

Professionele expertise van hogere functionarissen

Verschillen tussen zelf-beoordelingen en chef-beoordelingen

Beatrice van der Heijden

Motivatie en aanleiding voor de studie

Het management in arbeidsorganisaties heeft in toenemende mate aandacht voor de bewaking van een blijvende inzetbaarheid van zijn menselijk potentieel. Individuele werknemers die in staat zijn om blijvend een waardevolle bijdrage te leveren aan de organisatie, zijn degenen die in het bezit zijn van professionele kennis en vaardigheden die 'up-to-date' zijn. Bovendien beschikken zij over het vermogen om nieuwe professionele expertise op te bouwen in aangrenzende gebieden.

Het gevaar van obsoletie (of veroudering), dat wil zeggen de mate waarin werknemers een gebrek hebben aan kennis en vaardigheden die noodzakelijk zijn om effectief functioneren mogelijk te maken, is groter naarmate de vereiste kwalificaties aan meer verandering onderhevig zijn. Gezien de toenemende informatisering, technologische veranderingen en internationalisatie van de economie, is het belangrijk om aandacht te besteden aan de ontwikkeling van adaptiviteit en flexibiliteit in het functioneren.

Het is van belang om te achterhalen welke kennis en vaardigheden (expertises) verder moeten worden uitgebouwd in het licht van het beantwoorden aan de strategische doelstellingen van de afdeling en de organisatie waarin een medewerker werkzaam is. Pas na een gedegen operationalisatie van het concept professionele expertise kan de verdere ontwikkeling van expertise in complexe, dynamische professionele domeinen worden bestudeerd. Theoretische bijdragen vanuit de leertheorie, 'human resource management' en arbeids- en organisatiepsychologie zijn gebruikt voor de conceptualisatie van professionele expertise en om een theorie voor expertise-ontwikkeling te ontwerpen (zie Van der Heijden, 1998; 2000 voor een volledig overzicht van de studie en het meetinstrument voor professionele expertise).

Dr. B.I.J.M. van der Heijden is als Universitair Hoofddocente verbonden aan de vakgroep HRM van de faculteit Technologie en Management, Universiteit Twente. Ze doet onderzoek en geeft onderwijs op het terrein van de loopbaanontwikkeling, in het bijzonder employability. Op dit moment leidt zij onder andere een grootschalig Europees onderzoek naar het bevorderen van de inzetbaarheid van software engineers.

In dit artikel wordt ingegaan op een gedeelte van de genoemde studie, dat wil zeggen de verschillen in inschatting door een medewerker en zijn of haar directe chef ten aanzien van de hoeveelheid kennis en vaardigheden die de individuele medewerker bezit. Het achterhalen van eventuele verschillen is van groot belang gezien het veelvuldig gebruik van de methodologie van 360° feedback. Veelal gaat men uit van de gedachte dat meerdere beoordelingsbronnen tezamen tot een meer betrouwbare uitspraak omtrent iemands functioneren zullen leiden. Vaak zonder de overeenkomsten en verschillen in visie van de verschillende informanten kritisch te beschouwen.

Pas na een gedegen toetsing van de mate waarin zelf-beoordelingen en chef-beoordelingen overeenkomen dan wel afwijken, kan in samenspraak een individueel ontwikkelingsplan worden geformuleerd. Het bewaken van de individuele inzetbaarheid en de verdere loopbaanontwikkeling is pas mogelijk nadat zowel de medewerker als zijn of haar direct leidinggevende overeenstemming kunnen bereiken over de waardering van het huidige kennis- en vaardigheidsarsenaal. Dit laatste kan worden bereikt door beide informanten te confronteren met de discrepanties in de waarnemingen, zoals bijvoorbeeld met behulp van Van der Heijdens (1998) methodiek kunnen worden verzameld.

In deze bijdrage zal ook worden ingegaan op de implicaties die deze discrepanties hebben en worden praktische aanbevelingen voor het management in arbeidsorganisaties behandeld. De betrouwbaarheid van de uitspraken, door de medewerker zelf of door zijn of haar direct leidinggevende, omtrent de deskundigheid van individuele medewerkers vormt de kern in het betoog.

Zelf-beoordelingen versus chef-beoordelingen

De meeste studies over organisationeel gedrag zijn gebaseerd op zelf-beoordelingen. Ondanks de grote populariteit van retrospectieve studies (waarin wordt teruggekeken op een voorbije periode), zijn ook veel onderzoekers tamelijk sceptisch over de resultaten die zijn verzameld met behulp van vragenlijsten waarin mensen vragen over zichzelf en hun prestaties krijgen voorgelegd. Het grootste probleem is dat informanten mogelijk niet in staat zijn om zich een accurate voorstelling van het verleden te maken. Bovendien zijn mensen van nature geneigd om een rooskleuriger beeld van zichzelf te geven. Dat wil zeggen, zelf-beoordelingen lijden aan een inflatie 'bias'.

Sceptici noemen meerdere 'biases' die allen zijn gericht op het beschermen van de mate van zelfwaardering. Zelf-beoordelaars zijn milder ten-

einde zichzelf positief en sociaal wenselijk af te schilderen (Huber en Power, 1985; Tsui en Ohlott, 1988). Deze neiging wordt het 'leniency-effect' genoemd.

Echter, ondanks twijfels ten aanzien van de validiteit van zelf-beoordelingen en het gebruik in 'settings' van toegepast onderzoek, is er reden om aan te nemen dat individuen in een goede positie zijn om een valide inschatting te maken van hun eigen kennis en bekwaamheden. De persoon die de functie uitoefent, is het meest bekend met de baan en lijkt daardoor uitermate geschikt om een vragenlijst over zijn of haar functioneren in te vullen. Studies naar werkgerelateerde variabelen hebben aangetoond dat degenen die de meeste bekendheid met de baan en de beoordeelde bezitten, oordelen verschaffen met een grotere betrouwbaarheid en minder fouten (Miller, 1996). Wie heeft immers een beter begrip van iemands werk en iemands functioneren dan de werknemer zelf?

Het evalueren van werkprestaties houdt in dat de respondent een inschatting moet maken van multikarakteristiek werkgedrag. Theorieën van zelfperceptie geven aan dat een medewerker degene is die toegang heeft tot een brede verzameling van gedragingen onder variabele omstandigheden en tijdsperioden. Bovendien heeft iedere persoon een uniek inzicht in zijn of haar innerlijke toestand, gevoelens en eigenschappen, alsook in de veranderingen over de tijd. Levine (1980, pagina 261) vat de voordelen van zelf-beoordelingen als volgt samen: 'We all have a good deal of time to verify, logically and rationally, our feelings about ourselves and our performance, because we have lived with ourselves since we can remember.'

Naar zijn mening zijn zelf-beoordelingen geschikt om vier redenen. Ten eerste, mensen ervaren direct de effecten van hun acties in een aantal 'real-life' situaties. Ten tweede verzamelen ze informatie over de dingen in de wereld om hen heen door anderen te observeren. Ten derde hebben ze ook beoordelingen van anderen over hun prestaties ontvangen. Tot slot, ze hebben tijd gehad om gevoelens en gedachten over zichzelf te verifiëren. Fox en Dinur (1988) zijn van mening dat individuen de bekwaamheid bezitten om op een betrouwbare manier een zelf-evaluatie te maken op een manier die gelijkend is op die van anderen en op een manier dat toekomstig gedrag kan worden voorspeld.

Overeenkomstig duiden meerdere literatuuroverzichten over zelf-beoordelingen (Mabe en West, 1982; Shrauger en Osberg, 1981) op gunstige evaluaties. Als men zich bewust is van de valkuilen die inherent zijn aan de methodologie, kunnen zelf-beoordelingen zeer waardevol zijn als we proberen complexe fenomenen, zoals bijvoorbeeld professionele exper-

tise, te onderzoeken. Er is een aantal mogelijkheden om de validiteit van zelf-beoordelingen te vergroten.

Mabe en West (1982) hebben gevonden dat de validiteit groter is in studies waarin proefpersonen de verwachting hebben dat hun zelf-beoordeling wordt gevalideerd tegen externe criteria, als ze reeds ervaring hebben met zelf-evaluatie, als ze worden geïnstrueerd dat ze zichzelf moeten vergelijken met anderen en als anonimiteit wordt gegarandeerd. Zij beargumenteren dat individuen worden beïnvloed door twee motieven als ze zichzelf beoordelen: de wens om een *accuraat* beeld van zichzelf te geven en de wens om zichzelf *gunstig* te presenteren.

In de onderhavige studie zijn zowel instructies aangaande 'cross-checking' met de chef-beoordelingen, als anonimiteit, gebruikt als manieren om de validiteit van de resultaten te vergroten. Als zodanig is getracht de voordelen van zelf-beoordelingen te combineren en de ongunstige condities te verminderen door het inbouwen van bepaalde maatregelen. De resulterende onderzoeksopzet is zeer bruikbaar gebleken en gemakkelijk toepasbaar en verschaft een breed gefundeerd inzicht in de mate van professionele expertise van individuele medewerkers.

Professionele expertise; een begripsbepaling

Er is een multidimensionele benadering van het concept professionele expertise gehanteerd. Een dergelijke benadering is nuttig om ten minste drie redenen. Ten eerste is het mogelijk om niveaus van competentie op verschillende domeinen te meten en te vergelijken. Op de tweede plaats is het mogelijk om te achterhalen voor welke dimensies van expertise een verdere uitbouw noodzakelijk is in het licht van het beantwoorden aan de strategische doelen van de afdeling en de organisatie waarin een medewerker werkzaam is. Tot slot is een multidimensionele benadering geschikt voor onderzoek naar de relatie tussen beïnvloedende factoren en bepaalde dimensies van expertise.

In het onderzoek zijn vijf dimensies onderscheiden. Op de eerste plaats de dimensie *sociale erkenning en status*. Expertise kan slechts bestaan bij gratie van erkenning door belangrijke mensen in de organisatie. Al bezit een medewerker nog zo veel kennis en vaardigheden, hij of zij wordt pas als expert gezien op het moment dat belangrijke sleutelfiguren in de organisatie de expertise erkennen.

De tweede dimensie bestaat uit verschillende typen kennis die iemand op een bepaald deskundigheidsgebied bezit. De verschillende typen die

hier worden onderscheiden, zijn: declaratieve kennis ('weten dat'), procedurele kennis ('weten hoe') en conditionele kennis ('weten wanneer en waar') (Alexander, Schallert en Hare, 1991). Deze zogenoemde *kennisdimensie* hangt sterk samen met de derde in de vragenlijst onderscheiden dimensie; de *meta-cognitieve kennisdimensie* ('weten over weten' of 'weten dat men weet'). Er bestaat een grote verscheidenheid aan namen voor dit type kennis: meta-cognitieve kennis, meta-kennis, zelfkennis, regulatieve kennis, meta-cognitieve strategische kennis, enzovoort.

De vierde dimensie bestaat uit de vaardigheden die een medewerker bezit om zijn of haar professionele taken uit te voeren. Als de taken en bijbehorende verantwoordelijkheden van een werknemer eenmaal zijn beschreven, is het duidelijk welke vaardigheden zijn vereist om deze taken en verantwoordelijkheden uit te voeren. Iemand kan slechts een expert worden genoemd als hij of zij de bekwaamheid laat zien, zoals blijkt uit waarneembaar gedrag, om kwalitatief goed te presteren in een bepaald taakdomein of expertisegebied. Deze dimensie wordt hier de *vaardigheidsdimensie* genoemd.

De laatste dimensie heet *groei en flexibiliteit*. Als een medewerker in staat is om op meer dan één domein zich de nodige expertise eigen te maken, komt dat onder andere tot uitdrukking in een hoge score op deze dimensie in de vragenlijst. De flexpert (Van der Heijden, 1996) is in staat om binnen aangrenzende of totaal nieuwe gebieden dan het eigen, vertrouwde werkgebied, kennis en vaardigheden op te doen of hij of zij bezit een strategie om nieuwe expertisegebieden te leren beheersen.

In tabel 1 is voor iedere dimensie van professionele expertise een voorbeeld gegeven van een item uit de vragenlijst.

| Sociale erkenning | Kennisdimensie | Meta-cognitieve kennisdimensie | Vaardigheidsdimensie | Groei en flexibiliteit |
|---|--|---|---|---|
| 'Ik trad in het afgelopen jaar voor ... op als een soort mentor voor het bijbrengen van vaardigheden' | 'Ik beschik over ... feitelijke of vak-technische kennis op mijn gebied' | 'Ik ben ... in staat om ook in complexe situaties het overzicht te bewaren' | 'Ik was in het afgelopen jaar over het algemeen ... in staat om snel beslissingen ten aanzien van de werkaanpak te nemen' | 'Ik acht mezelf ... in staat om flexibel om te gaan met technologische veranderingen' |
| 1 niemand | 1 zeer weinig | 1 zeer slecht | 1 zeer slecht | 1 zeer slecht |
| 2 één collega | 2 tamelijk weinig | 2 tamelijk slecht | 2 tamelijk slecht | 2 tamelijk slecht |
| 3 twee collega's | 3 niet zo veel | 3 niet zo goed | 3 niet zo goed | 3 niet zo goed |
| 4 drie collega's | 4 tamelijk veel | 4 tamelijk goed | 4 tamelijk goed | 4 tamelijk goed |
| 5 vier collega's | 5 veel | 5 goed | 5 goed | 5 goed |
| 6 vijf of meer collega's | 6 zeer veel | 6 zeer goed | 6 zeer goed | 6 zeer goed |

Tabel 1. Voorbeelden van items uit de vijf dimensies van professionele expertise

Het meten van professionele expertise

Het begrip 'professionele expertise' is geoperationaliseerd door middel van een vragenlijst bestaande uit 78 items, elk verwijzend naar een attribuut of gedrag dat wordt toegeschreven aan expertise in verschillende domeinen. Ieder item verwijst tevens naar een bepaalde periode in de professionele loopbaan van de medewerker (het afgelopen jaar) of naar het functioneren in het algemeen.

De vragenlijsten zijn voorgelegd aan zowel medewerkers van minimaal middelbaar functioneringsniveau als aan hun direct leidinggevenden, werkzaam in acht Nederlandse arbeidsorganisaties uit zowel de profit- als non-profitsector (AKZO Nobel, Hewlett Packard, Ministerie van Justitie, Ministerie van Verkeer en Waterstaat, Philips Communication en Processing Services, Rabobank, Unilever Research Laboratorium en Universiteit Twente). Voor de chefs zijn alle items uit de vragenlijst aangepast en omgezet in zinnen zoals bijvoorbeeld: 'Deze medewerker is in staat om ook in complexe situaties het overzicht te bewaren.' De vragenlijsten zijn als zodanig nominaal identiek.

Noch in het geval van de zelf-beoordelingen, noch in het geval van de beoordelingen door de direct leidinggevenden is de volledige range van de zespuntsschalen gebruikt. Deze bevinding wordt toegeschreven aan een zogenaamde centrale antwoordtendentie, hetgeen impliceert dat extreme beoordelingen worden vermeden. In tabel 2 zijn de gemiddelden en standaardafwijkingen (standaarddeviaties) voor de zelf-beoordelingen en chef-beoordelingen gegeven.

| | Totale steekproef medewerkers (N = 558) | | Totale steekproef chefs (N = 454) | |
|------------------------|--|-----------|--------------------------------------|-----------|
| | Gem. | Std. dev. | Gem. | Std. dev. |
| Kennis | 4.32 | .45 | 4.19 | .67 |
| Meta-cognitie | 4.61 | .45 | 4.34 | .67 |
| Vaardigheden | 4.66 | .44 | 4.41 | .70 |
| Sociale erkenning | 4.20 | .52 | 4.06 | .79 |
| Groei en flexibiliteit | 4.07 | .51 | 3.78 | .67 |

Tabel 2. Descriptieve statistieken voor de zelf-beoordelingen en chef-beoordelingen

Omdat van een groot aantal medewerkers zowel hun eigen oordeel als dat van hun direct leidinggevende is verzameld, namelijk 313, kunnen interessante vergelijkingen tussen de twee bronnen van beoordelaars worden gemaakt. Ten behoeve van de volledigheid is in tabel 3 van de paren van beoordelaars vermeld van welke organisatie ze afkomstig zijn.

| Organisatie | Aantal en responsepercentage van de paren |
|--|---|
| Akzo Nobel | 56 (17.9%) |
| Hewlett Packard | 35 (11.2%) |
| Ministerie van Justitie | 40 (12.8%) |
| Ministerie van Verkeer en Waterstaat | 40 (12.8%) |
| Philips Communications and Processing Services | 35 (11.2%) |
| Rabobank Organisatie | 64 (20.4%) |
| Unilever Research Laboratorium | 30 (9.6%) |
| Universiteit Twente | 13 (4.2%) |
| Totaal: 313 paren | |

Tabel 3. Oorsprong van de paren van beoordelaars

Alvorens het hoofdonderzoek op te starten zijn door middel van een pilotstudie de psychometrische kwaliteiten van de samengestelde vijf meetschalen oftewel dimensies van expertise onderzocht. Zowel de uitkomsten van de pilotstudie als die van het hoofdonderzoek wijzen uit dat een betrouwbaar en valide meetinstrument is ontwikkeld. Zie Van der Heijden (1998, 2000) voor een uitgebreide verhandeling omtrent de psychometrische kwaliteiten van het instrument en de methodologie die is gehanteerd om een en ander te bepalen.

Het effect van beoordelaar op inschattingen van de mate van professionele expertise

Men moet zich realiseren dat de medewerkers die deel hebben genomen aan het onderzoek, in de hogere regionen van de organisatie werkzaam zijn. Op deze niveaus respecteren zowel medewerkers als hun direct leidinggevenden dat ze separate verantwoordelijkheden hebben en zien en spreken ze elkaar vaak slechts incidenteel. Bovendien is uit eerder onderzoek (Boerlijst, Van der Heijden en Van Assen, 1993) gebleken dat de periode gedurende welke medewerker en chef met elkaar te maken hebben, relatief kort is. Leidinggevenden veranderen doorgaans om de drie à vier jaar van afdeling. Dientengevolge is er een gebrek aan interesse in de verdere ontwikkeling van de medewerker van de kant van het management. Een aanzienlijke kloof tussen de gezichtspunten van beide partijen is dan ook niet verwonderlijk.

Hypothese 1: De correlaties tussen zelf-beoordelingen en de corresponderende chef-beoordelingen zijn tamelijk laag.

Over het algemeen verschillen zelf-beoordelingen een halve tot een hele standaardafwijking van de beoordelingen door naaste collega's of chefs (Arnold en MacKenzie Daveys, 1992). Mensen zijn geneigd om zichzelf in een positief licht af te schilderen. Deze neiging, welke het 'leniency-effect' wordt genoemd, is eerder in het artikel al uiteengezet.

Campbell en Lee (1988) schrijven het verschil in zelf-beoordelingen en chef-beoordelingen toe aan: (1) verschillende percepties van de baan en dienstengevolge verschillende beoordelingscriteria; (2) verschillende cognitieve processen en (3) verschillende affectieve processen (bijvoorbeeld de neiging om zichzelf positief te presenteren, zelfontkenning en projectie).

Hypothese 2: Zelf-beoordelingen zijn systematisch hoger in vergelijking met chef-beoordelingen.

In tabel 4 worden de uitkomsten van de paarsgewijze vergelijkingen gepresenteerd. Als eerste worden het gemiddelde en de standaardafwijking gepresenteerd. Vervolgens worden de resultaten van de paarsgewijze T-toetsen en de paarsgewijze correlaties gegeven.

| | N | Gemiddelde | Standaarddeviatie | t | Significantie (eenzijdig) | Correlatie | Significantie (tweezijdig) |
|-------------------------------------|-----|------------|-------------------|------|---------------------------|------------|----------------------------|
| Paar 1 | | | | | | | |
| – Kennis medewerker | 313 | 4.32 | .45 | 2.31 | p<.05 | .28 | .000*** |
| – Kennis chef | | 4.23 | .67 | | | | |
| Paar 2 | | | | | | | |
| – Meta-cognitie medewerker | 313 | 4.63 | .45 | 5.57 | p<.001 | .09 | n.s. |
| – Meta-cognitie chef | | 4.39 | .67 | | | | |
| Paar 3 | | | | | | | |
| – Vaardigh. medewerker | 313 | 4.69 | .44 | 5.66 | p<.001 | .23 | .000*** |
| – Vaardigh. chef | | 4.46 | .70 | | | | |
| Paar 4 | | | | | | | |
| – Sociale erkenning medewerker | 313 | 4.24 | .52 | 2.85 | p<.01 | .31 | .000*** |
| – Sociale erkenning chef | | 4.11 | .79 | | | | |
| Paar 5 | | | | | | | |
| – Groei en flexibiliteit medewerker | 313 | 4.08 | .51 | 7.08 | p<.001 | .27 | .000*** |
| – Groei en flexibiliteit chef | | 3.81 | .67 | | | | |

(n.s. = niet significant, * = p<.05, ** = p<.01, *** = p<.001)

Tabel 4. Paarsgewijze T-toetsen en paarsgewijze correlaties

Op één uitzondering na hangen de zelf-beoordelingen en chef-beoordelingen voor alle paren significant met elkaar samen. De correlatie ligt steeds tussen $r = .23$ en $.31$. Dit gegeven is van belang voor de validiteit, oftewel de 'juistheid' van het meetinstrument (Cronbach, 1990). Bovendien is deze uitkomst in overeenstemming met de resultaten uit vergelijkbare studies. Mabe en West (1982) verschaften een excellent overzicht en een meta-analyse van de validiteit van zelf-beoordelingen van bekwaamheid.

De uitzondering betreft de corresponderende beoordelingen op de meta-cognitieve kennisdimensie die een correlatie hebben van $.09$. Eerder is reeds gesteld dat indien de hypothese omtrent de relatieve onwetendheid van chefs aangaande de 'ups en downs' van hun medewerkers klopt, lage correlaties kunnen worden verwacht. Ze zijn duidelijk laag. Hypothese 1 wordt met deze uitkomsten bevestigd.

Voortbordurend op de ongerelateerdheid van de beoordelingen op de meta-cognitieve kennisdimensie, zou een mogelijke verklaring kunnen liggen in het feit dat meta-cognitie een attribuut is met een sterke innerlijke gerichtheid en zelf-reflectieve lading. Het is wellicht dié persoonsgebonden factor die het moeilijkst op een valide manier is in te schatten. Chefs, hoewel verantwoordelijk, zijn vaak te veel op een afstand om hun medewerkers meer dan oppervlakkig te kennen. Dit kan de reden zijn van het feit dat meta-cognitie de dimensie is die het moeilijkst in te schatten is. Uiteraard is verder onderzoek nodig om nader uit te zoeken of medewerkers in dit geval inderdaad beter zijn uitgerust om hier een oordeel over uit te spreken. Zelfreflectie is namelijk ook geen bezigheid die mensen geneigd zijn veel uit te oefenen als het ingewikkelde cognitieve vaardigheden en gedragsuitkomsten betreft.

De paarsgewijze T-toetsen bevestigen dat, voor iedere schaal, de zelfbeoordelingen systematisch hoger liggen dan de corresponderende chefbeoordelingen. Met andere woorden, medewerkers denken beter over zichzelf dan hun direct leidinggevenden dat doen, of ze geven op z'n minst een positiever beeld van zichzelf. Dit betekent dat hypothese 2 ook wordt bevestigd.

'Multitrait-Multimethod'-analyse

Voor het bepalen van betrouwbaarheids- en validiteitsaspecten van de expertisemetingen, is een 'Multitrait-Multimethod'-analyse (MTMM) gemaakt (zie Van der Heijden (1998, 2000) voor een volledige uitwerking van deze benadering). Omdat medewerkers op twee manieren zijn beoordeeld, namelijk door zichzelf en door hun chef, is het mogelijk om deze benadering te hanteren.

Op de eerste plaats zijn de schaalbetrouwbaarheden aan de hand van Cronbachs α bepaald voor iedere dimensie én voor de twee beoordelingsmethoden. In tabel 5 zijn de betrouwbaarheidscoëfficiënten gegeven.

| Schaal | Items | Alfa zelf (N = 558) | Alfa chef (N = 454) |
|------------------------|-------|---------------------|---------------------|
| Kennis | 17 | .83 | .93 |
| Meta-cognitie | 15 | .86 | .94 |
| Vaardigheden | 12 | .84 | .94 |
| Sociale erkenning | 15 | .83 | .94 |
| Groei en flexibiliteit | 19 | .87 | .93 |

Tabel 5. Betrouwbaarheidscoëfficiënten volgens Cronbachs α

Alle vijf schalen blijken zeer betrouwbaar te zijn. Het is interessant om te observeren dat de α 's voor de chefs allen hoger zijn dan de corresponderende α 's voor de medewerkers. Deze uitkomst maakt het noodzakelijk om nog wat verder te reflecteren op wat dit zou kunnen betekenen.

Het feit dat de betrouwbaarheden voor de vijf dimensies voor de medewerkers lager zijn (dat wil zeggen lagere inter-item correlatiecoëfficiënten hebben) wil niet per se zeggen dat medewerkers een minder valide beeld van zichzelf geven, in vergelijking met leidinggevend, maar een meer gedifferentieerd zelfbeeld. Ieder item verwijst immers naar verschillend gedrag, een verschillende gedragsstijl of een verschillende uitkomst.

Een andere argumentatie is dat de chef-beoordelingen sterk worden beïnvloed door een 'response set,' bijvoorbeeld de mate waarin ze de medewerker in kwestie mogen als persoon. Het is bekend dat dergelijke factoren een belangrijke rol spelen op het moment dat van iemand wordt verwacht dat hij of zij een betrouwbaar en valide oordeel geeft van iemand anders, maar eigenlijk niet voldoende op de hoogte is om zijn of haar oordeel te baseren op 'real-life' observaties van die persoon, of op relevante 'real-life' gebeurtenissen.

Evaluatie van een persoon, louter op basis van één enkel attribuut, of het nou een gunstig of ongunstig attribuut is, wordt het *halo-effect* genoemd. Met andere woorden, een 'halo' verblindt de waarnemer voor andere attributen die ook zouden moeten worden meegenomen, om een complete en accurate indruk van de andere persoon te krijgen. Eén enkel kenmerk wordt centraal gesteld en gebruikt als basis voor het beoordelen van alle andere prestatiemetingen. Het zou bijvoorbeeld kunnen gebeuren dat een laag verzuimpercentage van een persoon kan leiden tot een hoge beoordeling voor de mate van productiviteit, kwaliteit en nauwgezetheid van het werk – afgezien van het feit of dit daadwerkelijk het geval is.

Empirische studies hebben aangetoond dat een 'halo-effect' minder prominent is bij zelf-beoordelingen dan bij het beoordelen van anderen (Hoffman, Nathan en Holden, 1991). Verder blijkt een 'halo-effect' meer prominent aanwezig bij chef-beoordelingen dan bij beoordelingen door collega's of ondergeschikten (Tsui en Ohlott, 1988). Het effect van 'undersampling' kan mogelijk een verklaring zijn. Chef-beoordelingen zijn relatief gezien op een kleinere hoeveelheid informatie gebaseerd (Vogelaar, De Vries en Kuipers, 1994).

Bovendien worden de chef-beoordelingen voor een bepaalde dimensie, over het algemeen, meer beïnvloed door beoordelingen in een andere dimensie. Cooper (1981) noemt dit verschijnsel ‘engulfing’, hetgeen in dit geval betekent dat de beoordelingen voor een bepaalde dimensie van expertise de beoordelingen van een andere dimensie kleuren. De mogelijkheden om medewerkers te observeren zijn beperkt, hetgeen inhoudt dat het oordeel op minder informatie is gebaseerd vergeleken met de zelf-beoordelingen. Managers moeten hun beoordelingen baseren op globale werkresultaten of op incidenten die toevallig onder hun aandacht komen.

Het voorgaande brengt ons tot de laatste hypothese:

Hypothese 3: De inter-schaalcorrelaties voor de chef-beoordelingen zijn systematisch hoger vergeleken met de inter-schaalcorrelaties voor de zelf-beoordelingen.

In de volgende twee tabellen, 6 en 7, worden de patronen van de correlaties voor beide typen beoordelaars gegeven.

| | Kennis | Meta-cognitieve kennis | Vaardigheden | Sociale erkenning | Groei en flexibiliteit |
|------------------------|--------|------------------------|--------------|-------------------|------------------------|
| Kennis | .83 | .71 | .68 | .78 | .77 |
| Meta-cognitieve kennis | | .86 | .75 | .65 | .67 |
| Vaardigheden | | | .84 | .72 | .65 |
| Sociale erkenning | | | | .83 | .72 |
| Groei en flexibiliteit | | | | | .87 |

(Voor alle correlaties $p < .001$) (tweezijdig)

Tabel 6. Inter-schaalcorrelaties zelf-beoordelingen (N = 558)

| | Kennis | Meta-cognitieve kennis | Vaardigheden | Sociale erkenning | Groei en flexibiliteit |
|------------------------|--------|------------------------|--------------|-------------------|------------------------|
| Kennis | .93 | .87 | .87 | .91 | .86 |
| Meta-cognitieve kennis | | .94 | .87 | .85 | .81 |
| Vaardigheden | | | .94 | .89 | .80 |
| Sociale erkenning | | | | .94 | .86 |
| Groei en flexibiliteit | | | | | .93 |

(Voor alle correlaties $p < .001$) (tweezijdig)

Tabel 7. Inter-schaalcorrelaties chef-beoordelingen (N = 454)

Als men de inter-schaalcorrelaties van beiden bronnen vergelijkt, kan men zien dat de correlaties tussen de verschillende schalen voor de chef-beoordelingen systematisch hoger liggen vergeleken met die van de zelf-beoordelingen. Het lijkt zo te zijn dat chefs meer consistent zijn

in hun antwoorden. Als ze van mening zijn dat hun medewerkers bijvoorbeeld veel kennis bezitten, volgen hun oordelen met betrekking tot de andere dimensies van expertise meer of minder logisch dit patroon.

Dit resultaat is in overeenstemming met hypothese 3 betreffende het verschil in inter-schaalcorrelaties voor beide typen beoordelaars. De sterkte van het 'halo-effect' is groter als het chef-beoordelingen betreft (Hoffman, Nathan en Holden, 1991). De antwoorden van de medewerkers volgen deze neiging naar consistentie in mindere mate. Dat wil zeggen dat de schalen voor hen meer logisch onafhankelijk zijn.

Conclusies en discussie

We veronderstellen dat de aard van de beoordelaar – de medewerker of de chef – sterk verschillende informatie over de medewerker oplevert. De correlaties tussen de zelf-beoordelingen en de chef-beoordelingen zijn tamelijk laag. Bovendien is het verschil tussen de twee typen beoordelingen significant, dat wil zeggen medewerkers denken positiever over zichzelf dan hun direct leidinggevenden dat doen.

De enige verklaring die we op dit moment kunnen geven, is het gebrek aan communicatie tussen de twee partijen. De afstand tussen de medewerker en zijn of haar direct leidinggevende is groot. Op de hogere functioneringsniveaus in het bijzonder, werkt een medewerker zeer onafhankelijk en vaak zelfs solitair. Teneinde meer inzicht te krijgen in het potentieel van de afdeling zouden direct leidinggevenden moeten proberen om actief de kloof tussen henzelf en hun medewerkers te verkleinen. Door regelmatige werkbijeenkomsten, functioneringsgesprekken en liefst ook het invoeren van potentieelbeoordelingen gaat het praten over individuele bekwaamheden en tekortkomingen van beide partijen tot een vast onderdeel van de HRM-activiteiten behoren.

Verder onderzoek is nodig om de communicatiepatronen tussen individuele medewerkers en direct leidinggevenden beter te begrijpen en om te bepalen in welke mate ze de oordelen over de mate van professionele expertise beïnvloeden. Zoals al eerder aangegeven, zou het verschil voor een (aanzienlijk) deel kunnen worden verklaard door het zogenoemde 'halo-effect', dat door het gebrek aan inzicht in het feitelijke functioneren de waarneming door de direct leidinggevende in sterke mate kan kleuren.

Hoewel het meetinstrument bruikbaar is om uit te zoeken waar expertise en expertgedrag reeds aanwezig zijn in de organisatie en waar ze

verbeterd moeten worden, is het niet onverstandig om te proberen de validiteit te verbeteren. Een interessante verbetering zou kunnen bestaan uit onderzoek waarin aan beoordelaars wordt gevraagd om concrete voorbeelden te geven van prestaties of gedragingen. Dergelijke voorbeelden kunnen wellicht illustreren waarom iemand een bepaald oordeel geeft naar aanleiding van een bepaald item uit de vragenlijst. We verwachten dat 'response sets', zoals het 'halo-effect', hierdoor op zijn minst minder voor zullen komen. De differentiatie tussen de betekenissen van de items zal toenemen en de homogeniteit van de schalen zal verminderen, maar de beoordelingen zullen wel explicieter zijn gefundeerd op empirisch toetsbare observaties van gedrag en prestaties.

Als dit succesvol verloopt, zal de validiteit van het instrument toenemen en meer inzichtelijk worden. De set van demonstratieve voorbeelden die door de beoordelaars wordt gegeven, kan wellicht ook meer inzicht geven in de intrinsieke betekenis van expertise en expertgedrag, zoals ze in arbeidsorganisaties voorkomen. Het zou de aanzet kunnen zijn van een databestand van expertise.

Implicaties en praktische aanbevelingen

Op de hogere functioneringsniveaus is de correlatie tussen zelf-beoordelingen en beoordelingen door de direct leidinggevenden laag. Het feit dat de verschillende typen beoordelaars veelal sterk verschillende informatie verschaffen over de beoordeelde impliceert dat het toepassen van de methodiek van 360° feedback als HRM-instrument in geen geval mag inhouden dat de beoordelingen door de verschillende partijen worden opgeteld (zie ook Van der Heijden, 1998).

De zelf-beoordelingen en de beoordelingen door de chef moeten veeleer naast elkaar worden beschouwd om te bestuderen welke perceptieverschillen er bestaan en welke gevolgen ze hebben voor het gangbare personeelsmanagement. De aard van de beoordelaar – medewerker of chef – maakt dat sterk verschillende informatie wordt bovengehaald vanwege een gebrek aan communicatie tussen beide partijen. Op hogere functioneringsniveaus wordt veelal zeer zelfstandig en solitair gewerkt en moet nog een aanzienlijke kloof worden overbrugd, wil de chef gedegen inzicht in het functioneren van zijn medewerkers krijgen.

Het blijkt bijvoorbeeld dat nogal wat chefs slecht op de hoogte zijn van de netwerken en informele circuits waarin hun medewerkers zich bewegen. Voorzover ze dat wel zijn, betreft het vooral de netwerken die chefs zelf van belang achten. Het is van belang dat chefs een actieve

belangstelling tonen voor de (informele) netwerken waaraan hun medewerkers deelnemen. Deze netwerken kunnen namelijk voor de betreffende medewerker een rijkdom aan ervaringen betekenen en dus een uitbreiding van hun competenties.

Indien de beoordelaar beter op de hoogte is van en meer inzicht heeft in het arsenaal aan kennis en vaardigheden van zijn of haar ondergeschikten, zal hij of zij minder vatbaar zijn voor bijvoorbeeld het 'halo-effect'. Een mogelijkheid om beoordelingssituaties in deze te verbeteren kan worden gezocht in trainingsprogramma's waarin zowel direct leidinggevenden als hun ondergeschikten participeren en waarin ze onder een kritische begeleiding feedback uitspreken over elkaars functioneren. Indien men waarderingen openlijk bespreekbaar maakt, kunnen mogelijke foutieve veronderstellingen aan het licht worden gebracht en worden beide partijen gedwongen om op feiten gebaseerde uitspraken te doen ten aanzien van de concrete inbreng in de betreffende afdeling.

In veel situaties is het noodzakelijk om de inzichten en opinies van het management te corrigeren, omdat veel foutieve beoordelingen over medewerkers mede zijn toe te schrijven aan 'verkeerde' of 'onjuiste' opvattingen van hun chefs (Boerlijst, Van der Heijden en Van Assen, 1993). Opvattingen die niet kloppen met de feiten. Als chefs zich in hun gedrag en reacties bijvoorbeeld laten leiden door de opvatting dat de bijdrage van oudere werknemers heel sterk afneemt met de leeftijd, valt te verwachten dat de beoordeling over dergelijke werknemers negatiever zal zijn. Medewerkers die niet aan het 'ideale' stereotype beantwoorden, zal men in eerste instantie in de gewenste richting trachten te sturen of te beïnvloeden. Als het management weinig vertrouwen heeft in de zin van dergelijke acties, zal het de betreffende groepering eerder links laten liggen of terzijde schuiven. De invloed van leeftijd is een belangrijke variabele in de studie van Van der Heijden (1998) en is uitgebreid beschreven in een ander artikel (Van der Heijden, in druk).

De houding en opstelling van het management ten opzichte van bepaalde categorieën die wel of niet aan bepaalde ideaalvoorstellingen voldoen, zijn weliswaar te begrijpen, maar niet altijd te rechtvaardigen. Het stereotiepe beeld dat men van bepaalde groepen in de organisatie of onderneming heeft – en dat geldt natuurlijk ook voor relevante groepen daarbuiten – is niet altijd 'juist' in de zojuist omschreven zin. Bovendien is het onverstandig om voorbij te gaan aan individuele verschillen, omdat organisaties het juist moeten hebben van de bijdragen van individuele medewerkers. Stereotypering leidt in de praktijk van het management tot collectivisatie waarbij de individualiteit van medewerkers verloren gaat. Dit betekent vaak dat zij zichzelf ook niet meer als individu betrokken of aangesproken voelen.

In het personeelsbeleid zou aandacht moeten bestaan voor het duidelijk krijgen van eventuele stereotypen die in de organisatie worden gebezigd, en welke beleids- en actieconsequenties men daaraan verbindt. Dergelijke stereotypen kunnen een gewenst beleid bevorderen of belemmeren, met name een beleid dat is gericht op het behoud en de ontwikkeling van flexibiliteit en creativiteit van mensen in de organisatie. De 'juistheid' of 'genuanceerdheid' van dergelijke stereotypen moeten worden geverifieerd aan feitelijke en objectief te controleren waarnemingen omtrent deze categorieën. Bovendien moeten degenen die een te somber beeld hebben ten aanzien van de deskundigheid van hun ondergeschikten, worden geconfronteerd met het beeld dat de medewerkers van zichzelf hebben. Eventueel kunnen in het kader van een 360° beoordelingsmethodiek ook waarderingen door collega's en klanten worden meegenomen.

Vervolgens zullen het beleid en de acties welke uit de beoordelingen voortvloeien, met name op hun merites moeten worden bekeken, zo nodig worden herzien en bijgestuurd. Het is aanbevelenswaardig om de door chefs en andere leden van het management gebezigde stereotypen kritisch met hen te bespreken. In deze gesprekken moet men niet alleen over de inhoud en feitelijke (on)juistheid van de waarnemingen spreken, maar ook over de effecten en over het gedrag dat zij kunnen uitlokken bij chefs én bij de betreffende medewerkers. Een deskundige op het gebied van Human Resource Management en communicatieve vaardigheden is in deze een pré.

Deze gesprekken en het inzamelen van informatie over individuele medewerkers zijn bovendien activiteiten die wegbereiders zijn voor iemands mobiliteit als ze daar ook betrekking op hebben en als ze ook daadwerkelijk goed worden uitgevoerd. Het is van groot belang om te beseffen dat de, vaak grote hoeveelheid, informele informatie die veel chefs over hun medewerkers 'achter de hand' kunnen hebben, in het kader van een 'human resources'-beleid van elementair belang is.

Samengevat kan worden gesteld dat het in sterke mate de vraag is in hoeverre chefs een gedegen beeld hebben van de mogelijkheden van hun menselijk potentieel en in hoeverre hun inschatting van de mate van professionele expertise berust op betrouwbare informatie. De loopbaan en de ontwikkeling van de medewerker vormen in veel gevallen geen onderwerp van gesprek met de chef. Het gevolg is dat de gemiddelde chef weinig van de medewerker en zijn of haar loopbaan weet en dus weinig met die gebrekkige kennis kan doen. Dit heeft uiteraard gevolgen voor de beoordelingspraktijk.

Zoals eerder aangegeven, zouden er veel nauwere contacten over deze zaken moeten zijn tussen de medewerker en zijn of haar chef (Stoker en Van der Heijden, in druk). Niet minder belangrijk zijn de contacten tussen de chef en andere echelons van het management om ook de strategische aspecten van het ontwikkelingsmanagement tot hun recht te laten komen. Op deze manier krijgen ook andere delen van het management inzicht in de individuele 'human resources' die een organisatie in huis heeft. Dit laatste is met name van belang met het oog op 'onderhoud' van individuele functies en van de sociale infrastructuur.

S U M M A R Y

In this contribution the outcomes of a study regarding the differences between self-ratings and corresponding supervisor ratings on five different scales of professional expertise are discussed. The results show that employees think somewhat better of themselves than supervisors do, or at any rate they give a rosier image. The tendency to present themselves positively is called the *leniency effect*, common in performance ratings.

The scale reliabilities are very high both for the self-ratings and for the supervisor ratings. Those made by the supervisor are all higher than the corresponding ones made by their employees. It might be that the self-ratings reflect a reliable, but somewhat more differentiated self-image.

In addition to the difference in scale reliabilities, another discrepancy between the two types of ratings concerns the fact that the correlations between the five scales for the supervisor ratings are systematically higher compared with the ones by the employees. The supervisors are more consistent in answering. That is to say, the ratings made by supervisors are more coloured by the ratings in another dimension.

After a presentation of the outcomes of the study, implications and recommendations for personnel management are given.

Literatuurlijst

- Alexander, P.A., D.L. Schallert en V.C. Hare (1991). 'Coming to terms: How researchers in learning and literacy talk about knowledge'. In: *Review of Educational Research*, 61(3), pp. 315-344.
- Arnold, J. en K. Mackenzie Daveys (1992). 'Self-ratings and supervisor ratings of graduate employees' competences during early career'. In: *Journal of Occupational and Organizational Psychology*, 65, pp. 235-250.

- Boerlijst, J.G., B.I.J.M. van der Heijden en A. van Assen (1993). *Veertig-plussers in de onderneming*. Assen: Van Gorcum/Stichting Management Studies.
- Campbell, D.J. en C. Lee (1988). 'Self-appraisal in performance evaluation: Development versus evaluation'. In: *Academy of Management Review*, 13(2), pp. 302-314.
- Cooper, W. (1981). 'Ubiquitous halo'. In: *Psychological Bulletin*, 90, pp. 218-244.
- Cronbach, L.J. (1990). In: *Essentials of psychological testing*. New York: Harper Collins Publishers.
- Fox, S. en Y. Dinur (1988). 'Validity of self-assessment: a field evaluation'. In: *Personnel Psychology*, 41, pp. 581-592.
- Hoffman, C.C., B.R. Nathan en L.M. Holden (1991). 'A comparison of validation criteria: objective versus subjective performance measures and self- versus supervisor ratings'. In: *Personnel Psychology*, 44, pp. 601-619.
- Huber, G.P. en D.J. Power (1985). 'Retrospective reports of strategic level managers: Guidelines for increasing their accuracy'. In: *Strategic Management Journal*, 6, pp. 171-180.
- Levine, E.L. (1980). 'Introductory remarks for the symposium Organizational applications of self-appraisal and self-assessment: Another look'. In: *Personnel psychology*, 33, pp. 259-262.
- Mabe, P. en S. West (1982). 'Validity of self-evaluation of ability: a review and meta-analyses'. In: *Journal of Applied Psychology*, 67, pp. 280-296.
- Miller, M.J. (1996). *Rater individual differences and accuracy in performance appraisal*. A thesis submitted to the Office of Graduate Studies of Texas A&M University.
- Shrauger, S.J. en T.M. Osberg (1981). 'The relative accuracy of self-prediction and judgments by others in psychological assessment'. In: *Psychological Bulletin*, 90, pp. 322-351.
- Stoker, J.I. en B.I.J.M. van der Heijden (in druk). 'Self-perceptions of employees and supervisors and their consequences for competence development'. In: *Journal of Career Development*.
- Tsui, A.S. en P. Ohlott (1988). 'Multiple assessment of managerial effectiveness: interrater agreement and consensus in effectiveness models'. In: *Personnel Psychology*, 41, pp. 779-803.
- Van der Heijden, B.I.J.M. (1996). *Life-long expertise development: goal of the nineties*. Book of Proceedings of the Fifth Conference on International Human Resource Management. Hyatt Islandia San Diego, California, USA, 24-28 June, 1996.
- Van der Heijden, B.I.J.M. (1998). *The measurement and development of professional expertise throughout the career. A retrospective study among higher level Dutch professionals*. Ph.D.-thesis. University of Twente, Enschede, the Netherlands. Enschede: PrintPartners Ipskamp.
- Van der Heijden, B.I.J.M. (2000). 'The development and psychometric evaluation of a multidimensional measurement instrument of professional expertise'. In: *High Ability Studies. The Journal of the European Council for High Ability*, 11(1), 9-39.

- Van der Heijden, B.I.J.M. (in druk). 'Age and assessments of professional expertise. The relationship between higher level employees' age and self-assessments or supervisor ratings of professional expertise'. In: *International Journal of Selection and Assessment*.
- Vogelaar, A.L.W., R.E. de Vries en H. Kuipers (1994). 'Beoordelingen van prestaties en persoonskenmerken: Halo, 'leniency' en overeenstemming [Performance and personality ratings: Halo, leniency, and agreement]'. In: *Gedrag en Organisatie*, 7(3), 129-144.