlijke sensitiviteit W	3.276	.684	-.011	-.006	.054	.646**	.524**						
7.Stresstolerantie W	3.509	.758	.146	.106	.059	.430**	.572**	.423**					
8.Probleemanalyse en-oplossing P	3.428	.715	-.086	-.076	.097	.340*	.495**	.405**	.627**				
9.Plannen en organiseren P	3.014	.700	-.126	-.037	.130	.288*	.466**	.286*	.411**	.676**	.328*		
10.Interpersoonlijke sensitiviteit P	3.184	.618	-.161	.106	.069	.265	.273	.359**	.277*	.290*	.294*		
11. Stresstolerantie P	3.235	.717	.041	.103	.121	.243	.320*	.192	.612**	.461*	.328*	.484*	

Noot. Ondanks het feit dat er eigenlijk geen gemiddelden, standaarddeviaties en correlaties mogen berekend worden van categorische variabelen, werd dit hier toch gedaan als eerste indicatie voor de interpretatie. Hiervoor werden de categorische variabelen dummygecodeerd als volgt: ‘gestandaardiseerde oefeninginstructies’: 0 = zonder, 1 = met; ‘geslacht’: 0 = man, 1 = vrouw.

De scores die worden weergegeven zijn de gemiddelde scores van de assessoren op de verschillende dimensies.

W = Werkgesprek; P = Presentatie

* $p < 0.05$ ** $p < 0.01$

RESULTATEN

Descriptieve Resultaten

In Tabel 1 zijn de gemiddelden, standaardafwijkingen, en correlaties van de verschillende variabelen die in deze studie werden gebruikt terug te vinden. Deze correlaties werden berekend aan de hand van Pearson correlaties met het statistische programma SPSS 16.

De controlevariabelen, geslacht en leeftijd, correleren met geen enkele dimensie significant.

Werkgesprek

Binnen het werkgesprek correleren alle dimensies significant met elkaar. De hoogste correlatie is tussen *interpersoonlijke sensitiviteit* en *probleemanalyse en -oplossing* (.646 met $p < .05$). De laagste correlatie is terug te vinden tussen *stresstolerantie* en *interpersoonlijke sensitiviteit* (.423 met $p < .05$).

Presentatie

Binnen de presentatie correleren alle dimensies eveneens significant met elkaar. De hoogste correlatie is tussen *plannen en organiseren* en *probleemanalyse en -oplossing* (.676 met $p < .05$). De laagste correlatie is terug te vinden tussen *interpersoonlijke sensitiviteit* en *probleemanalyse en -oplossing* (.290 met $p < .05$).

Normaalverdelingen

De grafieken van de gemiddelde scores op de verschillende dimensies per AC oefening zijn terug te vinden in Bijlage 6. Deze lijken allemaal een normaalverdeling te volgen.

Constructvaliditeit

Zoals reeds vermeld werd voor de constructvaliditeit en de interbeoordelaarsbetrouwbaarheid een generaliseerbaarheidsanalyse uitgevoerd. Deze analyse werd uitgevoerd met het statistische programma SPSS 16.

De VCn die worden verkregen door de generaliseerbaarheidsanalyse weerspiegelen voor ieder facet de bijdrage aan de totale variantie. De geschatte VCn hangen af van de omvang van de metingen. Daarom is het van belang om deze componenten te interpreteren door middel van hun relatieve belangrijkheid. Om dit te bekomen werd de procentuele bijdrage van iedere VC berekend. De procentueel verklaarde variantie per facet wordt bekomen door de som te nemen van alle VCn en dan iedere VC afzonderlijk te delen door deze som.

De resultaten van deze analyse zijn terug te vinden in Tabel 2. De generaliseerbaarheidsanalyse die in deze studie werd uitgevoerd heeft vier facetten: kandidaat (K), assessor in kandidaat genest (A(K)), dimensie (D) en oefening (O). Dit design is een mix van gecrost en geneste factoren. De score van de kandidaten was het meetobject van deze analyse.

Met betrekking tot de constructvaliditeit worden achtereenvolgens de convergente validiteit en de discriminante validiteit besproken.

Tabel 2.

Generaliseerbaarheidsanalyse (N=51)

Effect	ZGO		MGO	
	VC	VC%	VC	VC%
K(andidaat)	.180	12.1	.043	4.7
A(assessor(Kandidaat))	.317	21.3	.175	18.9
D(imensie)	.001	.1	.007	.7
O(efening)	0	0	0	0
KD	.056	3.8	.060	6.4
KO	.082	5.5	.043	4.6
DO	.006	.4	.010	1.1
KDO	.120	8.1	0	0
KA(K)DO	.728	48.9	.589	63.6

Noot. ZGO = Zonder Gestandaardiseerde Oefeninginstructies; MGO: Met

Gestandaardiseerde Oefeninginstructies; VC = Variantie Component

Convergente Validiteit

Voor de convergente validiteit werd evidentie afgeleid uit de VC die geassocieerd is met de interactie van Kandidaat X Oefening (KO) enerzijds en de VC die geassocieerd is met de interactie van Kandidaat X Dimensie X Oefening (KDO) anderzijds (Kane, 1982). Een lage waarde van de Kandidaat X Oefening (KO) VC veronderstelt een invariantie van de scores die de kandidaten krijgen doorheen de oefeningen. Een lage waarde van de Kandidaat X Dimensie X Oefening (KDO) VC veronderstelt dat er invariantie is van de scores die kandidaten krijgen op de verschillende dimensies doorheen de oefeningen. Anders gezegd uit deze twee VCn kan worden afgeleid of verschillende scores van kandidaten op dezelfde dimensie met elkaar overeenstemmen en dit over verschillende oefeningen heen.

In de conditie zonder gestandaardiseerde oefeninginstructies bedraagt de waarde van de procentuele VC Kandidaat X Oefening (KO) 5.5 procent van de totale verklaarde variantie, de procentuele VC Kandidaat X Dimensie X Oefening (KDO) bedraagt 8.1 procent van de totale verklaarde variantie. In de conditie met gestandaardiseerde oefeninginstructies bedraagt de waarde van de procentuele VC Kandidaat X Oefening (KO) 4.6 procent van de totale verklaarde variantie, de procentuele VC Kandidaat X Dimensie X Oefening (KDO) is verwaarloosbaar klein.

De waarden voor de VCn zijn in de conditie met gestandaardiseerde oefeninginstructies dus lager dan in de conditie zonder gestandaardiseerde oefeninginstructies. Er kan dus besloten worden dat in de conditie met gestandaardiseerde oefeninginstructies de convergente validiteit hoger is dan in de conditie zonder gestandaardiseerde oefeninginstructies. De scores die de kandidaten krijgen op de verschillende dimensies doorheen de verschillende oefeningen vertonen dus minder variaties in de conditie met gestandaardiseerde oefeninginstructies dan zonder gestandaardiseerde oefeninginstructies.

Discriminante Validiteit

Evidentie voor discriminante validiteit kan worden afgeleid uit de VC die geassocieerd is met de interactie van Kandidaat X Dimensie (KD) (Kane, 1982). Een hoge waarde van deze VC geeft aan dat er substantiële verschillen zijn in de scores van de kandidaten op de verschillende dimensies. Met andere woorden, de score van een kandidaat op een bepaalde dimensie verschilt van de score van de kandidaat op een andere dimensie.

In de conditie zonder gestandaardiseerde oefeninginstructies bedraagt de procentuele VC Kandidaat X Dimensie (KD) 3.8 procent van de totale verklaarde variantie. In de conditie met gestandaardiseerde oefeninginstructies bedraagt de waarde van de procentuele VC Kandidaat X Dimensie (KD) 6.4 procent van de totale verklaarde variantie.

De waarde van de VC in de conditie met gestandaardiseerde oefeninginstructies is dus hoger dan in de conditie zonder gestandaardiseerde oefeninginstructies. Daaruit kan besloten worden dat in de conditie met gestandaardiseerde oefeninginstructies de discriminante validiteit hoger is dan in de conditie zonder gestandaardiseerde oefeninginstructies. Hier kan worden verondersteld dat in de conditie met gestandaardiseerde oefeninginstructies meer gedifferentieerd wordt tussen dimensies dan in de conditie zonder gestandaardiseerde oefeninginstructies.

Interbeoordelaarsbetrouwbaarheid

Voor de interbeoordelaarsbetrouwbaarheid werden generaliseerbaarheidscoëfficiënten (ρ^2) berekend. Dit is een intra-klas-correlatiecoëfficiënt gelijkaardig met het klassieke betrouwbaarheidscoëfficiënt. Het wordt gedefinieerd als de verhouding tussen de universele score variantie en de verwachte geobserveerde score variantie (Brennan, 1992).

In deze studie reflecteert het generaliseerbaarheidscoëfficiënt een generaliseerbaarheid van de scores van kandidaten over oefeningen, assessoren en dimensies heen. Een lage waarde van het generaliseerbaarheidscoëfficiënt weerspiegelt een lage interbeoordelaarsbetrouwbaarheid.

In de conditie zonder gestandaardiseerde oefeninginstructies is het generaliseerbaarheidscoëfficiënt .61.

In de conditie met gestandaardiseerde oefeninginstructies is het generaliseerbaarheidscoëfficiënt .45.

De interbeoordelaarsbetrouwbaarheid is dus lager in de conditie met gestandaardiseerde oefeninginstructies dan in de conditie zonder de gestandaardiseerde oefeninginstructies. In de conditie zonder gestandaardiseerde oefeninginstructies kunnen de score van de kandidaten over oefeningen, assessoren en dimensies heen meer gegeneraliseerd worden dan in de conditie met gestandaardiseerde oefeninginstructies.

DISCUSSIE

Hieronder wordt dieper ingegaan op de gevonden resultaten van deze studie en worden mogelijke verklaringen aangereikt. Vervolgens worden enkele methodologische reflecties besproken. Daarna volgt een bespreking van theoretische en praktische implicaties, evenals mogelijkheden voor toekomstig onderzoek. Er wordt afgesloten met een algemene conclusie van deze studie.

Bespreking Resultaten

Het opzet van deze studie was om na te gaan of gestandaardiseerde oefeninginstructies een gunstige bijdrage kunnen leveren tot de constructvaliditeit van ACs. Dit omdat er in het verleden vaak tegenstrijdige resultaten werden gevonden met betrekking tot de constructvaliditeit, de zogenaamde paradox. De gestandaardiseerde oefeninginstructies werden ontwikkeld op basis van de TAT van Tett en Guterman (2000). Deze theorie werd hierdoor prescriptief in plaats van descriptief gebruikt. Daarnaast werd ook de invloed van deze gestandaardiseerde oefeninginstructies op de interbeoordelaarsbetrouwbaarheid nagegaan.

Ten eerste geven de resultaten van de huidige studie aan dat gestandaardiseerde oefeninginstructies een gunstige bijdrage kunnen leveren aan de constructvaliditeit van ACs. Dit zowel met betrekking tot de convergente als discriminante validiteit. Ten tweede blijkt dat het gebruik van gestandaardiseerde oefeninginstructies eerder een negatieve invloed heeft op de interbeoordelaarsbetrouwbaarheid van ACs.

Constructvaliditeit

De eerste hypothese die voorop werd gesteld had betrekking op de twee componenten van de constructvaliditeit.

Convergente validiteit Op basis van de TAT van Tett en Guterman (2000) ging deze studie ervan uit dat AC oefeningen met gestandaardiseerde oefeninginstructies een hogere convergente validiteit zouden hebben dan AC oefeningen zonder gestandaardiseerde oefeninginstructies (Hypothese 1a). Voor deze hypothese werd

evidentie gevonden. In de conditie waar gestandaardiseerde oefeninginstructies werden gebruikt was er een hogere convergente validiteit. Verschillende metingen van dezelfde dimensie stemmen dus overeen.

Dit positief effect wordt waarschijnlijk bekomen door een verhoging van de TAP (Tett & Burnett, 2003). De TAP betreft de mogelijkheid van een situatie om het trekrelevante gedrag uit te lokken. Het omvat een kwalitatief luik, de trekrelevantie, en een kwantitatief luik, de sterkte van de situatie.

Een situatie is trekrelevant als er cues worden aangeboden, voor de uitdrukking van het gedrag, die relevant zijn voor die trek (Tett & Gutermann, 2000). Gestandaardiseerde oefeninginstructies fungeren in een AC oefening dus mogelijk als trekrelevante cues. Zo kan de trekrelevantie van de situatie verhogen en vervolgens de kans dat de assessee het gedrag stelt verhogen.

Om over verschillende AC oefeningen heen convergente validiteit te bekomen, is het van belang dat er gelijkende trekrelevante cues aanwezig zijn (Schneider & Schmitt, 1992). De gestandaardiseerde oefeninginstructies waren voor zowel het rollenspel als voor de presentatie gebaseerd op dezelfde definities van Arthur et al. (2003). Deze gelijkende gestandaardiseerde oefeninginstructies werden aangeboden bij twee verschillende praktijksimulaties. Indien er wordt van uitgegaan dat ze in beide praktijksimulaties de trekrelevantie verhoogden en dus de waarschijnlijkheid dat de assessee het relevante gedrag stelde, kan dit een gunstig effect gehad hebben op de convergente validiteit. In de conditie zonder gestandaardiseerde oefeninginstructies waren deze trekrelevante cues in geen van de praktijksimulaties aanwezig. Deze bevindingen zijn in lijn met eerder onderzoek van Haaland en Christiansen (2002; zie ook Lievens et al. 2006). Zij toonden reeds aan dat een verhoging van de TAP een gunstige invloed kan hebben op de convergente validiteit.

Een ander aspect van de TAP is de sterkte van de situatie. Een mogelijk gevaar van het gebruik van gestandaardiseerde oefeninginstructies is dat de situatie te sterk kan worden. Dit is echter nadelig omdat op die manier individuele verschillen worden onderdrukt (Haaland & Christiansen, 2002). Doordat in deze studie een positief effect werd waargenomen van gestandaardiseerde oefeninginstructies op de convergente

validiteit is het waarschijnlijk dat de situatie niet te sterk was en individuele verschillen tussen assessee's nog steeds geobserveerd konden worden.

Discriminante validiteit Voor de discriminante validiteit werd eveneens op basis van de TAT van Tett en Guterman (2000) vooropgesteld dat AC oefeningen met gestandaardiseerde oefeninginstructies een hogere discriminante validiteit zouden hebben dan AC oefeningen zonder gestandaardiseerde oefeninginstructies (Hypothese 1b). Deze hypothese kan aan de hand van de resultaten bevestigd worden. In de conditie waar gestandaardiseerde oefeninginstructies werden gebruikt, was er een hogere discriminante validiteit. De meting van verschillende dimensies met één instrument stemmen niet overeen.

De TAT stelt dat in een situatie meerdere trekrelevante cues voorzien moeten worden om een trek te kunnen activeren en ervoor te zorgen dat die trek tot uiting kan komen.

Vooreerst werden de gestandaardiseerde oefeninginstructies van de dimensies opgesteld aan de hand van meerdere relevante gedragingen per dimensie (Brannick, 2008). Dit aan de hand van de definities van Arthur et al. (2003).

Daarnaast kregen de assessee's in de gestandaardiseerde oefeninginstructies per dimensie ook meerdere relevante cues mee. Hierdoor kan het opnieuw mogelijk zijn dat zo een verhoging van de TAP werd verkregen, wat een gunstig effect had op de discriminante validiteit. Dit kan verklaard worden doordat de assessoren meerdere relevante gedragingen te zien kregen per dimensie en op die manier misschien beter konden differentiëren tussen de verschillende dimensies in één oefening (Brannick, 2008). Deze bevindingen zijn in lijn met eerder onderzoek van Lievens et al. (2006). Zij vonden reeds dat de discriminante validiteit beter is indien er meer dimensiegerelateerde gedragingen zijn die geobserveerd kunnen worden.

Met betrekking tot de constructvaliditeit kan er besloten worden dat voor beide hypothesen evidentie werd gevonden. En dat het gebruik van gestandaardiseerde oefeninginstructies een gunstig effect kan hebben op de constructvaliditeit van AC oefeningen.

Interbeoordelaarsbetrouwbaarheid

De tweede en laatste hypothese die werd geformuleerd betreft de interbeoordelaarsbetrouwbaarheid. Op basis van de TAT werd vooropgesteld dat in een AC oefening met gestandaardiseerde oefeninginstructies de interbeoordelaarsbetrouwbaarheid hoger zou liggen dan in een AC oefening zonder gestandaardiseerde oefeninginstructies. Er werd uitgegaan van de idee dat een verhoging van de TAP de observeerbaarheid van de trekken zou kunnen verbeteren, wat op zijn beurt gunstig zou kunnen zijn voor de interbeoordelaarsbetrouwbaarheid (Lievens & Van Keer, 1999). De gevonden resultaten spreken deze redenering echter tegen. De algemene interbeoordelaarsbetrouwbaarheid is laag in beide condities. Daarnaast is in de conditie met gestandaardiseerde oefeninginstructies de interbeoordelaarsbetrouwbaarheid lager dan in de conditie zonder de gestandaardiseerde oefeninginstructies.

Een mogelijke verklaring voor de algemeen lage interbeoordelaarsbetrouwbaarheid is dat bij deze studie gewerkt werd met onervaren assessoren, laatstejaarsstudenten Bedrijfspsychologie en Personeelsbeleid aan de Associatie Universiteit Gent. Het werken met onervaren assessoren kan een effect hebben op de interbeoordelaarsbetrouwbaarheid (Barr & Hitt, 1986). Doordat de assessoren niet vertrouwd waren met de te beoordelen situatie kan het zijn dat zij de trekrelevante gedragingen niet percipieerden als relevant voor een bepaalde dimensie.

Daarnaast werd aan de assessoren in deze studie gevraagd om de assessee te beoordelen op een bepaalde dimensie in een oefening. Onderzoek heeft reeds aangetoond dat bij deze zogenaamde binnen-oefening methode de interbeoordelaarsbetrouwbaarheid lager ligt (Lievens, 2001). Dit kan eveneens een mogelijke verklaring zijn voor de lage algemene interbeoordelaarsbetrouwbaarheid.

Een andere mogelijke verklaring is het ontbreken van een discussiefase na het observeren van de assessee in de praktijksimulaties. De interbeoordelaarsbetrouwbaarheid blijkt hoger te zijn nadat assessoren hun observaties met elkaar hebben gedeeld en zo tot een consensus trachten te komen. Zo kunnen ook observaties, die door de ene assessor opgemerkt werden en door de andere niet, gedeeld

worden (Lievens, 2001; Schmitt, 1977). Door het opzet van deze studie was een discussiefase echter niet mogelijk.

Tenslotte waren de assessoren niet op de hoogte van de gestandaardiseerde oefeninginstructies. Dit is een mogelijke verklaring voor de daling van de interbeoordelaarsbetrouwbaarheid in de conditie met gestandaardiseerde oefeninginstructies. Doordat de assessoren niet op de hoogte waren konden de gestandaardiseerde oefeninginstructies niet zorgen voor een beperking van de cognitieve overload bij de assessoren. Het is mogelijk dat zij de relevante gedragingen wel opmerkten maar dat dit niet voldoende was om die gedragingen ook correct te identificeren. Daarnaast is het mogelijk dat door de gestandaardiseerde oefeninginstructies de assessee's meer verschillende trekrelevante gedragingen gingen stellen. Aangezien de assessoren hiervan niet op de hoogte waren kan dit de cognitieve overload verhoogt hebben in plaats van verlaagt. Daardoor hadden de gestandaardiseerde oefeninginstructies op de observaties en de beoordeling van de assessoren misschien niet het gewenste effect. Wat op zijn buurt de interbeoordelaarsbetrouwbaarheid verlaagt kan hebben.

Met betrekking tot de interbeoordelaarsbetrouwbaarheid werd er geen verbetering gevonden en werd de vooropgestelde hypothese dus niet bevestigd.

Methodologische Reflecties

Hieronder worden een aantal reflecties besproken met betrekking tot de methodologische beperkingen van deze studie. Achtereenvolgens wordt de steekproef, de procedure, en de data-analyse besproken. Mogelijke oplossingen voor deze beperkingen worden besproken in de paragraaf Toekomstig Onderzoek.

Steekproef

Een eerste beperking van deze studie betreft de steekproef. Deze bestond voor zowel de deelnemers, assessoren als rollenspelers uitsluitend uit laatstejaarsstudenten van de Associatie Universiteit Gent. Hier rijzen vragen met betrekking tot de externe validiteit. Zijn de resultaten van deze studie representatief voor een meer algemene

populatie van assessoren? Uit vroeger onderzoek blijkt bijvoorbeeld dat studenten en managers significant verschillen in de beslissingen die ze nemen met betrekking tot de beoordeling in selectie (Barr & Hitt, 1986). Verder kan ook de vraag gesteld worden of de resultaten generaliseerbaar zijn naar andere populaties van deelnemers, zoals sollicitanten en mensen die reeds werken. Tenslotte is het ook mogelijk dat de rollenspelers in deze studie zich anders gedroegen dan professionele rollenspelers.

Procedure

Met betrekking tot de procedure kunnen eveneens enkele opmerkingen gemaakt worden.

Ten eerste waren zowel de assessoren als de rollenspelers die deelnamen aan deze studie onervaren. Opnieuw werden hier laatstejaarsstudenten ingeschakeld, deze studeerden allen Bedrijfspsychologie en Personeelsbeleid aan de Associatie Universiteit Gent. Ook al kregen zij een grondige praktijkgerichte training in observeren en scoren, wat de beoordelingen normaal ten goede komt, (Woehr & Hufcutt, 1994; Lievens, 1998) kan hun gebrek aan relevante ervaring met ACs een invloed gehad hebben op de resultaten.

Ten tweede waren alle assessoren laatstejaarsstudenten Bedrijfspsychologie en Personeelsbeleid, wat eventueel een vertekening van de resultaten met zich mee kan brengen. Zo blijken psychologen als assessoren een gunstig effect te hebben op de constructvaliditeit van ACs (Sagie & Mangezy, 1997).

Ten derde moesten assessoren soms tot twintig oefeningen per dag observeren. Dit AC proces bestond slechts uit twee oefeningen en vertoonde dus al snel een repetitief karakter. Voor de assessoren zal het dan ook moeilijk geweest zijn om voortdurend met dezelfde focus en aandacht te observeren en de dimensies te scoren. Er zijn namelijk beperkingen aan de hoeveelheid informatie die mensen kunnen verwerken (Miller, 1956).

Ten vierde waren er in deze sollicitatietraining slechts twee AC oefeningen. De assessoren hadden dus maar een beperkte kans om het gedrag van de assessee te observeren. Normaal is er in een AC proces een grotere diversiteit aan informatiebronnen om gedrag te scoren en te interpreteren (International Task Force, 2009).

Data-analyse

De toegepaste analysemethode, namelijk een generaliseerbaarheidsanalyse, is een vernieuwende methode om de constructvaliditeit van ACs te onderzoeken. Om na te gaan of de resultaten van deze studie niet inherent zijn aan het gebruik van de generaliseerbaarheidsanalyse werd ook een multitrek – multimethode (MTMM) analyse uitgevoerd met het statistisch programma SPSS 16. De resultaten van de volledige analyse voor de conditie zonder gestandaardiseerde oefeninginstructies is terug te vinden in Bijlage 7, die voor de conditie met gestandaardiseerde oefeninginstructies in Bijlage 8.

In de MTMM-matrix is de monotrek – heteromethode correlatie een indicatie voor de convergente validiteit. Hoe hoger deze waarde hoe beter de convergente validiteit. In de conditie zonder gestandaardiseerde oefeninginstructies is de gemiddelde waarde van de monotrek - heteromethode correlatie .422. In de conditie met gestandaardiseerde oefeninginstructies is deze correlatie gelijk aan .483. De conditie met gestandaardiseerde oefeninginstructies blijkt een betere convergente validiteit te hebben dan de conditie zonder gestandaardiseerde oefeninginstructies.

De heterotrek – monomethode correlatie is een indicatie voor de discriminante validiteit. Hoe lager deze waarde hoe beter de discriminante validiteit. In de conditie zonder gestandaardiseerde oefeninginstructies is de gemiddelde waarde van de heterotrek – monomethode correlatie over de oefeningen heen .495. In de conditie met gestandaardiseerde oefeninginstructies is deze correlatie gelijk aan .423.

De resultaten geven aan dat er in de conditie met gestandaardiseerde oefeninginstructies een betere discriminante validiteit wordt gevonden dan in de conditie zonder gestandaardiseerde oefeninginstructies.

Algemeen kan er gesteld worden dat de resultaten van de MTMM-matrix parallel zijn met de resultaten verkregen aan de hand van de uitgevoerde generaliseerbaarheidsanalyse.

Theoretische en Praktische Implicaties

Vooreerst levert deze studie een bijdrage aan het, tot nu toe, schaars onderzoek rond de invloed van oefeninginstructies op de verschillende psychometrische eigenschappen van ACs. Ondanks het feit dat er op dit moment nog maar weinig onderzoek naar werd gedaan, gaven onderzoekers reeds aan dat oefeninginstructies een invloed kunnen hebben op het uitlokken van trekrelevant gedrag (Thornton & Mueller-Hanson, 2004). Uit de resultaten van de huidige studie blijkt dat de standaardisatie van oefeninginstructies een gunstig effect zou kunnen hebben op de constructvaliditeit.

Ten tweede wordt in deze studie zowel evidentie voor de convergente als discriminante validiteit gevonden. In vroegere onderzoeken werden deze vormen van constructvaliditeit ofwel niet gevonden ofwel waren de resultaten tegenstrijdig (Henderson et al., 1995; Joyce et al., 1994; Lievens, 2006; Schneider & Schmitt, 1992). Hiermee wordt voor een stuk tegemoetgekomen aan de paradox binnen ACs (Woehr & Arthur, 2003).

Ten slotte bouwt deze studie verder op het feit dat de TAT van Tett en Gutterman (2000) een goede invalshoek kan zijn om de constructvaliditeit van ACs te onderzoeken (Haaland & Christiansen, 2002; Lievens et al., 2006, 2009; Tett & Gutterman, 2000). In deze studie werd de TAT prescriptief angewend. Dit in tegenstelling tot vorige studies waar de TAT vooral descriptief werd angewend. Er werden op basis van de TAT gestandaardiseerde oefeninginstructies ontwikkeld. Deze studie gaat ervan uit dat deze gestandaardiseerde oefeninginstructies kunnen fungeren als trekrelevante cues, het kwalitatieve luik van de TAP. Bij het gebruik van gestandaardiseerde oefeninginstructies is het wel van belang ook rekening te houden met het kwantitatieve luik van de TAP, de sterkte van de situatie. Een situatie mag namelijk niet te sterk zijn, anders komen individuele verschillen niet tot uiting. Er mogen dus wel cues aanwezig zijn om trekrelevant gedrag uit te lokken, maar de situatie moet nog voldoende ambigue zijn om een verscheidenheid aan reacties te krijgen.

Het prescriptief toepassen van de TAT door middel van gestandaardiseerde oefeninginstructies zou er dus voor kunnen zorgen dat er een meer gunstige convergente en discriminante validiteit wordt bekomen bij ACs.

Deze studie geeft eveneens enkele nuttige praktische implicaties voor het AC proces.

Vooreerst kan het bij het design van ACs nuttig zijn om te werken met gestandaardiseerde oefeninginstructies om tot een kwalitatief beter oordeel te komen. Indien men vooraf gaat bepalen welke gedragingen relevant zijn voor de targetfunctie, kan men deze verwerken in de gestandaardiseerde oefeninginstructies. Zo kan men cues voorzien om het trekrelevante gedrag uit te lokken. Het belang van een goede afstemming tussen dimensies en oefeningen werd reeds aangegeven in de literatuur (Howard, 1997).

Ten slotte is het gebruik van gestandaardiseerde oefeninginstructies een goedkope aanpassing aan het design van ACs, die makkelijk te implementeren is, en een verbetering van de kwaliteit van de beoordeling kan bewerkstelligen.

Toekomstig Onderzoek

Inhoudelijk

Vooreerst kan er op inhoudelijk vlak, naar de toekomst toe, een uitbreiding van het huidige onderzoek plaatsvinden. Nu werd de invloed van gestandaardiseerde oefeninginstructies nagegaan op de constructvaliditeit en de interbeoordelaarsbetrouwbaarheid. Er kan echter ook worden nagegaan wat de invloed is van gestandaardiseerde oefeninginstructies op de observeerbaarheid van het gedrag. Eén van de grote voordelen van ACs is dat assessoren de kans krijgen om het complexe gedrag te observeren in een situatie die gelijkend is op de werkelijkheid. Dit zorgt echter ook voor een uitdaging, aangezien er niet altijd genoeg relevant gedrag per dimensie is om te observeren (Reilly et al., 1990). Soms moeten assessoren één gedraging gebruiken om meerder dimensies te scoren (Brannick, 2008). De vraag is nu of het gebruik van gestandaardiseerde oefeninginstructies een gunstig effect kan hebben op de observeerbaarheid van gedrag. Daarnaast kan onderzocht worden of deze gestandaardiseerde oefeninginstructies meerdere relevante gedragingen per dimensie gaan uitlokken, zoals wordt opgedragen door de *Guidelines* (International Task Force, 2009).

Uit een onderzoek van Kleinmann et al. (1996) bleek dat indien de assesses vooraf op de hoogte werden gebracht over de dimensies, dit een gunstig effect had op de constructvaliditeit. Een nadelig effect van het op voorhand meedelen van de dimensies was een verhoogde transparantie, wat de criteriumvaliditeit niet ten goede komt. Er werd in deze studie getracht om hieraan tegemoet te komen door het definiëren van de gestandaardiseerde oefeninginstructies in plaats van de dimensies uit te leggen. Toch kan het nuttig zijn om in toekomstig onderzoek na te gaan wat het effect is van deze gestandaardiseerde oefeninginstructies op de transparantie.

Ten slotte kan er in een volgend onderzoek gebruik gemaakt worden van andere dimensies, andere of meerdere oefeningen en andere bronnen van variantie.

Methodologie

Op het vlak van de methodologie worden de volgende voorstellen voor toekomstig onderzoek aangereikt.

Ten eerste kan de externe validiteit van de gevonden resultaten worden nagegaan. Het onderzoek werd uitgevoerd bij laatstejaarsstudenten van de Associatie Universiteit Gent die deelnamen op vrijwillige basis. Omwille van de specificiteit van deze steekproef moet voorzichtig worden omgesprongen met een generalisatie van de resultaten. Zo zou men in de toekomst random steekproeven kunnen gebruiken van mensen die al enige praktijkervaring hebben. Door de praktijkervaring van deze mensen zullen zij de situaties in de praktijksimulaties waarschijnlijk meer herkennen en daar ook anders op reageren dan onervaren studenten. Daarnaast kan worden nagegaan of de effecten van gestandaardiseerde oefeninginstructies niet wegvallen, omdat deze door de herkenbaarheid van de situatie als niet-relevant worden gezien. Uiteraard kan ook binnen ‘mensen met praktijkervaring’ nog een verscheidenheid van steekproeven worden gegenereerd.

Verder kan er worden nagegaan of de gevonden resultaten gerepliceerd worden wanneer men ervaren assessoren en rollenspelers inzet in plaats van laatstejaarsstudenten Bedrijfspsychologie en Personeelsbeleid, die een training kregen. Mensen die ervaring als assessor of rollenspeler hebben in ACs geven misschien zelf al genoeg trekrelevante cues aan de assesses. Dan hebben gestandaardiseerde oefeninginstructies in die context misschien geen effect op de constructvaliditeit.

Deze studie maakte voor de data-analyse gebruik van een generaliseerbaarheidsanalyse, wat een vernieuwende manier is om de constructvaliditeit binnen ACs te benaderen. Toekomstig onderzoek kan zich richten op het verder gebruik van deze statistische methode. Daarnaast kan het gebruik van een confirmatorische factoranalyse eveneens een invalshoek zijn om de constructvaliditeit binnen ACs te benaderen (Lievens, 2009).

Ten slotte kan er nagegaan worden of het gebruik van gestandaardiseerde oefeninginstructies nuttig kan zijn om de cognitieve overload bij assessoren te beperken. Indien de assessoren vooraf getraind worden en op de hoogte zijn van de gestandaardiseerde oefeninginstructies zullen zij misschien in staat zijn om de relevante gedragingen beter op te merken en te classificeren. Dit zou mogelijk een gunstig effect kunnen hebben op de interbeoordelaarsbetrouwbaarheid.

Conclusie

Als algemene conclusie kan gesteld worden dat gestandaardiseerde oefeninginstructies in deze studie een gunstig effect blijken te hebben op zowel de convergente als de discriminante validiteit. De TAT van Tett en Guterman (2000) kan mogelijk een verklaring bieden voor de gevonden resultaten. Een verhoging van de TAP lijkt een impact te hebben op de constructvaliditeit. Zowel de trekrelevantie als de sterkte van de situatie lijken van belang te zijn voor de constructvaliditeit.

Er werd geen evidentie gevonden dat het gebruik van gestandaardiseerde oefeninginstructies een verbetering zou kunnen betekenen voor de interbeoordelaarsbetrouwbaarheid. Toekomstig onderzoek kan daarom nagaan wat de invloed is van gestandaardiseerde oefeninginstructies op de observeerbaarheid van het gedrag in AC oefeningen en op de cognitieve overload voor assessoren, om zo eventueel de interbeoordelaarsbetrouwbaarheid te kunnen verbeteren.

REFERENTIES

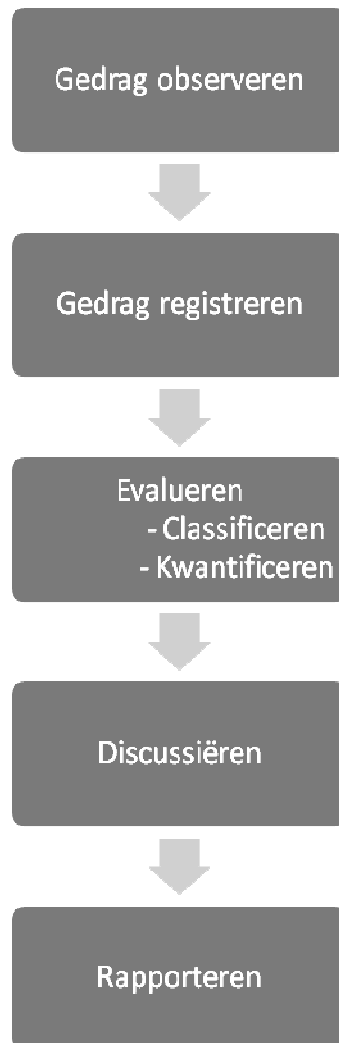
- Arthur, W., Day, E.A., McNelly, T.L., & Edens, P.S. (2003). A meta-analysis of the criterion-related validity of assessment centre dimensions. *Personnel Psychology*, 56, 125-154.
- Barr, S.H., Hitt M.A. (1986). A comparison of selection decision-models in managers versus student samples. *Personnel Psychology*, 39, 599-617
- Brannick, M.T. (2008). Back to Basics of Test Construction and Scoring. *Industrial and Organizational Psychology*, 1, 131-133.
- Brennan, R.L. (1992). *Elements of Generalizability Theory*. American College Testing Program: Iowa City, IA.
- Bycio, P., Alvares, K. M., & Hahn, J. (1987). Situation specificity in assessment centre ratings: a confirmatory analysis. *Journal of Applied Psychology*, 72, 463-474.
- Conway, J.M., Jako, R.A., & Goodman, D.F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 565-579.
- Damitz, M., Manzey, D., Kleinmann, M., & Severin, K. (2003). Assessment Centre for Pilot Selection: Construct and Criterion Validity and the Impact of Assessor Type. *Applied Psychology: An International Review*, 52, 193-212.
- Donahue, L.M., Truxillo, D.M., Cornwell, J.M., & Gerrity, M.J. (1997). Assessment centre construct validity and behavioural checklists: some additional findings. *Journal of Social Behavior and Personality*, 12, 85-105.
- Fletcher, C. & Kerslake C. (1992). The impact of assessment centres and their outcomes on participants' self-assessments. *Human Relations*, 45, 281-289.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, C. (1987). Meta-analysis of assessment centre validity. *Journal of Applied Psychology*, 74, 611-618.
- Gaugler, B.B. & Thornton, G.C. (1989). Number of assessment centre dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, 74, 611-618.
- Haaland, S. & Christiansen, N.D. (2002). Implications of Trait-Activation Theory for Evaluating the Construct Validity of Assessment Centre Ratings. *Personnel Psychology*, 55, 137-163.

- Henderson, F., Anderson, N., & Rick, S. (1995). Future competency profiling: validating and redesigning the ICL graduate assessment center. *Personnel Review*, 24, 19-31.
- Howard, A. (1997). A reassessment of assessment centers: challenges for the 21st century. *Journal of Social Behavior and personality*, 12, 13-52.
- International Task Force on Assessment Center Guidelines (2009). Guidelines and ethical considerations for assessment center operations. *International Journal of Selection and Assessment*, 17, 243-253.
- Jackson, D.J.R., Atkins, S.G., Fletcher, R.B., & Stillman, J.A. (2005). Frame of Reference Training for Assessment Centers: Effects on Interrater Reliability When Rating Behaviors and Ability Traits. *Public Personnel Management*, 34, 17-30.
- Jansen, P. & De Jongh, F. (1997). *Assessment Centres: A Practical Handbook*. Chichester, ENG: Wiley.
- Joyce, L. W., Thayer, P. W., & Pond, S.B. (1994). Managerial functions: an alternative to traditional assessment center dimensions? *Personnel Psychology*, 47, 109-121.
- Kane, M.T., (1982). A sampling model for validity. *Applied Psychological Measurement*, 6, 125-160.
- Kleinmann, M., Kuptsch, C., & Köller, O. (1996). Transparency: A necessary requirement for the construct validity of assessment centres. *Applied Psychology: An International Review*, 45, 67-84.
- Kok, B. & De Jongh, F. (2004). *Assessment doen: Hoe werkt het voor jou?* Utrecht: Het Spectrum.
- Larsen, R.J. & Buss, D.M. (2005) *Differential and Personality Psychology: Domains of Knowledge About Human Nature*. (2nd ed.). New York: McGraw-Hill.
- Lievens, F. (1998). Factors which improve the construct validity of Assessment Centers: A review. *International Journal of Selection and Assessment*, 6, 141-152.
- Lievens, F. (2001). An examination of the accuracy of slogans related to assessment centres. *Personnel Review*, 31, 86-97.
- Lievens, F. (2002). Trying to understand the different pieces of the construct validity puzzle of assessment centers: an examination of assessor and assessee effects. *Journal of Applied Psychology*, 87, 675-686.

- Lievens, F. (2006). Selectie en assessment van personeel. *Handboek Human Resource Management. Back to basic*. 121-166. Den Haag: Uitgeverij Lannoo.
- Lievens, F. (2009). Assessment centres: A tale about dimensions, exercises, and dancing bears. *European Journal of Work and Organizational Psychology*, 18, 102-121
- Lievens, F., Chasteen, C.D., Day, E.A., & Christiansen, N.D. (2006). Large Scale Investigation of the Role of Trait Activation Theory for Understanding Assessment Center Convergent and Discriminant Validity. *Journal of Applied Psychology*, 91, 247-258.
- Lievens, F. & Conway, J.M. (2001). Dimension and exercise variance in assessment center scores: a large-scale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, 86, 1202-1222.
- Lievens, F., De Fruyt, F., & Van Dam, K. (2001). Assessors use of personality traits in descriptions of assessment center candidates: A five factor model perspective. *Journal of Occupational & Organizational Psychology*, 74, 623-636.
- Lievens, F. & Van Keer, E. (1999). *Assessment Centers in Vlaanderen: Een wetenschappelijke Evaluatie*. Gent: Academia Press.
- Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological review*, 80, 252-253.
- Reber, A.S. (2002). *Woordenboek van de psychologie: uitgebreide en herziene editie*. Negende druk. Amsterdam: Uitgeverij Bert Bakker.
- Reilly, R. R., Henry, S., & Smither, J. W. (1990). An examination of the effects of using behaviour checklists on the construct validity of assessment center dimensions. *Personnel Psychology*, 43, 71-84.
- Riggio, R. E. & Mayes B. T. (Eds.). (1997). Assessment centers: Research and applications [special issue]. *Journal of Social Behavior and Personality*, 12, 3-12.
- Robie, C., Osburn, H.G., Morris, M.A., Etchegaray, J.M., & Adams, K.A. (2000). Effects of the Rating Process on Construct Validity of Assessment Center Dimensions Evaluations. *Human Performance*, 13, 355-370.

- Sackett, P. R. (1987). Assessment centers and content validity: some neglected issues. *Personal Psychology*, 40, 120-137.
- Sackett, P. R. & Dreher, G. F. (1982). Constructs and assessment centre dimensions: some troubling empirical findings. *Journal of Applied Psychology*, 67, 401-410.
- Sagie, A. & Mangezy, R. (1997). Assessor Type, Number of Distinguishable Dimension Categories, and Assessment Centre Construct Validity. *Journal of Occupational and Organizational Psychology*, 70, 103-108.
- Schmitt, N. (1977). Interrater Agreement in Dimensionality and Combination of Assessment Center Judgements. *Journal of Applied Psychology*, 62, 171-176.
- Schneider, J.R. & Schmitt, N. (1992). An Exercise Design Approach to Understanding Assessment Center Dimension and Exercise Constructs. *Journal of Applied Psychology*, 77, 32-41.
- Tett, R. P. & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88, 500-517.
- Tett, R. P. & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: testing a principle of trait activation. *Journal of Research in Personality*, 34, 397-423.
- Thornton, G. C. (1992) *Assessment centers in human resource management*. Reading, MA: Addison-Wesley.
- Thornton, G. C. (2006). *Assessment centers in human resource management*. Imprint reading (Mass.): Addison-Wesley, 2006. Descr. XV, 276p. Series Addison-Wesley series on managing human resources
- Thornton, G. C., III & Byham, W. C. (1982). *Assessment centers and managerial performance*. New York: Academic Press.
- Thornton, G.C. III & Mueller-Hanson, R.A. (2004). *Developing Organizational Simulations: A Guide for Practitioners and Students*. Mahwah, NJ: Erlbaum.
- Van Beirendonck, L. (1998). *Beoordelen en ontwikkelen van competenties, assessment centers, development centers en aanverwante technieken*. Leuven: Acco.
- Van Minden, J.J.R. (2006). *Alles over Assessment Centers*. Elfde herziene druk. Amsterdam/Antwerpen: Business Contact.

- Woehr, D.J. & Arthur, W. Jr. (2003). The construct-related validity of assessment center ratings: a review and meta-analysis of the role of methodological factors. *Journal of Management*, 29, 231-258.
- Woehr, D. J. & Huffcutt, A. I. (1994). Rater training for performance appraisal: a meta-analytic review. *Journal of Occupational an Organizational Psychology*, 67, 189-205.

BIJLAGEN**Bijlage 1**

Figuur 1.

Een voorbeeld van een beoordelingsproces van het AC (Lievens & Van Keer, 1999)

Bijlage 2

Tabel 3.

Overzicht van scoringsmethodes (Robie et al., 2000)

Scoringsmethode	Beschrijving
Gedragrapportering methode	De assessoren observeren het gedrag in alle oefeningen, daarna geven ze een algemene score op de dimensies.
Oefening-dimensie methode	De assessoren observeren het gedrag in alle oefeningen, daarna geven ze een beoordeling op de dimensies binnen elke oefening.
Binnen-oefening methode	De assessoren observeren het gedrag binnen elke oefening, onmiddellijk na elke oefening beoordelen ze de dimensies.
Cross-oefening methode	De assessoren observeren het gedrag binnen één oefening, ze beoordelen de dimensie van die ene oefening.

Bijlage 3

Tabel 4.

De trekken van het VFM (Larsen & Buss, 2005)

Trek van het VFM	Beschrijving
Openheid-Intellect	Creatief, intelligent, fantasierijk
Consciëntieusheid	Verantwoordelijk, ordelijk, overtuigend, principieel
Extraversie	Sociaal, open, avontuurlijk, risico nemend, graag praten
Altruïsme	Vriendelijk, niet jaloers, meewerkend, meevoelend met anderen
Neuroticisme	Zenuwachtig, opgewonden, gespannen

Bijlage 4

Tabel 5 .

Overzicht van de gestandaardiseerde oefeninginstructies

Presentatie	
Dimensie	Gestandaardiseerde Oefeninginstructie
Interpersoonlijke sensitiviteit	Hou bij het opstellen en presenteren van uw ideeën rekening met mogelijke gevoelens van de rollenspeler. Het is niet de bedoeling dat u de rollenspel bruuskeert.
Plannen en Organiseren	Werk dit voorstel uit. Tracht dit planmatig aan te pakken.
Probleemanalyse en -oplossing	Maak een analyse van deze tekst en haal er de kern uit. Tracht ideeën te bedenken en maak een overzicht van uw voorstel, dat u nadien zal presenteren.
Stresstolerantie	U mag er vanuit gaan dat de rollenspeler niet dadelijk akkoord zal gaan met uw voorstel en u onder druk zal zetten. Blijf kalm en beheerst.
Werkgesprek	
Dimensie	Gestandaardiseerde Oefeninginstructie
Interpersoonlijke sensitiviteit	Het is niet de bedoeling dat u de rollenspeler bruuskeert. Hou dus ook rekening met zijn mening en gevoelens.
Plannen en organiseren	Werk een concrete oplossing uit en maak afspraken op het einde.
Probleemanalyse en -oplossing	Het is de bedoeling dat u het eigenlijke probleem vindt van het minder functioneren van Dominique en er een oplossing voor vindt. U twijfelt eraan of u Dominique wel zult aanmelden voor de cursus 'Reclamebeheer'. U moet Dominique corrigeren en motiveren om het werk weer goed te doen.
Stresstolerantie	U mag er vanuit gaan dat de rollenspeler niet dadelijk akkoord zal gaan met uw voorstel en u onder druk zal zetten. Blijf kalm en beheerst.

Bijlage 5

Manipulation Check Vragenlijst

Beste,

Hierbij vindt u een aantal oefeninginstructies. Dit zijn vooraf bepaalde instructies die aan kandidaten bij het begin van de assessment center oefening voorgelegd worden, met als doel dimensiegerelateerd gedrag uit te lokken.

Wij vragen u de dimensie aan te duiden waaronder het gedrag dat met de oefeninginstructie wordt uitgelokt het best kan geclassificeerd worden. U kan de dimensie aanduiden door een kruisje te plaatsen. Indien u vindt dat de oefeninginstructie met geen enkele van de dimensies overeenstemt, kan u een kruisje plaatsen bij “niet van toepassing”. Gelieve bij elke afzonderlijke oefeninginstructie slechts één dimensie aan te duiden.

Ter verduidelijking definiëren we de dimensies als volgt:

Stresstolerantie/omgaan met stress (Stresstolerantie):

- Effectief handelen, ondanks werkdruk, tegenstand en ontgoocheling.

Analyse en probleemoplossing (Analyse en oplossing):

- Informatie verzamelen;
- Technische en professionele informatie begrijpen;
- Data en informatie accuraat analyseren;
- Bruikbare ideeën en oplossingen genereren;
- Verbeterende acties voor problemen en situaties selecteren;
- De beschikbare middelen op vernieuwende wijze toepassen;
- Innovatieve oplossingen genereren en herkennen.

Interpersoonlijke sensitiviteit/empatisch handelen (Sensitiviteit):

- De acties van het individu weerspiegelen begrip voor de gevoelens en de behoeften van anderen, zowel binnen als buiten de organisatie.

Plannen en organiseren (Plannen):

- Zowel eigen werk en middelen als die van anderen systematisch ordenen met het oog op efficiëntie;
- Zich voorbereiden op de toekomst;
- Anticiperen op de toekomst.

Bedankt voor uw medewerking!

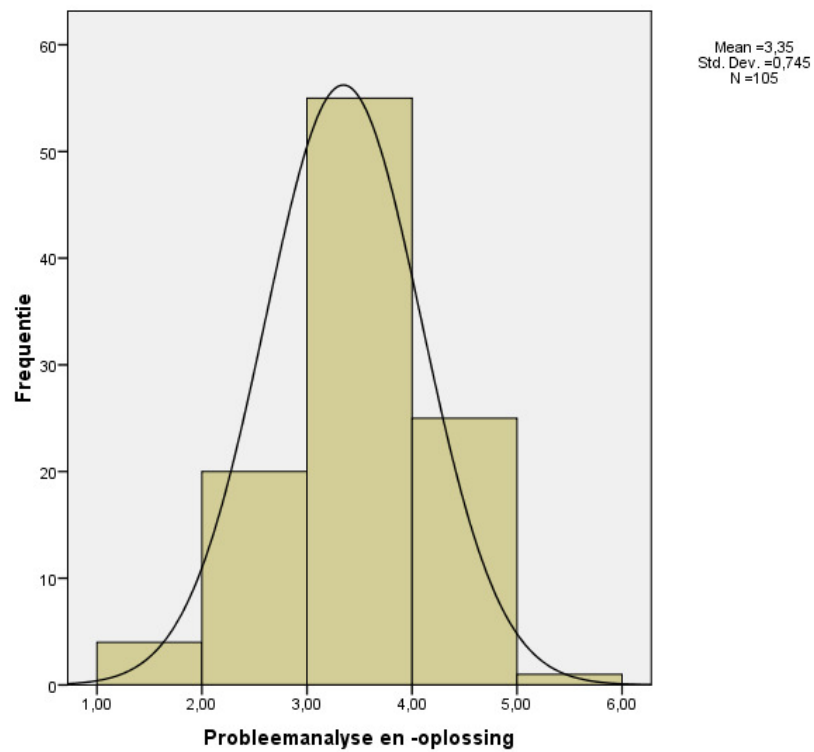
Assessment center oefening: Presentatie	Stress- tolerantie	Analyse en oplossing	Sensitiviteit	Plannen	Niet van toepassing
“Maak een analyse van deze tekst en haal er de kern uit. Tracht ideeën te bedenken en maak een overzicht van uw voorstel, dat u nadien zal presenteren.”					
“Onderneem tijdens uw presentatie actie uit eigen beweging, en tracht pro-actief te handelen.”					
“U mag er vanuit gaan dat de rollenspeler niet dadelijk akkoord zal gaan met uw voorstel en u onder druk zal zetten. Blijf kalm en beheerst.”					
“Werk dit voorstel concreet uit. Tracht dit planmatig aan te pakken.”					
“Zorg dat u uw gebaren en stem bewust gebruikt en op elkaar afstemt, zodat u uw boodschap duidelijk overbrengt.”					
“Draag bij het bedenken van uw ideeën sociale en ethische normen en waarden steeds hoog in het vaandel.”					
“Hou bij het opstellen en presenteren van uw ideeën rekening met de mogelijke gevoelens van de rollenspeler. Het is niet de bedoeling dat u de rollenspeler bruskeert.”					
“Durf bepaalde risico’s aangaan om uiteindelijk een voordeel te behalen”					

Gelieve onderstaande vragen in te vullen. De gegevens zullen vertrouwelijk behandeld worden en zullen niet ter beschikking van derden gesteld worden.

1. U bent: ☐ man ☐ vrouw
2. Uw leeftijd:jaar
3. Uw anciënniteit:jaar
4. Uw huidige functie:
5. Uw diploma('s) (zo precies mogelijk omschrijven)

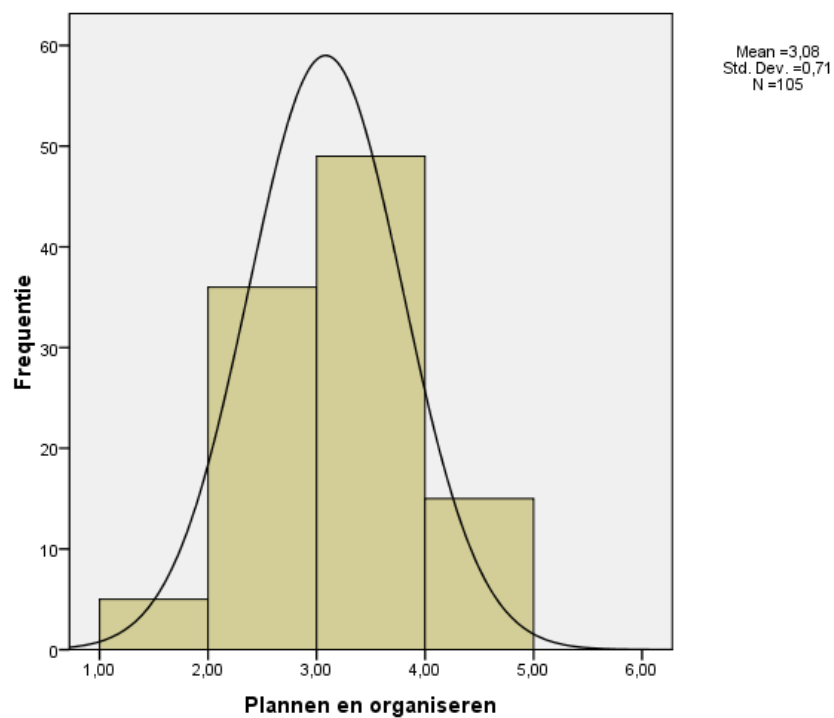
Bedankt voor uw medewerking!

Bijlage 6



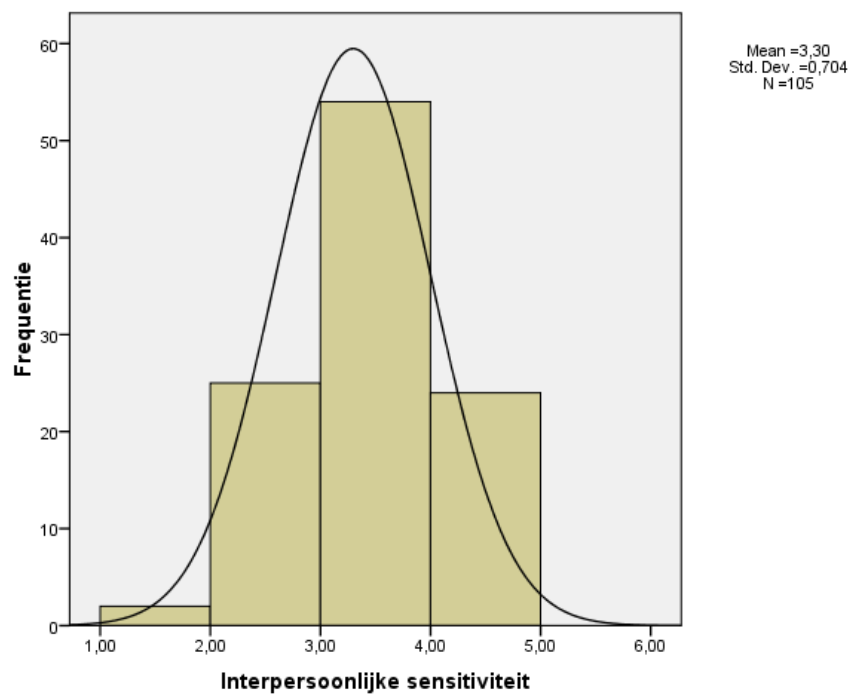
Figuur 2.

Normaalverdeling Probleemanalyse en- oplossing in het Werkgesprek



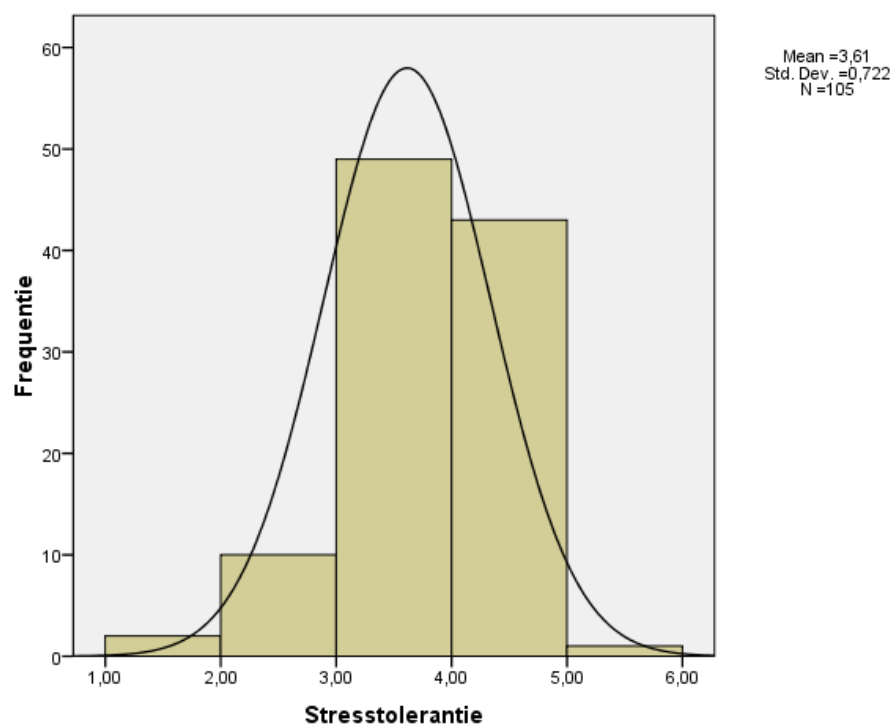
Figuur 3.

Normaalverdeling Plannen en organiseren in het Werkgesprek



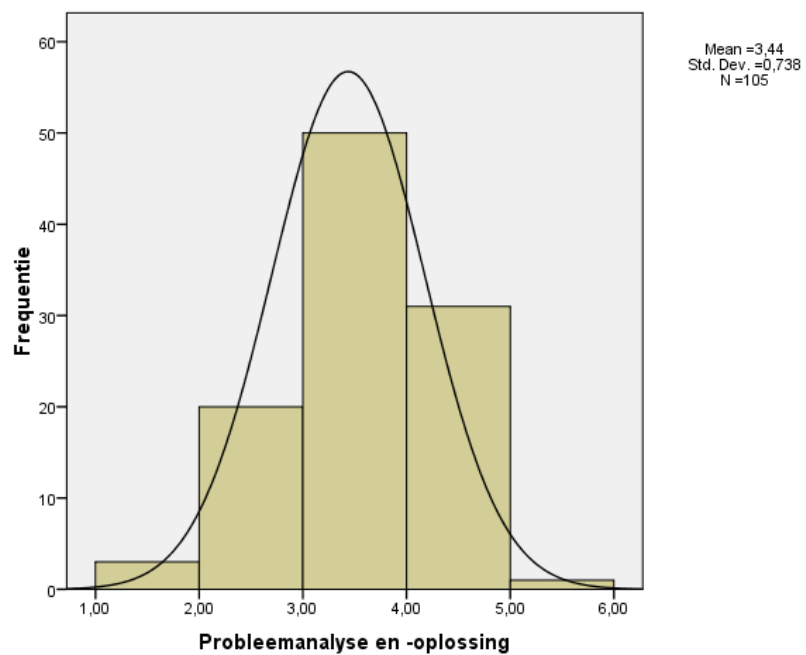
Figuur 4.

Normaalverdeling Interpersoonlijke sensitiviteit in het Werkgesprek



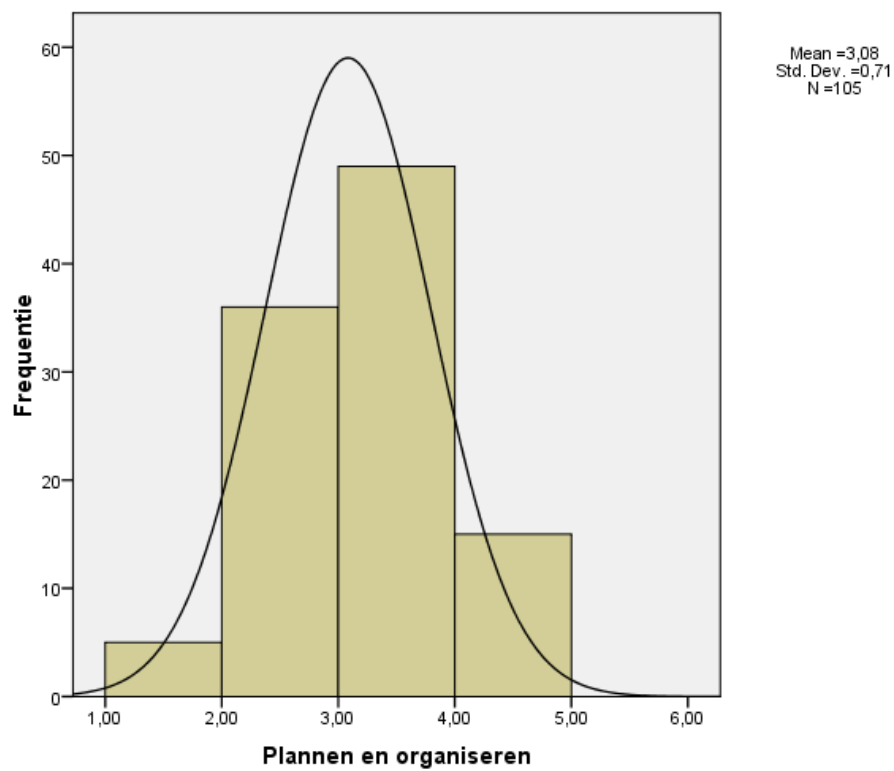
Figuur 5.

Normaalverdeling Stresstolerantie in het Werkgesprek



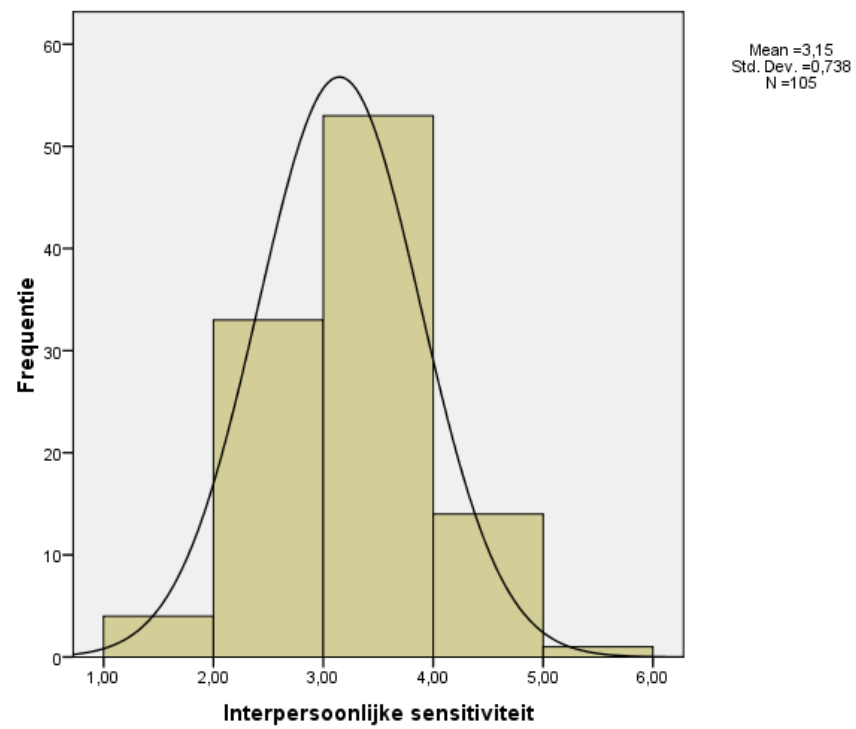
Figuur 6.

Normaalverdeling Probleemanalyse en- oplossing in de Presentatie



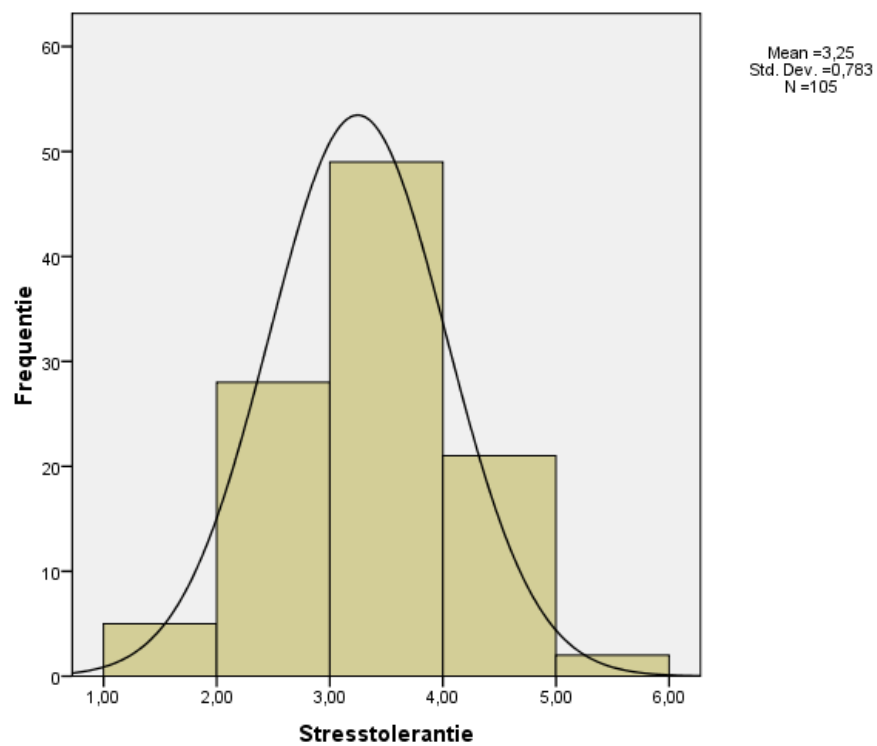
Figuur 7.

Normaalverdeling Plannen en organiseren in de Presentatie



Figuur 8.

Normaalverdeling Interpersoonlijke sensitiviteit in de Presentatie



Figuur 9.

Normaalverdeling Stresstolerantie in de Presentatie

Bijlage 7

Tabel 6.

MTMM matrix conditie zonder gestandaardiseerde oefeninginstructies (N=26)

Variabelen	Werkgesprek				Presentatie			
	1	2	3	4	5	6	7	8
Werkgesprek								
1. Probleemanalyse en oplossing								
2. Plannen en organiseren	<u>.464*</u>							
3. Interpersoonlijke sensitiviteit	<u>.666**</u>	<u>.489*</u>						
4. Stresstolerantie	<u>.428**</u>	<u>.588**</u>	<u>.391</u>					
Presentatie								
5. Probleemanalyse en oplossing	.252	.466*	.341	.639**				
6. Plannen en organiseren	.177	.443**	.319	.423*	<u>.654**</u>			
7. Interpersoonlijke sensitiviteit	.289	.378	.380	.449*	<u>.321</u>	<u>.425*</u>		
8. Stresstolerantie	.311	.325	.153	.613**	<u>.562**</u>	<u>.349</u>	<u>.610</u>	

Noot. De schuin gedrukte waarden zijn de monotrek - heteromethode correlaties;
de onderlijnde waarden zijn de heterotrek - monomethode correlaties.

* $p < .05$ ** $p < .01$

Bijlage 8

Tabel 7.

MTMM matrix conditie met gestandaardiseerde oefeninginstructies (N=25)

Variabelen	Werkgesprek				Presentatie			
	1	2	3	4	5	6	7	8
Werkgesprek								
1. Probleemanalyse en oplossing								
2. Plannen en organiseren	<u>.694**</u>							
3. Interpersoonlijke sensitiviteit	<u>.615**</u>	<u>.602**</u>						
4. Stresstolerantie	<u>.429**</u>	<u>.569**</u>	<u>.494*</u>					
Presentatie								
5. Probleemanalyse en oplossing	.468*	.565**	.504*	.615**				
6. Plannen en organiseren	.459*	.533**	.218	.394*	<u>.709**</u>			
7. Interpersoonlijke sensitiviteit	.222	.041	.317	.005	<u>.217</u>	<u>.110</u>		
8. Stresstolerantie	.125	.319	.243	.615**	<u>.263</u>	<u>.156</u>	<u>.228</u>	

Noot. De schuin gedrukte waarden zijn de monotrek - heteromethode correlaties;
de onderlijnde waarden zijn de heterotrek - monomethode correlaties.

* $p < .05$ ** $p < .01$