



Dermatologisch getest!

Zin en onzin van wetenschappelijke taal in onderzoek

Adviesbureaus bieden hun onderzoeksrapporten via eigen kanalen en via de media aan. Veel van die rapporten zijn daarom terug te vinden in de Online Bibliotheek van www.managementenconsulting.nl. Maar wat staat er eigenlijk in? En bovenal: waar baseren de auteurs hun uitspraken op? Zes ervaren onderzoekers (Floor Basten, Jaap Germans, Steven ten Have, Lieke Hoogerwerf, Dick Ruimschotel en Coert Visser) buigen zich over vier rapporten om de wetenschappelijke pretenties op hun merites te beoordelen. Niet één doorstaat de test.

Adviesbureaus versturen regelmatig persberichten waarin zij nieuw onderzoek aankondigen. Veel van deze onderzoeksrapporten plaatst *Management en Consulting* met toestemming van de onderzoekers in de Online Bibliotheek. Het is een omvangrijke verzameling geworden die een kijkje in de keuken van bureaus geeft. Nieuwsgierig geworden naar de kwaliteit van deze onderzoeken heeft de redactie zes ervaren onderzoekers gevraagd de deugdelijkheid ervan te verkennen. Het team is dusdanig samengesteld, dat een brede variatie in methodologische kennis vertegenwoordigd is.

De rapporten zijn vrij willekeurig uit het aanbod op de website geselecteerd. De enige voorwaarde was de afwezigheid van een externe opdrachtgever. De vier rapporten dus gaan over onderzoeken die door de adviesbureaus zelf zijn geïnitieerd.

Gezien de pretenties die de rapporten hebben, willen we het wetenschappelijk perspectief als richtlijn nemen. We hebben vooraf alle rapporten geanalyseerd op:

1. de relatie tussen vorm en inhoud;
2. relevantie van het onderwerp;
3. betrouwbaarheid en validiteit;
4. het verband tussen data en conclusies, en
5. de afbakening en definities.

Wetenschap is een gesprek over waarheid. Een belangrijk kenmerk van wetenschappelijke publicaties is daarom dat deze transparant zijn. Zonder transparantie kan het gesprek niet goed gevoerd worden, want de mogelijkheid om de bevindingen kritisch op hun waarde (waarheidsclaim) te beoordelen ontbreekt. Bij transparantie hoort dat de onderzoeksopzet besproken en verantwoord wordt, want daarmee laten de onderzoekers zien hoe ze de data verwerkt hebben tot een antwoord op hun onderzoeksvraag. Collega's moeten in staat zijn op grond van diezelfde data tot hetzelfde antwoord te komen. Dit principe, herhaalbaarheid, heeft als consequentie dat alle waarnemingen en de manier waarop die gedaan zijn beschreven zijn en dat definities nauwkeurig beschreven en afgebakend zijn. Alleen dan kunnen collega's beoordelen of voldaan is aan het validiteitscriterium: de vraag of gemeten is wat de onderzoekers wilden meten. Met betrouwbaarheid wordt bedoeld dat wanneer het onderzoek herhaald wordt, er min of meer dezelfde gegevens uitkomen (de mate van toeval). In publicaties zien we dit terug in een beschrijving van bronnen (literatuur, respondenten). Popularisering van wetenschappelijk onderzoek komt in veel tijdschriften voor. Vaak wordt een ingewikkeld wetenschappelijk verhaal uitgelegd in alledaagse taal. In die gevallen wordt verwezen naar de bron, de plek waar het ingewikkelde verhaal te vinden is. Daar waar zonder bronvermelding gemeld wordt dat 'onderzoek uitwijst dat...', wordt het verdacht. Even suspect zijn tendentieuze en suggestieve uitspraken die de betekenis van de resultaten pogen te veranderen. In principe moet elke waarheidsclaim weerlegd kunnen worden. Het is immers een claim en geen geloof. De discussiesectie is de plek om de onderzoeksresultaten in dat

licht te bespreken. Publicaties die een mogelijke weerlegging ontkennen en zonder enig voorbehoud resultaten weergeven, zijn in de regel niet wetenschappelijk.

(Dat wetenschap een gesprek is, is een premisse die aan deze exercitie ten grondslag ligt.)

Niet op afstuderen

Het onderzoek van Booz Allen Hamilton 'Profiles in Organizational DNA. Research and Remedies' is een beschrijving van de resultaten van online onderzoek. Deelnemers kunnen het DNA van hun organisatie bepalen aan de hand van een self-assessment. Er zijn volgens dit rapport analoog aan DNA vier bouwstenen die het functioneren van organisaties bepalen: informatie, besluitvorming, motivatie en structuur. Op basis hiervan zijn zeven profielen mogelijk, waarvan er drie verwijzen naar gezond en vier naar ongezond functioneren. Op basis van 4007 verzamelde assessments concluderen de auteurs dat het merendeel van alle organisaties ongezond is.

Waar stoelt deze conclusie op? Het valt ons op dat de cijfers in de grafieken niet kloppen en het is onduidelijk welke betekenis we moeten hechten aan verschillen tussen percentages. Dit kan door afrondingsverschillen komen maar dan is het onduidelijk welke betekenis kleine verschillen tussen percentages hebben. Het onderzoek is echter niet herhaalbaar, want de software waarmee de gegevens geanalyseerd worden is niet beschikbaar voor herhalingsonderzoek. Ook de verwijzing naar de ervaring van de auteurs maakt herhaling onmogelijk. Er is geen discussie van de resultaten, geen enkel voorbehoud en nul relativering.

De vraag wie om welke redenen de vragenlijst invult (dus de vraag naar mogelijke selectiviteit), wordt niet gesteld, maar de conclusies worden desondanks voor alle organisaties getrokken. Mensen die op de website terechtkomen, worden met de vraag of ze willen weten waarom organisaties slecht functioneren uitgenodigd hun eigen organisatie te diagnosticeren. Daar zit al een behoorlijke selectie van respondenten in. We vragen ons daarom af hoe betrouwbaar de conclusies zijn.

Verder zijn 'bouwstenen' contra-intuïtief en niet gebaseerd op literatuur, bijvoorbeeld over contingentie (gedrag variërend per situatie, afdeling of tijdstip). De profielen zijn classificaties zonder spreiding. We vragen ons af of glijdende schalen of dimensies niet beter de gezondheid een organisatie weergeven. Maar we vragen ons wel meer af: wat is de relatie tussen de bouwstenen en de profielen en waar zijn de aanbevelingen op gebaseerd?

De bevindingen in dit rapport zijn desondanks niet weerlegbaar, omdat er een cirkelredenering in het onderzoek zit. De auteurs definiëren de bouwstenen en de profielen, classificeren vervolgens de antwoorden op de self-assessments in die profielen en bepalen wat gezond is en wat niet. Ze bevestigen als het ware hun eigen model zonder andere gegevens of verklaringen toe te laten: bronnen buiten Booz Allen Hamilton schitteren in afwezigheid. Deze zelfbevestiging

zien we terug in dramatisch taalgebruik als 'Anyone who has worked in an organization of any size...' Het is ondoenlijk om aan de ongezonde predeterminatie te ontkomen, want het DNA van organisaties bestemt ons inherent voor tot verkeerd gedrag. De metafoer van het organisationeel DNA blijkt niet houdbaar, want dankzij biotechnologie kunnen organisaties wél ontsnappen aan fatalistische horizonten. En kijk nou: Booz Allen Hamilton heeft de kwalificaties in huis om aan DNA te sleutelen en organisaties te genezen! Onze conclusie over dit onderzoek: hier kun je niet mee afstuderen.

Vlees noch vis

Dan naar het onderzoeksrapport van Twynstra Gudde, 'Nationaal Projectmanagement Onderzoek ® ... voor bedrijven'. Dit rapport geeft de resultaten van een enquête over projectmanagement weer. We kunnen echter niet ontdekken wat er precies onderzocht is. Het lijkt te gaan over de effectiviteit van projectmanagement, maar het onderzoek gaat over de meningen van mensen die met projectmanagement te maken hebben. Daarbij zijn de definities die van belang zijn voor de resultaten vaag en slordig: elke verandering wordt gezien als een uiting van projectmanagement, terwijl veel veranderingen ook totstandkomen via daartoe meer geëigende instrumenten.

Het onderzoek beperkt zich tot de steekproef, die verder geen relatie lijkt te hebben met de populatie waaruit hij afkomstig is. We moeten zelf de vragenlijst reconstrueren uit de weergave van resultaten. Er worden geen significante resultaten gegeven, slechts interpretaties in termen van 'substantieel', 'aanzienlijk', 'redelijk' en 'in beperkte mate', of in termen van 'meer dan'. Hierdoor is het lastig de resultaten te beoordelen op hun waarde: we weten immers niet wat 'substantieel' is en hoe belangrijk 'meer dan' is. De respondenten is gevraagd een top 3 van hefbomen te kiezen uit een lijst; onduidelijk is of de respondenten deze hefbomen tegenkomen in hun praktijk of juist niet. Het rapport sluit af met een aantal aanbevelingen.

We vinden het rapport helder, verzorgd en elegant, hoewel de auteurs voor het overzicht beter tabellen hadden kunnen gebruiken. Dat had hen waarschijnlijk ook geholpen de cijfers voor de variabelen consistent weer te geven. Waar komen die variabelen trouwens vandaan? We vinden geen enkele bronvermelding of andere verwijzing naar een theoretische inbedding. Had er überhaupt iets anders dan dit uit het onderzoek kunnen komen? Een vergelijkend onderzoek, bijvoorbeeld tussen projectmanagement en andere methodieken, was spannender en wellicht meer ont-hullend geweest. Over de wetenschappelijke ambitie kunnen we het eens zijn: die is laag. Over de vraag of dat prettig is, verschillen we van mening. Want als de ambitie zo laag is, waarom dan de pretentie wetenschappelijk te zijn? Waarom niet gewoon een Handboek Soldaat maken, in plaats van de aanbevelingen cijfermatig te willen optuigen?

De vraag is: is dit rapport – gezien de methodologische makke – goede PR voor Twynstra Gudde(alle aanbevelingen

kunnen zonder dit bureau worden doorgevoerd en dat is op zich positief)? Met een meer kritische studie zou het bureau meer bereiken. We weten immers niets van de respondenten en we kunnen dus ook niet beoordelen of dit een evaluatieve studie onder de eigen klanten is. Aan de andere kant kan een kritische studie juist de eigen kracht laten zien en de positie versterken. Het bevreedt ons ook dat deze experts opeens aan de leek gaan vragen hoe ze hun projecten managen, zonder hiermee verder iets te doen. Kortom: dit rapport laat ons gepuzzeld achter. Ons oordeel: vlees noch vis.

Reclamefolder

Het KPMG-rapport 'Strategic Management of Intellectual Property: Enhancing Organisational Success by Unlocking the Values of Ideas' stemt ons niet vrolijker. Hier wordt het belang van het managen van intellectueel kapitaal onderstreept door zes mensen die telefonisch zijn geïnterviewd. Geen van hen spreekt het belang tegen en we kunnen stellen dat we hier letterlijk met een selectief gezelschap te maken hebben. Dat tast de betrouwbaarheid van de resultaten aan. Verder wordt een tabel gegeven met daarin drie ranglijsten van patenthoudende bedrijven, maar de toegevoegde waarde van deze tabel ontgaat ons. De opmerking in de lopende tekst dat sommige bedrijven wel duizend patenten per jaar aanvragen lijkt ons, gezien de kosten die daarmee gemoeid zijn en het aantal gehonoreerde aanvragen (216 voor de koploper) enigszins overdreven. Bovendien is de bronvermelding naar KPMG-workpapers vaag.

Net als het rapport van Booz Allen Hamilton wordt er een sense of urgency gecreëerd (gemiddeld staat op elke pagina een must, have to, need to, should) die moet leiden tot de aanschaf van een product. We hebben het dus over behoeftecreatie: het systeem is de oplossing, het niet hebben van dit systeem is het probleem. Er wordt ingespeeld op angst door een probleemscheppend verhaal. Maar ofschoon het rapport hamert op shareholder value, laat het nergens zien dat portfolio management van intellectueel eigendom door bijvoorbeeld winst uit patenten daadwerkelijk harde euro's oplevert. De validiteit is dus ver te zoeken. Ons oordeel: dit is een reclamefolder, vermomd als onderzoeksrapport. KPMG bevestigt dat het niet een wetenschappelijk onderzoek is maar een white paper.

Waarom?

Ten slotte het NOVIM-rapport 'Nationaal Onderzoek Verandermanagement (NOVIM)' over verandermanagement van Intermaat, een onderdeel van Boer & Croon. De opmerking 'puur feitelijk op basis van kundig onderzoek' in de inleiding is een handschoen die wij graag oppakken. Gaat het om feiten? Is het onderzoek kundig? Na de inleidende start wordt het rapport opgesplitst in twee delen die niets met de geschetste maatschappelijke context te maken hebben: het eerste gaat over interventievoorkuren en het tweede over stimulerende en blokkerende factoren bij verandering.



De onderzoekers

Floor Basten, oprichter van onderzoeksbureau Orleón (www.orleon.nl), is onderzoeker met een sociaal wetenschappelijke en taalkundige achtergrond. Haar specialiteit is interpretatief onderzoek. Voorbeelden van haar werk zijn metaforen binnen een academische opleiding bedrijfswetenschappen (proefschrift), actief burgerschap, ruimtelijke beleving en de taal van sociaal rendement.

Lieke Hoogerwerf is mede-initiatiefneemster van Netwerkimpuls (www.netwerkimpuls.nl). Als adviseur, onderzoeker en docent ondersteunt zij mensen en organisaties bij het ondernemen en organiseren vanuit netwerkprincipes. Voorheen was zij werkzaam als sociotechnisch adviseur, universitair docent bedrijfskunde en leermanager verandermanagement.

Coert Visser werkte bij PricewaterhouseCoopers en is sinds een jaar of vijf zelfstandig adviseur. Tot zijn belangrijkste werkzaamheden behoren: individuele coaching, assessment, begeleiding van (management) teams, begeleiden van werkconferenties, geven van trainingen en het adviseren in grote adviesprojecten. Hij is ook lid van de hoofdredactie van www.managementsite.net.



Dit artikel plaatst de redactie voor een dilemma. Wij willen het debat over de kwaliteit van onderzoeksrapporten aangaan. Niet alleen in deze aflevering maar ook in toekomstige artikelen. De rapporten hier besproken, zijn ons opgevallen vanwege de publiciteit die ze genereerden of vanwege een actuele in de belangstelling staande onderzoeksvraag. De bureaus zijn niet blij met het artikel en hadden graag gezien dat we het hadden ingetrokken. Het debat doen we daarmee geen goed. Met publicatie lopen we het risico dat we mensen en bureaus tegen ons in het harnas jagen met alle mogelijke consequenties. Zoals de opbouw van de Online Bibliotheek dat inmiddels een heel fraai onderzoeksarchief biedt. We willen laten zien wat bureaus doen; dat is een taak van het vakblad. Dit artikel is een resultaat van een discussie onder panelleden. Onderzoekers die hun werk serieus nemen in een 'ons kent ons' wereldje. Management en Consulting leent zich niet voor kinnesinne onder professionals of bureauïjd noch gaat het ons om uitsluitend een smakelijk verhaal te schrijven. Wel lenen wij ons voor het debat en de redactie nodigt u graag voor deelname uit. Stuur uw reactie voor 1 september naar redactie@managementenconsulting.nl.

De onderzoekers

Dick Ruimschotel is oprichter en directeur van Compliance Methodology Consultants (www.cmc@T11.net). Het bureau werkt op het grensvlak van compliance, integriteit en rechtshandhaving. Ruimschotel is jurist en sociaal wetenschapper (methodoloog) met interesse in klassieke onderwerpen als causaliteit en modellen.

Steven ten Have is organisatieadviseur en mede-eigenaar van advieskantoor Ten Have & Company (www.tenhavecompany.com). Het bureau richt zich op vraagstukken op het gebied van strategie, verandering en besturing. Verder is Ten Have hoogleraar aan de VU Amsterdam, leerstoel Strategie en Verandering.

Jaap Germans (niet op de foto) werkte 25 jaar bij Philips als onder meer kwantitatief onderzoeker bij Natlab, adviseur en hoofd van een intern adviesbureau. Vervolgens was hij ruim tien jaar zelfstandig organisatieadviseur en tevens 'schoolhoofd' bij de TIAS voor de Mastersopleiding Management van Leren en Ontwikkeling. Germans studeerde wiskunde en is bij een sociaal psycholoog gepromoveerd.

Waar baseren de auteurs hun uitspraken op?

Het deel van de interventievoorkuren is eigenlijk geen onderzoeksrapport, want er wordt een vernieuwing van een bestaand instrument gepresenteerd. Bij die presentatie worden resultaten vermeld van een online onderzoek naar interventievoorkuren, terwijl de instructies bij de vragenlijst in de bijlage gaan over feitelijk gedrag. Het is dus onduidelijk wat gemeten wordt.

De 'kleurentheorie' (volgens ons is dit helemaal geen theorie maar een classificatievoorstel met vage verwijzing naar arbitraire kleuren) wordt nergens onderbouwd. Is het model uitputtend? Valide? Betrouwbaar? Zijn de categorieën consistent? We tasten volledig in het duister, maar de eigenschappen van de onderzoekers doen er ook niet zo toe. Hun kundigheid wel. De vragenlijst telt 17 stellingen waar telkens 8 punten over verdeeld kunnen worden. Per categorie (5 in totaal) kan dus gemiddeld 27 punten worden gescoord. Blijkbaar is dit gemiddelde plus of min 2 'normaal', eindigt de categorie 'laag' bij 24 en start de categorie 'hoog' bij 30.

Een nogal gevoelig instrument, maar slordig toegepast: niet alle stellingen zijn opgeteld $17 \times 8 = 136$. De willekeur in kleuren maakt niet uit. Waar het om gaat, is dat de categorieën een overzicht geven van de geschiedenis van het adviesvak tot een bepaald tijdstip en dat deze categorieën de nationale standaard lijken te worden voor het heden en daarna.

Het tweede deel van het NOVVM-rapport lijkt geen verband te hebben met het eerste. Hier wordt wel een onderscheid gemaakt tussen 'nu' en 'wenselijk'. Een weergave in tabellen in plaats van een herhaling van de variabelen zou fijn zijn geweest. Wat ons opvalt, is dat er schijnbaar zeer simpele oplossingen zijn voor vastzittende situaties. De auteur vraagt zich bijna wanhopig af waarom men maar niet in staat is tot de meest basale ingrepen. Misschien legt de diagnose niet de 'echte' blokkerende en stimulerende factoren bloot? De vraag is – en dat geldt voor de kleurentest net zo – wie deze vragenlijsten invult en waarom; is hier sprake van een ontladende catharsis of gaat het om geïdentificeerde groepen? Is het verstandig om de schijn te wekken dat het onderzoek onder academische vlag is uitgevoerd omdat beide hoogleraren vanuit de universiteit opereren, maar tegelijkertijd hun werkerterrein voor het grootste deel in de adviespraktijk hebben? Dit soort achtergrondinformatie zou niet relevant moeten zijn, maar sommigen van ons vinden het buitengewoon vervelend dat deze kwaliteit zich vertoont onder academische vlag. Ons gehele oordeel over dit onderzoek: waarom?

Dermatologisch getest

In de reclamewereld is het voor sommige producten gebruikelijk om met een zweem van wetenschappelijkheid de geloofwaardigheid van het product te onderstrepen. Een kreet als 'dermatologisch getest' moet de kwaliteit garanderen en de kritische consument over de streep trekken. Het ziet er naar uit dat deze marketingtruc de managementwereld binnensluipt. Natuurlijk, adviesbureaus willen hun tijd en producten verkopen, en daar is helemaal niets mis mee. Het

wordt wél een beetje kwalijk als het verkoopargument gestoeld is op wetenschappelijk onderzoek dat niet deugt. Waarom eigenlijk?

Is er sprake van devaluatie van onderzoek wanneer adviesbureaus rapporten publiceren met wetenschappelijke claims die ze niet waar kunnen maken? Zijn die rapporten niet veelal een gespreksonderwerp bij de koffiemachine of op de borrel, waarna weer wordt overgegaan tot de orde van de dag? Misschien maken ze iets wakker in mensen dat toch al aandacht wilde, zijn ze niet meer dan een ontstekingsmechanisme. In elk geval, zo menen we, zijn het pogingen om gesprekken te bepalen, gesprekken waar soms managers maar al te graag aan meedoen. Anders wordt het, wanneer hoogleraren er hun signatuur aan verlenen en de rapporten een navenante status geven. In dit kader is het zinvol om een drietal gesprekken te onderscheiden, elk met een eigen thema. Het wetenschappelijke gesprek gaat over waarheidsclaims en wordt dus gestructureerd door de in dit artikel genoemde criteria. Het adviesgesprek gaat over veranderingen en verloopt langs de lijnen van diagnose en interventie. Het managementgesprek gaat over de praktijk en thematiseert hoe de veranderingen gewerkt hebben. Deze gesprekken kunnen in één en hetzelfde rapport best samenkomen, maar dan is het onze aanbeveling om ze helder af te bakenen. Dus nee, liever geen wetenschappelijke termen in een adviesgesprek alleen om te camoufleren dat het om een adviesgesprek gaat, net zo min als er adviestermen in een wetenschapsgesprek mogen voorkomen zonder de onderbouwde herkomst daarvan.

Tips

Enkele tips voor adviesbureaus die voornemens zijn rapporten op basis van eigen onderzoek te publiceren. Wie meent een wetenschappelijk verantwoord rapport te kunnen schrijven en tevens vermoedt dit zo te kunnen dat managers het een spannend verhaal vinden om te lezen, moet dit vooral doen. Wie bevroedt dat de lezer afhaakt in een al te taaie brei van wetenschappelijke paraferalia, of wie verwacht dat het lezerspubliek niet gecharmeerd is van enig voorbehoud van bureauezijde, kan twee rapporten produceren. In de marketingversie staat dan een samenvatting van de opzet en de resultaten én een verwijzing naar het tweede rapport. Dit tweede rapport is dan het volgens de wetenschappelijke criteria opgezette onderzoeksrapport, waarin de methode inzichtelijk is en de resultaten besproken worden in een discussiesectie. Wie twijfelt aan de eigen wetenschappelijke kwaliteiten hoeft zich nergens voor te schamen – onderzoek is namelijk ook een vak – en kan zich beperken tot één document, dat dan productbrochure heet en bij voorkeur door iemand anders is geschreven.

Is deze selectie van rapporten representatief voor alle rapporten die door bureaus worden opgesteld? Voldoet dit artikel aan de criteria die erin besproken zijn? We hopen u voldoende richtlijnen geboden te hebben om dat zelf te beoordelen. ■