

KPI's en de valkuilen van ons denken

PSYCHOLOGIE

De mens is niet zo rationeel als we graag zouden geloven. Studies hebben aangetoond dat we vaak op systematische en op voorspelbare wijze denkfouten maken. Onze redeneringen steunen niet zozeer op de regels van de logica, als wel op eenvoudige denkstrategieën die leiden tot vooroordelen en illusies.

Door Maarten van Eck

Door de geschiedenis heen hebben tal van filosofen en logici zich over de geïdealiseerde wereld van de logica gebogen. Zij stelden zich de vraag hoe de mens 'zou moeten denken', en probeerden het correcte denken en redeneren in regels uit te drukken. Pas sinds de laatste twintig jaar getroosten wetenschappers zich de moeite met experimenten na te gaan hoe het in werkelijkheid met ons redeneren is gesteld. Toonaangevend voor dit empirisch onderzoek is het werk van de Amerikaanse psychologen Amos Tversky en Daniel Kahneman.

Tversky en Kahneman onderzochten hoe mensen in situaties van onzekerheid oordelen en beslissen; de situaties dus waarmee we in het dagelijkse leven geconfronteerd worden. Of we nu het terugverdieneffect van een nieuw software product inschatten, de kansen van een nieuwe propositie, het inschatten van sales targets; meestal beschikken we over onvolgende informatie en kunnen we aan gebeurtenissen slechts een bepaalde graad van waarschijnlijkheid toekennen.

Tversky en Kahneman constateerden dat mensen bij dit soort problemen niet de regels van de logica en de waarschijnlijkheidsleer toepassen, maar wel vuistregels of strategieën die het hen mogelijk maken op een eenvoudige manier een oplossing te vinden. Deze vuistregels – heuristieken – zijn soms efficiënt, maar kunnen ons vaak ook op een dwaalspoor zetten, doordat we ze niet op de juiste manier of niet bij het juiste soort probleem toepassen.

Het is daarom van belang wanneer een organisatie KPI's wil inzetten om te sturen, dat men bij het bepalen van deze Kritische Performance Indicatoren beseft hoe psychologische processen het gedrag dat een KPI moet bewerkstelligen kan breken of maken.

Absoluut versus relatief

Hoeveel procent van alle McDonald's vestigingen bevindt zich in de Verenigde staten? U weet dat daar McDonald's is begonnen, de Amerikanen graag fast-food eten en dat het land groot is. Deze informatie tezamen gebruikt u om een uiteindelijk antwoord te formuleren. De meeste antwoorden liggen tussen de 20 en 40 procent, maar het juiste antwoord is 40 procent. En wat is vervolgens het aantal vestigingen van McDonald's in de Verenigde staten? Dan lopen de antwoorden behoorlijk uiteen: van 100 tot 20.000. Wanneer we om een relatief gegeven vragen dan zitten we met ons antwoord er vaak niet ver naast maar met een absoluut getal gaan we de mist in; het juiste antwoord is 12.804 vestigingen. Het horen van een absoluut getal roept sterkere emoties op dan een relatief getal. "Zoveel, dat had ik niet verwacht!" Een absoluut antwoord is belangrijk om mensen echt iets bewust te maken. Dat tachtig procent van de klanten geen klachten heeft doet ons betrekkelijk weinig. Maar wanneer we weten dat het aantal klachten de honderdduizend overstijgt dan hebben we als snel een beeld van een ontevreden menigte en weten we dat we hier actie op moeten ondernemen.

TIP. Wanneer KPI's worden gebruikt om prestaties te meten, moet zoveel mogelijk worden ingespeeld op de gevoelsmatige impact. Op een absoluut getal reageren we sterker dan op een relatief getal. Wanneer je gedrag wil kunnen beïnvloeden moet je werken met KPI's die mensen raken.

Framing

Een ander voorbeeld is dat de manier waarop we een probleem oplossen vaak afhankelijk is van de wijze waarop het probleem wordt geformuleerd of omkaderd. Psychologen spreken hierbij van 'framing' van keuzes. Dat zelfs hooggevoelden zich door deze heuristiek laten beetnemen mocht blijken toen Amerikaanse artsen werd gevraagd of ze in een specifiek geval al dan niet een chirurgische ingreep zouden aanbevelen. Als hen werd verteld dat de ingreep tot een sterftcijfer van 7 procent (binnen de vijf jaar na de operatie) zou leiden, dan bevalen ze die ingreep slechts met enige aarzeling aan. Wanneer hen echter werd gezegd dat de overlevingskans 93 procent bedroeg, dan waren ze veel meer geneigd de ingreep warm aan te bevelen. Of het glas als halfvol of als halffleg wordt beschreven, maakt voor onze hersenen dus wel degelijk een verschil uit en kan dus het gedrag van een KPI ten positieve of negatieve beïnvloeden.

TIP. Bepaal de omkadering van een KPI door na te denken over de psychologische impact.

Aversie tegen verlies

Kahneman toonde aan dat mensen verliezen veel zwaarder laten wegen dan dat ze eenzelfde waarde stijging waarderen. Heel basaal gezegd is de emotionele impact van het verliezen van 50 euro emotioneler dan het winnen van 50 euro.

TIP. Van dit inzicht kunnen we gebruik maken door KPI's zo te bepalen dat we bijvoorbeeld niet winst meten maar het gederfde verlies. Hierdoor zal een medewerker eerder geneigd zijn om tot actie over te gaan om van het ongemakkelijke gevoel af te komen. Al is het natuurlijk ook wel van belang om een balans aan te brengen: dus negatieve en positieve KPI's om te voorkomen dat een medewerker gedemotiveerd wordt.

De conjunction fallacy

De conjunction fallacy is een valkuil waarbij we aannemen dat specifieke condities waarschijnlijker zijn dan een algemene conditie. Een voorbeeld zal dit beter kunnen verduidelijken.

"Linda is 31 jaar oud, single, slim, en heeft een uitgesproken mening. Ze is afgestudeerd in pedagogiek. Tijdens haar studententijd maakte ze zich druk om discriminatie en sociale onrechtvaardigheid, en heeft ze meegelopen in demonstraties tegen het gebruik van kernenergie."

Wat is nu waarschijnlijker:

A. Linda is werkzaam bij een gemeente;

B. Linda is werkzaam bij een gemeente en actief in een feministische beweging?

Wat bleek nu; 85 procent van de personen koos voor optie B. En hoewel dit intuïtief inderdaad de goede keus lijkt, is het rekenkundig een verkeerd antwoord. Optie A is namelijk het meest waarschijnlijk.

Dit komt omdat de kans op twee samenvallende gebeurtenissen (de 'conjunction') altijd kleiner is dan de kans op één van die gebeurtenissen. Stel dat je naar aanleiding van de beschrijving schat dat er een kans van 10 procent is dat Linda bij de gemeente werkt, en dat je de kans dat ze in een feministische beweging actief is op 80 procent schat. Het lijkt dan wel alsof mogelijkheid B het meest waarschijnlijk is, maar deze kans is maar 8 procent. Immers; $0,10 \times 0,80 = 0,08 = 8$ procent.

Dit lijkt gevoelsmatig niet te kloppen, maar als je erover redeneert blijkt het wel te kloppen. Stel dat we 100 mensen hebben die aan de beschrijving van Linda voldoen. Van deze 100 mensen werken er tien bij de gemeente. Doordat we ook iemand moeten hebben die in een feministische beweging actief is, hebben we maar een groep van tien mensen waaruit we kunnen kiezen. Omdat een klein percentage waarschijnlijk politiek actief is, hebben we een of twee mensen uit de gehele groep van 100 mensen die zowel bij de gemeente werken als actief zijn in een politieke beweging.

Mogelijkheid B doet echter waarschijnlijker aan omdat deze representatiever is voor Linda zoals ze uit de beschrijving naar voren komt. Nu is dit nog een onschuldig voorbeeld, maar uit andere onderzoeken is ook gebleken dat politieke adviseurs kansen op ingrijpende gebeurtenissen verkeerd schatten. Ondanks dat een gedetailleerd en specifiek scenario erg waarschijnlijk lijkt, wordt de daadwerkelijke kans erop steeds kleiner door het toevoegen van meer details. Dit effect treedt ook op in KPI-modellen. Ondanks dat bepaalde voorspellingen erg waarschijnlijk lijken, zijn ze eigenlijk door de hoeveelheid aan details juist onwaarschijnlijk. Met andere woorden: een analyse waarin veel factoren voorkomen heeft ook een grotere kans om juist niet uit te komen.

Zo kan conjunction fallacy ook optreden bij het ontwikkelen van een KPI scorecard. Ondanks dat het misschien lijkt dat een scorecard met vijf of meer indicatoren een beter systeem is dan een simpele versie met twee indicatoren, hoeft dat niet zo te zijn. Sterker nog, door het toevoegen van extra indicatoren vergroot je alleen de kans dat de toekomst een volledig andere uitkomst geeft.

TIP. Hoe kan een medewerker zich nu wapenen tegen deze valkuil? Kennis van deze valkuil zal zeker helpen. Ook als een medewerker zich realiseert dat samenvallende gebeurtenissen altijd een kleinere kans hebben kan men hiermee rekening houden. Mocht een gedetailleerd scenario erg waarschijnlijk overkomen, bereken dan eens de kans dat het scenario zo uitkomt. Grote kans dat u het scenario een stuk minder waarschijnlijk vindt!

Law of small numbers

De law of small numbers (of overhaaste generalisatie) houdt in dat we conclusies trekken op basis van te weinig gegevens. De foutieve aanname die we hierbij maken is dat een kleine steekproef uit een grote populatie een nauwkeurige weergave is van diezelfde populatie. Echter, het is zo dat des te kleiner je steekproef is, des te groter het gevaar is dat je steekproef behoorlijk afwijkt van de grotere populatie. Bij een kleine steekproef is er meer spreiding om het gemiddelde (grotere standaardafwijkingen) dan optreedt bij een grotere steekproef. Het probleem is echter dat we, ondanks dat we dit waarschijnlijk wel weten, met beslissingen hier geen rekening mee houden. Een IT-operations afdeling lijkt misschien een uitstekende beheerpartij omdat het zes van de afgelopen tien maanden zijn SLA's heeft gehaald, maar de steekproef van tien maanden is hiervoor veel te klein om deze uitspraak te doen. Ook bij een patroon van technische analyse dat de afgelopen zes keer erg goed uitpakte, valt er geen conclusie te trekken of dit patroon daadwerkelijk een goed signaal geeft.

TIP. Een manier om met deze valkuil om te gaan is om bij een KPI te zetten of de data erachter representatief zijn, zodat we weten dat de getrokken conclusie statistisch onderbouwd is.

Recency effect

De valkuil van de law of small numbers wordt nog eens versterkt door het recency effect. Bij deze valkuil vinden we recente data belangrijker dan gegevens die al 'oud' zijn. Het is daarom ook niet verwonderlijk dat we de laatste zes keer dat een bepaald patroon voorkwam dusdanig belangrijk vinden, dat we geneigd zijn om te concluderen dat het patroon werkt. Een gevolg van deze valkuil is dat we, naast dat we een op het oog goede strategie kunnen uitvoeren die eigenlijk helemaal niet werkt, (te) snel vertrouwen winnen maar ook (te) snel verliezen. Zo kunnen we al vertrouwen krijgen in een strategie als die de afgelopen zes keer goede signalen gaf. Aan de andere kant kunnen we net zo gemakkelijk dit vertrouwen weer verliezen als de afgelopen zes signalen van de strategie verliesgevend waren. Grote kans dat je in het

laatste geval zult concluderen dat je strategie niet meer werkt. Bij het recency effect kennen we teveel gewicht toe aan de meest recente data, zonder veel van de oudere data mee te nemen in de overweging. Dit speelt vooral bij complexe beslissingen waarbij men veel variabelen moet overwegen, zoals bij beleggen. Het effect is eenvoudig te verklaren doordat we de meest recente gebeurtenissen nu eenmaal beter onthouden dan zaken die langer geleden zijn gebeurd.

TIP. Een oplossing om deze valkuil te vermijden is om bij een KPI te zetten over welke tijdsperiode deze is berekend en of in het verleden grote afwijkingen zijn geweest ten aanzien van de waarde van de huidige KPI.

Conclusie

Ondanks dat de psychologische effecten hierboven als losstaand worden beschreven, is de scheiding in de praktijk een stuk minder duidelijk. Zo zorgt het recency effect er bijvoorbeeld voor dat we als ankerwaarde vooral een meest recente waarde kiezen. Voeg daarbij dat we door de law of small numbers te snel conclusies trekken op basis van te weinig data en we vinden twee recente ankerwaarden al snel 'erg belangrijk'. Dit maakt het niet gemakkelijker om deze psychologische effecten te doorzien. Maar om goed doordachte beslissingen te nemen op basis van wat een KPI Scorecard of Performance Dashboard toont, begint het in eerste instantie met kennis over deze psychologische effecten. Want wanneer men weet dat een dergelijk effect bestaat kan men met het beoordelen van resultaten hier rekening mee houden. In tweede instantie kunt u met de praktische tips in dit artikel uw voordeel doen en een grotere kans creëren dat KPI's goed worden geïnterpreteerd en bij de gebruiker het gewenste gedrag opleveren.

Bronnen

Kahneman D., J. Knetsch & R. Thaler (1986), Fairness and the assumptions of economics, *Journal of Business* 59, S285-S300.
Kahneman D., J. Knetsch & R. Thaler (1990), Experimental tests of the endowment effect and the Coase theorem, *Journal of Political Economy* 98, 1325-1348.
Kahneman D. & D. Lovallo (1993), Timid choices and bold forecasts: A cognitive perspective on risk taking, *Management Science* 39, 17-31.
Kahneman D. & A. Tversky (1972), Subjective probability: A judgment of representativeness, *Cognitive Psychology* 3, 430-454.
Kahneman, D. & A. Tversky (1973), On the psychology of prediction, *Psychological Review* 80, 237-251.
Kahneman, D. & A. Tversky (1979), Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263.
Skepsis.

Maarten van Eck is Senior BI consultant bij Capgemini.