

Beoordelen onder topprofessionals: een mijnenveld

Dr. ir. M. Wanrooy is senior adviseur op het gebied van bestuur, leiderschap en conflict-begeleiding bij HRD-adviesbureau GITP.

Het wordt steeds normaler dat topprofessionals zoals rechters, medisch specialisten, hoogleraren en partners bij advocaten- en accountantskantoren worden beoordeeld op hun presteren. Maar de weg is hobbelig en vol valkuilen. Waarom is beoordelen onder professionals zo gevoelig, en wat is de beste aanpak?

Het evalueren van topprofessionals heeft een grote vlucht genomen. In dit artikel wordt onderzocht welke ontwikkelingen daaraan ten grondslag liggen en welke veranderingen zich daarbij hebben voorgedaan. Daarna komen dilemma's en gevoeligheden aan de orde en wordt beschreven welke beoordelingssystemen er zijn. Ten slotte wordt besproken welke valkuilen zich voordoen en hoe die kunnen worden vermeden. Dit artikel is bedoeld voor iedereen die met beoordelen onder topprofessionals te maken heeft en zich afvraagt hoe dit beter kan.

Urgentie neemt toe

Nog maar tien jaar geleden werden topprofessionals vrijwel niet beoordeeld. Had je eenmaal de positie bereikt van partner, hoogleraar of rechter, dan werd je verder met rust gelaten. Je status was betrekkelijk onaantastbaar en de onderlinge code was dat men elkaars domein respecteerde en gevoeligheden uit de weg ging. Eventuele fouten werden gedoogd of gecamoufleerd. Natuurlijk was men – zoals het goede professionals betaamt – uiterst kritisch naar elkaar, maar men sprak elkaar niet aan en reputaties werden beschermd. Die periode ligt achter ons, maar nog steeds is men uiterst omzichtig als het om beoordelen of feedback geven gaat; kritiek wordt nog weinig openlijk gedeeld en het onvoldoende functioneren van collega's wordt niet aangepakt of wordt afgeschermd. Dat kan men zich echter steeds minder permitteren; de druk om topprofessionals te beoordelen neemt in alle genoemde sectoren toe.

Dat is het gevolg van een aantal ontwikkelingen:

- Het gezag van de topprofessional neemt af. De witte jas van de medicus is geen garantie voor zijn of haar deskundigheid. Medische missers komen geregeld in het nieuws en ook rechters, advocaten of onderzoekers blijken niet onfeilbaar. Een prominent onderzoeker valt van zijn voetstuk vanwege grootschalige fraude, een beroemde stafpleiter wordt publiekelijk aan de schandpaal genageld. De voorbeelden spreken voor zich.
- De buitenwereld eist transparantie en wenst dat de van oorsprong gesloten beroepsgroepen verantwoording afleggen. Daar hoort ook bij dat men getoetst wordt op kwaliteit, en daar is meer voor nodig dan jaarlijks het verplichte aantal bijscholingspunten halen. De AFM laat geregeld zijn tanden zien. Er wordt extern toezicht in de advocatuur ingesteld. Accreditatie en *peer reviews* worden scherper, ook omdat in het (semi)publieke domein veel gemeenschapsgeld wordt besteed en er kritisch wordt gekeken naar doelen en middelen. Er wordt levendig gediscussieerd over de vraag of de rechterlijke macht zich meer in de pers moet laten zien – of juist niet.
- De druk van onderaf neemt toe. De jongere generatie pikt het niet langer dat de 'gearriveerde' toppers niet worden beoordeeld terwijl zij zelf een zwaar traject moeten doorlopen voordat ze tot de elite kunnen toetreden. Ook topprofessionals zelf geven toe dat het vreemd is dat in alle lagen onder hen jaarlijks beoordelings-

gesprekken worden gevoerd maar dat er voor henzelf relatief weinig is geregeld.

- Organisaties kunnen zich steeds minder veroorloven dat bepaalde collega's niet aan de maat presteren. Niet alleen de externe roep om verantwoording speelt hierbij een rol. Aansprakelijkheid en reputatieschade liggen op de loer. Bezuinigingen zetten de zaak op scherp. Het economisch tijt zorgt voor teruglopende resultaten en professionals verlangen van elkaar hoge inzet en prestaties. Dat geldt op meerdere gebieden tegelijk: professie, commercie (acquisitie) en leidinggeven. Sommigen – en niet alleen ouderen – houden die topsport niet vol.¹ Steeds vaker wordt de behoefte gevoeld aan een evaluatiesysteem, om signalen tijdig te bespreken, onderpresteren aan te pakken en uitval te voorkomen.
- Conflicten onder topprofessionals komen veel voor.² Ze belasten de onderlinge verhoudingen, bederven het werkplezier en kosten veel tijd en geld. Vaak ligt (vermeend) disfunctioneren ten grondslag aan conflicten. Geregelde evaluatie werkt preventief; dat besef lijkt steeds vaker door te dringen.³

Veranderingen

In de diverse sectoren is intussen al veel veranderd. Het tijdperk van 'met rust gelaten worden' is voorbij. Hoogleraren, bijvoorbeeld, worden tegenwoordig in eerste instantie voor drie of vijf jaar benoemd. Zij moeten zich dus telkens opnieuw bewijzen om in aanmerking te komen voor een herbenoeming. Dat geldt ook voor directeurs van onderzoeksinstituten, zoals de KNAW: alleen op basis van een evaluatie (mèt 360°-feedback uit de directe werkomgeving) worden zij voor een nieuwe periode benoemd. Ook tussendoor voeren toponderzoekers jaarlijkse resultaat- en ontwikkelingsgesprekken met hun afdelingsvoorzitter, decaan of directeur. Overigens zijn de periodieke *peer reviews* in de wetenschap ook vaak aanleiding om – zij het impliciet – het functioneren van toponderzoekers ter discussie te stellen.

Bij de meeste grote accountants- en advocatenkantoren is een vorm van partnerevaluatie ingesteld, soms rudimentair, soms zeer uitgebreid. Veelal is de implementatie gepaard gegaan met heftige discussie over de belangrijke vragen die hierbij onvermijdelijk opkomen. Welke criteria past men toe? Hoe wordt gemeten? Wie beoordeelt wie? Geregeld worden aanpassingen gedaan om de systemen verder te optimaliseren. In de

kleine en middelgrote kantoren is overigens van een systematische partnerevaluatie nog lang geen sprake.⁴

In de medische sector is sinds 2008 het IMFS (Individueel Functioneren van Medisch Specialist) beschikbaar. Dat is een vast protocol met behulp waarvan medici elkaar kunnen evalueren. Het IMFS wordt nog niet algemeen toegepast⁵ en niet alle facetten van het functioneren kunnen worden getoetst. Maar vrijwel ieder ziekenhuis is bezig met een of andere vorm van evaluatie van medisch specialisten. Er zijn maatschappen die naast of in plaats van het IFMS een eigen systeem van jaargesprekken hebben ingevoerd.

Leden van de rechterlijke macht zijn voor het leven door de Kroon benoemd. Hun autonomie is verregaand beschermd om hun rechterlijke onafhankelijkheid te waarborgen. In dat domein is het beoordelen van professionals een precaire aangelegenheid, want 'vonnissen wijzen' kan en mag niet door hiërarchische hoger geplaatsten worden beoordeeld. Toch worden ook hier functioneringsgesprekken steeds meer op geformaliseerde wijze gevoerd. Het laat zich raden dat discussies hierover breed gevoerd worden, want men beschouwt het als een inperking van het beroepsdomein. Er is echter consensus over het feit dat rechters op passende wijze geëvalueerd moeten worden in een jaarlijks gesprek met hun afdelingsvoorzitter of president.

Wat maakt beoordeling van topprofessionals zo gevoelig?

In alle genoemde beroepsgroepen lopen de ontwikkelingen in de richting van *meer beoordelen* dus parallel. Opvallend is dat overal verhitte discussies plaatsvinden en de invoering veel stof doet opwaaien. Ook bij gewone arbeidsverhoudingen met een baas-medewerkers situatie leidt een beoordelingsgesprek vaak tot spanningen, maar onder topprofessionals is 'de hitte' vele malen groter. Wat maakt beoordelen onder topprofessionals tot zo'n brandbaar thema? Dat is een samenspel van verschillende factoren.

Problematisch is om te beginnen de afwezigheid van hiërarchie. Horizontale verhoudingen zijn kenmerkend voor de werkcontext van professionals en er is vooral bij partnerships geen duidelijk hoger niveau waar de beoordelingsbevoegdheid

De druk om topprofessionals te beoordelen neemt in alle sectoren toe

1. M.J. Wanrooy (2012).

De partner als alleskunner, hoe lang is dat nog houdbaar? *De Accountant*, mei, nr. 3, pp. 16-18.

2. M.J. Wanrooy (2011).

'Conflicten tussen topprofessionals: oorzaken, dynamiek en preventie'. *Holland Management Review*, jrg. 28, nr. 137, mei-juni, pp. 8-15.

3. J. Rosingh, M. van der Kemenade & J. Sijmons (2012). *'Disfunctioneren tijdig aangepakt'*, *Medisch Contact*, 3 februari, pp. 264-266.

4. S. van Brunschot (2012). *'Onder de Loep: de partnerevaluatie'*, *www.advocatie.nl*, 1 aug.

5. H. Koning, (2012). *'IFMS gebaat bij collegiaal leerklimaat'*, *Zorgmarkt*, juni, pp. 24-26.

neergelegd kan worden. Daarnaast bestaat het typerende 'Ik ben eigen baas'-paradigma. Of dat nu wel of niet feitelijk juist is – in maatschappen is

Vaak ligt (vermeend) disfunctioneren ten grondslag aan conflicten

de partner daadwerkelijk

mede-eigenaar – maakt niet uit; de professional voelt zich koning op het eigen vakgebied en de vraag 'Wie ben jij dat je mij beoordeelt' steekt dan al snel de kop op. Hoewel sterk aan erosie onderhevig, speelt ook nog steeds het gevoel mee dat men de meesterproef heeft afgelegd, nu tot de goden op de Olympos behoort en daarmee een onaantastbare status heeft verkregen. Op het eigen professionele domein laat men – al of niet terecht – geen inbreng van anderen toe, en dat geldt niet alleen voor rechters. Door beroepssocialisatie, sterk beïnvloed door een gesloten opleidingscultuur, zoals bij advocaten of medisch specialisten, ontwikkelt zich een vorm van vaktrots die soms tot gecultiveerde eigenwijsheid (en een groot ego) leidt.

Dit samenspel van factoren wordt bovenal gedomineerd door angst. Veel topprofessionals hebben een bovengemiddelde behoefte aan erkenning en zijn gevoelig voor beoordelen. Blijft erkenning uit, dan voelen ze zich kwetsbaar. Dat is ook niet zo verwonderlijk, want toppers hebben wat te verliezen: hoe hoger je bent gekomen, des te dieper kun je vallen. Reputaties kunnen sneuvelen. Maister noemt professionals *insecure overperformers*, die leiden aan het *imposter syndrome*: de angst om op een dag ontmaskerd te worden.⁶ De andere kant van dezelfde medaille is dat topprofessionals zeer kritisch zijn op zichzelf (daarom behalen ze hoge prestaties; ze zijn nooit tevreden met zichzelf) en tegelijkertijd ook uiterst kritisch op anderen. Juist hierdoor is het zo lastig om kritiek te geven en te ontvangen. Topprofessionals ervaren kritiek op hun presteren al snel als persoonlijk en kwetsend, en zij weten of denken – vaak terecht – dat anderen het ook zo zullen ervaren. Deze geprojecteerde onzekerheid verklaart voor een belangrijk deel waarom het voor professionals zo lastig is om elkaar aan te spreken op hun functioneren.

Er spelen overigens nog meer zaken mee: in een maatschap of partnership is men voor de collegiale sfeer, voor de vakuitoefening én voor de inkomsten sterk van elkaar afhankelijk. Men kijkt wel uit om de onderlinge relaties te belasten met risicovolle feedback, men wil geen ruzie, er staat simpelweg teveel op het spel. Daarnaast is er nog een geheel andere reden waarom men elkaar wei-

nig aanspreekt: gecalculeerd eigenbelang – als ik jou met rust laat, laat jij dan mij met rust?

Extra complicerend bij de invoering van beoordelingssystemen is dat de prestatiedruk in alle sectoren toeneemt. De 'rat race' is een passende metafoor. Wetenschappers moeten publiceren ('publish or perish'), advocaten en accountants/fiscalisten moeten klanten scoren en hun targets halen, medici moeten aan productienormen voldoen. Een beoordelingssysteem wordt gezien als een strop om de nek, terwijl het water al tot aan de lippen staat. Daarbij komt dat invoering soms plaatsvindt omdat enkele onderpresteerders moeten worden aangepakt. Of dat juist is of niet, maakt niet uit. Het wordt zo gepercipieerd en daarom is het verzet vaak taai en hevig. Er komt nog meer olie op het vuur wanneer de individuele *beloning* wordt gekoppeld aan de beoordelingsresultaten. De discussie over criteria en meetmethoden wordt dan helemaal op het scherpst van de snede gevoerd. In die situaties wordt men het moeilijk eens over de criteria, iedereen pleit voor maatstaven waarop men zelf goed scoort. Toch moeten de productietijgers aan de ene kant en de sociaal voelende teamleden anderzijds het eens worden met elkaar over: wat verwachten we van elkaar en welke criteria nemen we op in het profiel?

Case 1

In een werkgroep die een voorstel moet voorbereiden voor een nieuw evaluatiesysteem raakt de discussie flink gepolariseerd. Collega A wil het criterium omzet centraal stellen ('We zijn een commercieel bedrijf en daarin draait het nu eenmaal om omzet en winst, daar leven we van'). Collega B wil innovatie en bijdrage aan de organisatie centraal stellen ('We moeten meer samenwerken en op lange termijn hangt ons voortbestaan af van onze innovatiekracht').

Het blijkt dat collega A solistisch werkt, deel uitmaakt van een lucratieve praktijk in een gunstig marktsegment en hoge omzetten scoort. Collega B, daarentegen, haalt zijn targets niet maar speelt wel een belangrijke rol als bestuurder en 'verbinder' in de organisatie. Zij weerspiegelen twee stromingen in het team en het valt niet mee om consensus te bereiken over de uiteindelijke set criteria.

Wat zijn de dilemma's?

Er doet zich een aantal fundamentele dilemma's voor bij het ontwikkelen van een evaluatiesysteem. Het allerbelangrijkste daarvan is de tegenstelling tussen vrijheid en beheersing. Professionele arbeid

6. D.H. Maister (1993). *Managing the professional Service Firm*. New York: The Free Press.

gedijt het beste als er ruimte is voor creativiteit. Het opleggen van een systeem waarin criteria en meetbare indicatoren vóóraf worden bepaald, verdraagt zich slecht met die noodzakelijke vrijheidsgraden. Ook wil je de intrinsieke motivatie van de professional (de 'gouden brandstof' in professionele organisaties) niet smoren door die professionals het gevoel te geven aan verplichte normen en prestatieafspraken te moeten voldoen. Er dreigt een gevaar: als je de gewenste output van een professional vooraf vastlegt, transformeer je eigenlijk de intrinsieke motivatie (ik wil het uit mezelf doen) tot extrinsieke motivatie (ik doe het omdat het moet). Daarmee dood je de passie van de professional, zelfs al zou de uiteindelijke prestatie, bijvoorbeeld drie artikelen per jaar, hetzelfde zijn.

Ook is het de vraag of het plezier van *creatie* of van *ondernemerschap* – wat vaak geldt als hoogste doel – niet wordt geschaad door het stellen van eisen en concrete output-afspraken, zoals aantallen innovaties of aantallen nieuwe klanten. Als je niet oppast wordt het systeem ervaren als afdwingen en klemrijden, in plaats van als een stimulans voor excellente prestaties. De enige radicale oplossing voor dit dilemma is: geen eisen vastleggen. Maar dat is weinig bevredigend. De nadelen zijn simpelweg te groot: vrijblijvendheid, ongelijke prestaties, geen sturing.

De oplossing moet dus gezocht worden in het bouwen van een 'robuust' systeem, dat wil zeggen niet te fijnmazig. Er zijn dan wel richtlijnen en bandbreedtes geformuleerd, maar men moet aan de verleiding weerstand bieden om alle gewenste prestaties precies, objectief en meetbaar te definiëren. Vooral voor professionals die streven naar onderbouwing, objectiviteit en perfectie, is dit een lastige opgave. Door de robuustheid blijft het evaluatiesysteem imperfect en zal er altijd ruimte voor interpretatieverschillen blijven bestaan; daar moet men tegen kunnen. Anders sterft het systeem in schoonheid en verwordt het tot een naargeestig bureaucratisch monster met 'geneuzel' op de vierkante centimeter.

Een tweede dilemma betreft de harde (meetbare) en de zachte (onmeetbare) criteria. Omdat harde criteria zoals omzet, winst en aantallen verrichtingen goed te definiëren en te meten zijn, zijn deze criteria in vrijwel elk systeem dominant. Zachte criteria, zoals collegialiteit, samenwerken, initiatief of stressbestendigheid, zijn echter minstens even belangrijk; veel conflicten over disfunctioneren gaan over deze zachte aspecten. Deze zogeheten 'gedragscompetenties'⁷ zijn echter moeilijker

te meten en daardoor spelen ze – ten onrechte – vaak een ondergeschikte rol. De oplossing voor dit tweede dilemma ligt in dezelfde richting als hiervoor geopperd voor het

eerste: focus niet alleen op datgene wat meetbaar en objectiveerbaar is,

Het tijdperk van 'met rust gelaten worden' is voorbij

maar heb de moed om robuuste, subjectieve parameters op te nemen in het systeem. Daar komt nog een overweging bij: uiteindelijk is het gesprek dat je met elkaar voert, *naar aanleiding van* de criteria, belangrijker dan de exacte meetbaarheid van prestatie of de vraag of iemand op een bepaald criterium een 3, 5 of 4 scoort.

Een derde dilemma is van een andere orde. Hier gaat het niet over de criteria of het systeem als zodanig, maar over de vraag *wie met wie* de gesprekken voert. Er is niet altijd een (duidelijk) hoger hiërarchisch niveau, en dat maakt een vorm van horizontale beoordeling onvermijdelijk. In beide gevallen (beoordeling door een 'baas' of door een 'collega') zal de professional moeite hebben om

Case 2

Bij een voorstel over een nieuw evaluatiesysteem worden in de inleidende notitie de volgende doelen van het systeem beschreven:

1. De ontwikkeling van de partners stimuleren door reflectie en feedback te bieden ten aanzien van hun functioneren.
2. De professionaliteit van de partners verbeteren.
3. De kwaliteit van de dienstverlening en de bedrijfs-economische prestaties verbeteren.
4. Een toetsingskader bieden voor het toekennen van variabele beloning.

beoordeeld te worden, maar de macht van een hoger niveau kan een eventueel verschil van mening gemakkelijker beslechten; dat scheelt. Een ander punt is dat het expertisegebied van een professional vaak erg smal is. 'Wie heeft zowel de kennis als het inzicht om mij *op mijn domein* te beoordelen?', vraagt de topprofessional zich terecht af. Het is onontkoombaar dat *peers* bij de beoordeling worden betrokken, maar dat kan ten koste gaan van de onafhankelijkheid van de beoordeling; wanneer directe collega's een evaluatiegesprek met elkaar voeren, zullen zij geneigd zijn elkaar uit de wind te houden en is de kans op 'politiek spel' groter. Een oplossing voor dit probleem kan zijn om een *combinatie* van gespreks-

7. H.A. Hoekstra & E. van Sluijs (1999). *Management van competenties*. Assen: Van Gorcum.

partners bij de beoordeling te betrekken – bijvoorbeeld een voorzitter of directielid ‘op afstand’ en een collega uit de directe werkomgeving. Ook 360°-feedback kan behulpzaam zijn (op dit laatste aspect kom ik later in dit artikel nog terug).

Doelen

Laten we voorop stellen dat het primaire doel van partnerevaluatie *ontwikkeling, professionalisering* en verbetering van prestaties is. Dit wordt nog wel eens uit het oog verloren. Beoordelingssystemen raken overschaduwd door discussies over de ‘foto’,

Professionals voelen zich vaak koning op hun eigen vakgebied – wie heeft dan ‘het recht’ om hen te beoordelen?

dat wil zeggen de beoordeling zelf. De focus versmalt zich tot de score op de verschillende criteria. De spanning wordt extra groot als de *beloning* afhankelijk is van deze score. Men vergeet eenvoudigweg dat evaluatie

in essentie een hulpmiddel is voor *ontwikkeling* – het gaat om de film, niet om de foto. De andere zaken, zoals het vaststellen van een score of *ranking* in verband met het toekennen van beloning, zijn te beschouwen als nevendoelen.

Welk systeem?

Er zijn vele systemen en procedures mogelijk: licht en zwaar, beperkt of uitgebreid, smalle (alleen harde) criteria of brede (ook zachte) criteria. De variaties zijn eindeloos. We kunnen alle mogelijke systemen indelen met behulp van een hoofddimensie LICHT-ZWAAR en vier aspecten, die afgezet worden op deze hoofddimensie. De vier aspecten zijn:

A. Wie zijn de actoren?

In de lichte variant zijn alleen directe collega’s betrokken. De evaluatie lijkt in dat geval op onderlinge professionalisering door middel van een collegiaal groepsgeprek; de verhoudingen zijn gelijkwaardig. In de zware variant speelt een hoger hiërarchisch niveau de hoofdrol en is er sprake van een top-down-beoordelingsrelatie.

B. Wat zijn de criteria?

In de lichte variant kan worden gekozen voor niet meer dan enkele simpele *output*-parameters, zoals omzet, uren of productie, of deze kunnen worden aangevuld met andere meetbare outputcriteria, zoals aantal gepubliceerde artikelen, aantal uitgebrachte offertes of aantal nieuwe klanten. In beide

gevallen blijft de beoordeling beperkt tot een meting van de *harde* criteria, die het voordeel hebben dat ze goed meetbaar zijn. In de zware variant worden ook zachte criteria betrokken, zoals *reputatie* of *bekendheid in netwerken*, maar ook typische *gedragscriteria* als collegialiteit, initiatief, begeleiding van junioren of stressbestendigheid.

Bij de zware variant zijn dus twee sporen te onderscheiden, namelijk resultaten en competenties:

- De *resultaten* zijn onder te verdelen in een aantal resultaatgebieden, zoals omzet, winstgevendheid, externe profilering en kwaliteit. Elk gebied kan worden voorzien van indicatoren die gemeten worden. Omzet en winstgevendheid worden bijvoorbeeld gemeten via de indicator euro’s en percentages per jaar. Externe profilering kan worden gemeten – ‘benaderd’ is misschien beter uitgedrukt, want exact meten kan niet – via de indicatoren aantallen publicaties, lezingen of klantbijeenkomsten die men heeft gerealiseerd. De parameter *kwaliteit* kan in beeld worden gebracht via de indicatoren afnemersfeedback, reviews, aantallen klachten of verloren klanten.
- *Competenties* zijn onder te verdelen in specifieke gedragscompetenties als ondernemerschap, leidinggeven of plannen & organiseren. Deze gedragscompetenties zijn veel moeilijker objectief te meten, maar met behulp van *gedragsindicatoren* (bijvoorbeeld: ‘voert voortgangsprekken’ of ‘stelt takenlijst op’) kunnen de competenties worden geconcretiseerd.

Zowel resultaten als competenties kunnen op het niveau van de indicatoren gescoord worden aan de hand van een schaalverdeling, bijvoorbeeld een cijfer 1 t/m 5, of in woorden: onvoldoende, matig, voldoende, goed, zeer goed. Het totaal-

Case 3

Ten behoeve van de partnerevaluatie bij een advocatenkantoor wordt een formulier ingevuld. Voor het resultaatdomein *Business Development* worden de partners beoordeeld op de *output*-indicatoren:

- Aantal en aard publicaties en docentschappen
 - Aantal en aard relevante netwerkactiviteiten
 - Aantal nieuw ontwikkelde producten
 - Aantal (en aard) nieuwe cliënten
 - Aantal (en aard) vertrokken cliënten
 - Aanwezigheid en actieve betrokkenheid bij klant-events en acquisitieactiviteiten.
-

beeld van de scores levert een eindscore op. In het geval van variabele beloning kan deze eindscore worden gebruikt als basis voor de toekenning daarvan.

C. Foto of film?

In een eenvoudige opzet (lichte variant) wordt de foto vooral gemaakt op basis van de terugblik: hoe wordt je functioneren beoordeeld in het licht van je prestaties van het afgelopen jaar? In een uitgebreide opzet (zware variant) wordt niet alleen gekeken naar de status quo, maar wil men echt *ontwikkeling* bevorderen. De nadruk ligt dan op het benoemen van ontwikkelpunten en het maken van toetsbare afspraken voor de komende periode. Om deze ontwikkeling het nodige *gewicht* te geven, is het aan te raden om daadwerkelijk scores toe te kennen aan de mate waarin men progressie heeft laten zien ten opzichte van een jaar eerder. In hoeverre heeft men de verbeterafspraken gerealiseerd? Daarbij past het uitgangspunt dat *iedereen* – dus ook diegenen die een jaar eerder al ‘goed’ presteerden – zich kan verbeteren en zich dus ontwikkeldoelen behoort te stellen. De score die wordt toegekend aan getoonde progressie telt mee voor de eindscore. Op deze wijze wordt *ontwikkeling* daadwerkelijk onderdeel van de beoordeling (en beloning). In een uitgebalanceerd systeem komen dus *drie* domeinen voor waarop men wordt beoordeeld: resultaten, gedrag en ontwikkeling.

D. Welke bronnen worden gebruikt?

Bij de keuze van de criteria wordt ook bepaald welke bronnen men benut.

Bij productienormen kunnen de interne managementinformatiesystemen worden geraadpleegd. Bij het beoordelen

Heb de moed om robuuste, subjectieve parameters op te nemen in het systeem

van gedrag ligt het voor de hand dat feedback van collega's en eventueel klanten wordt betrokken. De wijze waarop dat gebeurt ligt gevoelig. In de praktijk wordt hier veel mee geëxperimenteerd:

- Laten we de beoordeelde zélf zijn feedbackgevers kiezen? Doet de beoordelaar dat? Zo ja, hoe (at random of gericht)?
- Maken we de feedbackgevers bekend of blijven ze anoniem?
- Verzamelen we mondelinge informatie van collega's of organiseren we schriftelijke feedback (bijvoorbeeld via een 360°-feedbackmethode)?
- Betrekken we ook de secretaresses erbij, of alleen directe collega's?
- En niet onbelangrijk: nemen we *klantfeedback* mee in de beoordeling?

Wat de beste keuze is hangt af van de cultuur, de gewoonten, de breedte en diepte die men met de beoordeling beoogt en vooral: de mate waarin men 'toe' is aan concrete, directe vormen van feedback op gedragsaspecten.

In figuur 1 vormen de bovengenoemde hoofdimensie LICHT – ZWAAR en de vier genoemde

Figuur 1. Overzicht beoordelingssystemen

	Licht: beperkt, open en vrijblijvend	← — — — — — →	Zwaar: uitgebreid, dicht en stringent
A. Actoren	Collega's onderling	Commissie van partners	Bestuur/Voorzitter/ Managing Partner
B. Criteria	Productie (uren, verrichtingen, omzet)	Productie én andere output	Productie én andere output én gedrag (Resultaten en Competenties)
C. Van foto naar film	Terugblikken	Terugblikken en huidige status quo vastleggen	Ook: vooruitblikken, afspraken maken en monitoren
D. Bronnen	Productiecijfers	Productiecijfers en andere meetbare resultaten (lezingen, nieuwe klanten, etc.)	Ook: feedback op gedrag (van collega's, klanten)

aspecten A t/m D samen een matrix, die een overzicht geeft van mogelijke varianten van beoordelingssystemen.

Hoe beoordeling invoeren?

Beoordelen leidt – vooral als er beloningsconsequenties aan verbonden zijn – tot pittige discussies, niet alleen omdat sommige collega's zich bedreigd voelen, of omdat niet alles objectief

De kwaliteit van de evaluatie staat of valt met het gesprek zelf

meetbaar te maken is, maar ook omdat het onderwerp aanleiding geeft tot een debat over ideologie en waarden. Aan de ene kant hoor ik in de praktijk geluiden als: 'Zo gaan we als collega's toch niet met elkaar om!', 'We gaan er toch vanuit dat iedereen zijn stinkende best doet', of 'Als we elkaar zó de maat gaan nemen, dan hoeft het van mij niet meer'. Men ervaart een beoordelingssysteem als aantasting van fundamentele waarden als collegialiteit, solidariteit en eigen verantwoordelijkheid. Aan de andere kant zijn er de pragmatici en degenen die bovengemiddeld presteren, die stellen: 'De bijdrage van ieder mag best zichtbaar worden gemaakt' of 'Overpresteren mag beloond worden' en sommigen zullen denken (maar niet hardop zeggen) 'Het wordt tijd dat collega X wordt aangepakt, want die presteert al jaren onder de maat'.

Het invoeren, of uitbreiden van een beoordelingssysteem betekent dat men behendig moet laveren tussen de klippen van voor- en tegenstanders. Dit is een echt hoofdpijndossier voor vele bestuurders, vooral omdat zij vaak maar een beperkt hiërarchische positie hebben en zich er voortdurend van moeten vergewissen dat ze op voldoende steun van de (dominante groep) van

Case 4

In een evaluatiesysteem bij een ingenieursbureau is het onderdeel *Leidinggeven* als volgt geconcretiseerd in gedragsindicatoren:

- Geeft duidelijk aan wat hij/zij van teamleden verwacht
- Biedt ruimte voor ieders inbreng
- Bevordert samenwerking tussen medewerkers en afdelingen
- Draagt organisatiewaarden uit
- Zorgt voor coaching en begeleiding van medewerkers.

collega's kunnen rekenen. Wat zijn de tips om draagvlak te creëren voor zo'n gevoelig onderwerp?

Zoals bij iedere door te voeren verandering is het neerzetten van de *urgentie* een belangrijke pijler voor commitment. Vervolgens is het zaak om zorgvuldig een aantal fasen in het proces te doorlopen: informeren, inspireren, overtuigen, consulteren, onderhandelen en uiteindelijk – wanneer nog maar een kleine minderheid tegen is – afdwingen. Dit kan door een officieel besluit met meerderheid te laten nemen, of door de tegenstrevers voor een *fait accompli* te stellen: mee doen of als dissident aan de kant blijven staan.

Wat helpt is om een externe gezaghebbende autoriteit (deskundige of *respected peer*) erbij te betrekken en hem of haar over de ervaringen elders te laten vertellen. Ook kan het helpen om een commissie van betrokkenen uit eigen kring samen te stellen; zo kan het bestuur voorkomen dat het iets oplegt dat niet gedragen wordt en op klassieke wijze als *not invented here* wordt afgekeurd. Tijdens het ontwikkelproces van het nieuwe systeem is het ook nuttig om consultatiesessies met de betrokkenen te houden, zodat inspraak gewaarborgd is. Verder is het altijd aan te bevelen om eerst in de vorm van een *pilot* te starten. Dan kan het systeem na een jaar worden geëvalueerd en bijgesteld. Tot slot is in de praktijk gebleken dat het verstandig is om eerst het evaluatiesysteem op orde te brengen, en daar enkele jaren ervaring mee op te doen, en er pas daarna eventueel variabele beloning aan te koppelen (zo men dat al wil).

Valkuilen

Tot besluit volgt hier een aantal valkuilen die ik in de praktijk vaak tegenkom:

- Elk systeem is maatwerk. U begaat een vergissing als u een methode van elders (of uit het boekje) regelrecht kopieert en geen rekening houdt met de bestaande cultuur, gewoonten en historie. Ook het ontwikkelen van een evaluatiesysteem is te beschouwen als een film; een statische oplossing is er niet. Of een bepaald ontwerp voldoet, kun je alleen bepalen door het te relateren aan waar je met elkaar vandaan komt en waar je naartoe gaat. De beste keuze is dus afhankelijk van context en ontwikkelingsfase.
- Er dreigt vrijwel altijd het gevaar van overdetailering. Door het streven naar objectiviteit en meetbaarheid heeft elk systeem de neiging uit te groeien tot een heel 'boekwerk', en iedere poging om een procedure te verfijnen kan leiden

tot een bureaucratisch geheel. Er ontstaan dan te veel (sub)criteria en het kost te veel tijd om de formulieren in te vullen; het systeem wordt een illusie van degelijkheid en blijkt uiteindelijk onwerkbaar te zijn. Het is de kunst om het systeem 'robuust' te houden en er eerder iets vanaf te halen dan aan toe te voegen. Een stuk of zes resultaatgebieden en zes tot acht competenties volstaan om de essentiële profieleisen in te vervatten en daarmee blijft het profiel ook nog goed hanteerbaar. Een *vijfpuntsschaal* is prima, niet te grof en niet te fijn. Formuleer niet meer dan twee ontwikkelafspraken; dat is genoeg om serieus de aandacht op richten. Maak de beloningsschalen niet te complex. De moraal is: *keep it simple*.

- Subjectiviteit valt niet uit te sluiten. Het is een misvatting dat subjectiviteit en (dus) discussies volledig uitgebannen kunnen worden. Je moet niet alleen rationeel, mathematisch en met meetbare grootheden willen werken, want als je niet oppast wordt er dan 'hard' gerekend met zachte feiten. Je moet altijd met gezond verstand én met gevoel blijven beoordelen. De *human touch* blijft belangrijk. Redelijkheid en billijkheid moeten de ultieme keuzen bepalen – en niet de optelsom van scores; die is alleen maar richtinggevend.
- Wees voorzichtig met het genereren van over-

dadige managementinformatie om het systeem te 'vullen'. Bijvoorbeeld: in een centrale registratie bijhouden wie hoeveel artikelen heeft geschreven, of wie precies hoeveel nieuwe klanten heeft binnengehaald. Laat het aan de beoordeelde over om zélf zijn portfolio ('bewijsstukken') op orde te brengen; dat ontslaat het bestuur of de organisatie van de verplichting om veel detailinformatie te verzamelen, wat mogelijk complex en in elk geval tijdrovend kan zijn.

- Voer niet een compleet beoordelingssysteem in wanneer het er alleen maar om gaat om één of twee onderpresteerders 'aan te pakken'. Het is *altijd* lastig om de confrontatie met onderpresteerders aan te gaan, met of zonder systeem. In de praktijk zie ik af en toe dat het bestuur 'vlucht' in het systeem. Zeker, een systeem voorziet in een ondergrond; verwacht er echter niet alle heil van, want een confronterend (of conflictueus) gesprek wordt het meestal toch. Het gaat er om, *oprecht* en *moedig* te zijn, met of zonder systeem.
- De kwaliteit van de evaluatie staat of valt met het gesprek zelf. Draait men om de hete brij heen? Durft of kan men geen kritische feedback geven? Heeft de beoordelaar voldoende gespreksvaardigheden in huis? Heeft men het lef om confrontaties aan te gaan? Men kan nog zo'n mooi scoresysteem hebben, als de gespreksdeelnemers elkaar sparen, is het effect van de evaluatie gering. Het systeem is uiteindelijk slechts een hulpmiddel; het gaat er om, *hoe* het gesprek wordt gevoerd. Nogmaals: wees *oprecht* en *moedig*. Dat bepaalt uiteindelijk het succes van beoordelen onder topprofessionals. ■

Er vloeit nog meer olie op het vuur wanneer de individuele beloning wordt gekoppeld aan de beoordelingsresultaten

Case 5

Bij de vernieuwing van een evaluatiesysteem in een groot accountantskantoor wordt besloten om de evaluatieprocedure uit te breiden met feedback van collega's. Maar hoe kan dit het beste worden gedaan? 'Koude' feedback verzamelen via een vragenlijst lijkt te afstandelijk, maar zomaar bij collega's langsgaan om even te vragen hoe zij vinden dat collega X functioneert wordt te vrijblijvend geacht.

Er wordt gekozen voor een tussenvorm: degenen die feedback moeten geven, krijgen een A4'tje met daarop de belangrijkste criteria uit het evaluatiesysteem en mogen enkele dagen nadenken over deze criteria in relatie tot de persoon over wie zij feedback moeten geven. Na die bedenktijd wordt hun gevraagd om mondeling een beoordeling over collega X te geven, aan de hand van de criteria op het A4'tje waarover ze hebben kunnen nadenken. De respondenten hoeven het A4'tje dus niet in te vullen, en evenmin wordt het formulier ingenomen, omdat het dan een te formeel karakter zou krijgen; men hecht vooral aan de 'warme', mondelinge informatie.