Statistisch redeneren kunnen we op allerlei manieren onderzoeken: wiskundig, didactisch, historisch, filosofisch, psychologisch, enzovoort. In dit artikel maakt **Arthur Bakker** een uitstapje naar de cognitieve psychologie.

Statistisch redeneren

Uitstapje naar de cognitieve psychologie

Statistiek is een glibberig terrein waar toeval en regelmaat verstoppertje met elkaar spelen. Iedereen kent wel statistische blunders van journalisten, wetenschappers, leerlingen, en misschien wel van zichzelf. In de cognitieve psychologie is onderzocht waardoor we statistische fouten maken en of het mogelijk is om statistisch redeneren door onderwijs te verbeteren. De antwoorden hierop laten in de geschiedenis van de cognitieve psychologie een pendelbeweging zien tussen optimisme en pessimisme. Ik bespreek enkele psychologische experimenten en inzichten, waar niet alleen leraren, maar ook slimme verkopers en beleidsmakers van kunnen profiteren.

Voor 1970

De meest optimistische kijk op het statistisch redeneren lijken Piaget en Inhelder (1951) te hebben. Zij ontdekten dat veel kinderen van zeven tot tien jaar al begrip van kans en de wet van de grote getallen hebben, zoals het onderstaande fragment illustreert. Het ging over een wijzer die je kon ronddraaien op een cirkelschijf en die dan stopte op een van de acht kleuren (Nisbett *et al.* 1983, p. 343).

Onderzoeker:	*Als ik tien of twintig keer draai, kan het dan zijn dat de wijzer op een bepaalde kleur niet één keer stopt?*
Kind van 7:	*Ja, dat kan. Het zou vaker zijn bij tien keer dan bij twintig.*
Onderzoeker:	*Zal de wijzer op alle kleuren komen?*
Kind van 10:	*Het hangt ervan af hoe vaak je draait.*
O:	*Waarom?*
K:	*Omdat als we vaker draaien, het meer kansen heeft om overal te komen.*

Op hun elfde jaar hebben veel kinderen volgens Piaget en Inhelder al gevoel voor niet-uniforme verdelingen zoals de 'normale' verdeling (bij het bord van Galton). Kinderen geven blijk van meer inzicht in het kansbegrip als het fysieke mechanisme dat het toeval genereert duidelijk herkenbaar is. In dit het geval zijn dat de wijzer op de cir-

Hoeveel olifanten?

kelschijf en het bord van Galton. Kennelijk helpt het om leerlingen te vragen naar het mechanisme dat de kansverdeling bepaalt.
Ook andere onderzoekers, zoals Peterson en Beach (1967), kwamen tot de conclusie dat mensen die intuïtief statistisch redeneren het lang niet slecht doen.

Tversky en Kahneman: 1970-1980

Dit rooskleurige beeld kwam onder druk te staan door het onderzoek van Amos Tversky en Daniel Kahneman. Hun experimenten lieten zien dat zelfs statistisch geschoolde studenten enorme inschattingsfouten maken. Tversky en Kahneman formuleerden enkele niet-wiskundige principes die bepalen hoe wij redeneren: het principe van *toegankelijkheid*, dat van *verankering* en dat van *representativiteit*. Deze principes zorgen ervoor dat wij dagelijkse onzekerheden kunnen behappen, maar ze leiden er ook toe dat wij in sommige situaties wiskundig gezien de mist ingaan.

Een opmerking vooraf. In veel experimenten wordt gevraagd om een schatting te geven. Is een schatting van een aantal of een percentage een statistische redenering? Ja, want een goede schatting is (meestal impliciet) gebaseerd op kennis van dataverwerking, het kansbegrip en statistische principes. Ter illustratie: hoe zou u het aantal olifanten op de foto schatten? Wat voor statistische kennis gebruikt u daarbij?

Principe van toegankelijkheid

Vraag rokers te schatten hoeveel procent van de bevolking rookt en ze zullen een veel hoger percentage schatten dan niet-rokers. Dit komt doordat ze meestal ook meer rokers kennen, waardoor hun beeld niet objectief is. Tversky en Kahneman (1974) noemen dit fenomeen het principe van toegankelijkheid. Mensen redeneren op grond van de informatie waartoe ze toegang hebben. In dit geval zijn dat de rokers die ze kennen.
De onderzoekers illustreren het verschijnsel ook met het volgende experiment. Ze vroegen proefpersonen het volgende:
Wat is de verhouding tussen het aantal Engelse woorden die met een k *beginnen en het aantal Engelse woorden die een* k *als derde letter hebben? Idem dito voor de letter* r.

De proefpersonen dachten dat er meer woorden met een *k* beginnen dan er woorden zijn met een *k* als derde letter. In feite zijn er drie keer zoveel Engelse woorden met een *k* als derde letter als woorden met een k als beginletter. Ook in het geval van de letter *r* en andere letters overschatten mensen het aantal woorden die met een bepaalde letter beginnen. Dit verklaren Tversky en Kahneman met het principe van toegankelijkheid. De mate waarin mensen toegang hebben tot bepaalde kennis in hun geheugen, is bepalend voor de schatting die zij maken. In dit experiment betekent het dat proefpersonen makkelijk woorden vinden met een bepaalde beginletter, omdat zij gewend zijn op de eerste letter te zoeken, zoals bij het gebruik van woordenboeken.

Hetzelfde verschijnsel vonden Tversky en Kahneman (1973) bij het volgende experiment. Er waren twee lijsten waarop precies evenveel mannen als vrouwen stonden, maar de ene lijst bevatte relatief bekendere vrouwen, de andere lijst relatief bekendere mannen. Proefpersonen kregen de lijsten te horen. Naderhand werd ze gevraagd of er meer mannen of meer vrouwen op de lijst stonden. Bij beide lijsten vergisten de proefpersonen zich. Ze dachten dat de lijst met relatief bekendere mannen ook meer mannen bevatte en de lijst met relatief bekendere vrouwen meer vrouwen. Dit verklaarden de onderzoekers als volgt. Beroemde namen zijn makkelijker te onthouden. Als de proefpersonen in hun geheugen graven, treffen ze daar meer bekende dan onbekende namen aan.

Over het algemeen houden mensen geen rekening met dit principe van toegankelijkheid. Dit kan, zoals bovenstaande experimenten laten zien, tot verkeerde conclusies leiden. Een slimme advocaat houdt rekening met het principe van toegankelijkheid en roept bij een pleidooi nog even de relevante informatie in het geheugen van de luisteraars op. Als die informatie vers in het geheugen zit, is de kans groot dat de rechter of jury die in de beslissing laat meewegen.

Principe van verankering

Mensen blijken erg makkelijk gemanipuleerd te kunnen worden. De onderzoekers vroegen proefpersonen om het percentage Afrikaanse landen in de NAVO te schatten. Voordat de proefpersonen mochten antwoorden, werd er een rad gedraaid met daarop de getallen 0 tot en met 100. Ze moesten dan eerst zeggen of het gevraagde percentage boven of onder dit getal lag. Daarna mochten ze hun schatting geven. Het bleek dat de mensen die eerst het getal 10 zagen als mediaan de schatting 25% gaven, en dat de mensen die eerst het getal 65 op het rad zagen als mediaan tot de schatting 45% kwamen, een beduidend hoger percentage. De onderzoekers hebben dit verschijnsel *het principe van aanpassing of verankering* genoemd. Mensen passen zich aan de eerste waarde aan, zelfs als volkomen duidelijk is dat het eerstgenoemde getal door zuiver toeval totstandkomt. Dit getal vormt een soort anker voor de vervolgredenering. Er wordt hun letterlijk een rad voor ogen gedraaid.

Een vergelijkbaar resultaat vonden Tversky en Kahneman toen ze middelbare scholieren binnen vijf seconden 8! lieten schatten.
De ene groep kreeg $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ te zien en de andere groep $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$. De eerste groep had als mediaan 2.250, de tweede groep 512 (en de

rekenmachine 40.320). Vermoedelijk rekenen leerlingen eerst een paar producten uit en extrapoleren dan. Het eerste product vormt een soort anker voor de rest van de schatting. Dit verklaart dat de groep die met $1 \times 2 \times 3$ begon tot een veel lagere schatting kwam dan de groep die met 8×7 begon. (In de paragraaf over de schatting van het aantal olifanten kom ik hierop terug.)

Een slimme verkoper maakt van dit soort inzichten gebruik. Een postorderbedrijf in de Verenigde Staten verkocht eerst één soort broodmachine voor $279. Later bood het ook een iets luxere machine aan voor $429. Uiteraard werd de duurdere weinig verkocht, maar de verkoop van de goedkope versie verdubbelde in korte tijd. De dure versie fungeerde waarschijnlijk als een soort anker waaraan de goedkope gerelateerd werd. Daardoor leek de goedkope versie in vergelijking tot de dure een uitstekende koop.

Principe van representativiteit

Een meer omvattend principe is dat van *representativiteit*. Het makkelijkst is dit principe te illustreren aan de hand van een voorbeeld uit de kansrekening.

Vraag proefpersonen bij muntwerpen naar de kans op de volgende series kop of munt:

KMKMMK, KKKMMM, KKKKKK.

Ze zullen waarschijnlijk de kans op de laatste serie het kleinst schatten en de kans op de eerste het grootst. Dit komt omdat de eerste representatief is voor een random serie en de andere twee niet. De serie KKKMMM is alleen representatief in het feit dat er evenveel kop als munt in voorkomt. Omdat de eerste serie het meest representatief is, acht men de kans daarop het grootst.
Hoewel dit principe van representativiteit in het dagelijks leven soms nuttig is, leidt het volgens Tversky en Kahneman tot veel soorten fouten, waarbij mensen met name geen rekening houden met de wet van de grote getallen, noch met regressie naar het gemiddelde of de basisverhoudingen binnen de populatie. Hieronder volgen daarvan enkele voorbeelden.

De wet van de grote getallen
Een fout die veel mensen begaan, is dat ze geen rekening houden met de steekproefgrootte. Ze zondigen in het bijzonder tegen *de wet van de grote getallen*. De zwakke wet van de grote getallen zegt dat de kans dat je een bepaald percentage afwijkt van de verwachting kleiner wordt als de steekproefgrootte toeneemt. Dit betekent bijvoorbeeld dat de kans dat het percentage 'kop' in een experiment van muntwerpen meer dan 10% afwijkt van de 50%, kleiner wordt als je vaker werpt.

Aan studenten werd gevraagd in welk ziekenhuis vaker meer dan 60% jongens per dag worden geboren: in een klein ziekenhuis waar gemiddeld 15 baby's per dag geboren worden, of in een groot ziekenhuis waar gemiddeld 45 baby's per dag geboren worden. Het resultaat was:

In het grote ziekenhuis	volgens 21 studenten
In het kleine ziekenhuis	volgens 21 studenten
Maakt niet uit	volgens 53 studenten

Vermoedelijk zien mensen zo'n groep baby's, ongeacht de grootte van het ziekenhuis, als een representatief deel van de populatie waarin het percentage jongens ongeveer 50% is. Dit verklaart waarom 53 van de 95 studenten 'maakt niet uit' kiezen in plaats van 'in het kleine ziekenhuis'. Ze houden, met andere woorden, geen rekening met de (zwakke) wet van de grote getallen.
Gokkers die een serie rood op de roulette gezien hebben, kiezen de volgende keer vaak voor zwart. Als er zwart bij zit, is de serie namelijk meer representatief voor een random-serie. Ze denken dat kans een zelfcorrigerend principe is, maar ze vergeten dat kans geen geheugen heeft.

Regressie naar het gemiddelde
Een andere fout vonden de onderzoekers onder andere bij trainers van piloten. Als piloten in opleiding na een uitstekende landing een compliment kregen, landden ze de volgende keer slechter, en als ze na een slechte manoeuvre een uitbrander kregen, presteerden ze de volgende keer beter. Voor de trainers was dit aanleiding om te geloven dat je beter kon straffen dan belonen. Ze zagen vermoedelijk iedere prestatie als representatief voor de capaciteiten en de inzet van hun piloten. Een statisticus zou eerder denken aan regressie naar het gemiddelde. De kans op een uitzonderlijk goede prestatie is heel klein, dus de volgende keer zal de manoeuvre wel minder goed zijn. Omgekeerd, als iemand een keer slecht gepresteerd heeft, is de kans groot dat de volgende prestatie beter is, zelfs als hij of zij niet eens extra zijn best doet.

Bent u wel eens teleurgesteld als u de tweede keer eet in een restaurant waar u de eerste keer een voortreffelijke maaltijd kreeg? Statistisch gezien is dit heel waarschijnlijk. Het is een voorbeeld van regressie naar het gemiddelde. Beginnersgeluk is ook zo'n fenomeen dat hierop berust. Het is heel waarschijnlijk dat iemand die een nieuw kansspel de eerste keer wint, daarna een paar keer niet wint.

Basisverhoudingen binnen de populatie
Over het algemeen houden mensen geen rekening met de grootte van groepen binnen een populatie. Als die groepen ongeveer even groot zijn, zoals mannen en vrouwen, dan is er niets aan de hand, maar als ze dat niet zijn, dan willen mensen dat nog wel eens over het hoofd zien.

Tversky en Kahneman (1973) vroegen twee groepen proefpersonen de kans te schatten dat een beschreven persoon een ingenieur of een advocaat was. De beschrij-

ving kwam uit een selectie van 100 beschrijvingen van ingenieurs en advocaten. Bij de ene groep proefpersonen (de I-groep) werd gezegd dat er 70 ingenieurs en 30 advocaten waren en bij de andere groep (de A-groep) dat er 30 ingenieurs en 70 advocaten waren. Eén voorbeeld van zo'n beschrijving voor de A-groep:
Jack is 45 jaar. Hij is getrouwd en heeft vier kinderen. Hij is conservatief, voorzichtig en ambitieus. Hij heeft weinig interesse voor politiek en besteedt zijn vrije tijd aan timmeren, zeilen en wiskundige puzzels. De kans dat Jack een van de 30 ingenieurs is, is ....%. (Bij de I-groep stond er *70 ingenieurs.*)

Vreemd genoeg schatte zowel de I- als de A-groep de kans dat Jack een ingenieur was ongeveer 90%, vermoedelijk vanwege de wiskundige puzzels. Toch stond er bij de I-groep dat 70 van de 100 mensen ingenieur was en bij de A-groep 30 van de 100. Kennelijk houden mensen geen rekening met de basisverhoudingen binnen de populatie waar het om gaat.
Dit vermoeden werd bevestigd toen de onderzoekers beschrijvingen gaven die geen relevante informatie bevatten. Beide groepen kwamen dan met de schatting van 50%, terwijl een statisticus voor de I-groep 70% zou verwachten en voor de A-groep 30%.
Dit experiment doet overigens enigszins gekunsteld aan, omdat een nietsvermoedende proefpersoon verwacht dat hij of zij de informatie uit de beschrijving moet halen.

Deze drie principes, van toegankelijkheid, verankering en representativiteit, hebben Tversky en Kahneman in de jaren tachtig aangevuld met andere. Hiervan behandel ik nog het *framing effect.*

Framing effect

Dit effect duidt op de grote invloed die de vraagstelling heeft op de beslissing. Shafir vroeg proefpersonen om jury te spelen in een echtscheidingszaak (Shafir, 1993; Shafir, Simonson & Tversky, 1993). Het kind van een echtpaar in scheiding moest naar een van de ouders. De ene groep proefpersonen moest beslissen aan wie het kind werd toegewezen en de andere groep wie het werd ontzegd. De uitkomst was:

	toegewezen	ontzegd door
Ouder A:		
modaal inkomen		
redelijke gezondheid		
normale werktijden		
redelijke band met kind		
normaal sociaal leven	36%	45%
Ouder B:		
bovenmodaal inkomen		
zeer goede band met kind		
actief sociaal leven		
reist veel voor werk		
goede gezondheid	64%	55%

Wie het kind krijgt, blijkt dus af te hangen van de vraagstelling. Als de jury moet toewijzen, krijgt ouder B het kind; als de jury moet ontzeggen, dan krijgt ouder A het kind.
Shafir verklaart dit fenomeen als volgt. Mensen nemen dat besluit dat ze het makkelijkst kunnen verantwoorden, tegenover zichzelf of tegenover anderen. In het geval van ouder B zijn er meer duidelijke redenen om het kind toe te wijzen (goede band) maar ook meer redenen om het te ontzeggen (vaak weg). Mensen die moeten toewijzen kiezen dus grotendeels voor 'kind naar ouder B' en mensen die moeten ontzeggen kiezen voor 'kind naar ouder A'.

Dit betekent dus dat wie een voorstel door een vergadering wil krijgen, dit framing effect zorgvuldig moet hanteren! De formulering van het probleem kan ervoor zorgen dat het eerder wordt geaccepteerd.
Als er duidelijke argumenten voor en tegen zijn, dan moet je vragen: wie stemt voor? Als het voorstel een beetje kleurloos is met weinig duidelijke voor- en nadelen, vraag dan: wie stemt tegen?
In het verlengde hiervan: denk niet dat het altijd beter is om meer argumenten te geven, want één dubieus argument kan tegen u werken. Een slecht argument creëert namelijk een slechter frame waardoor het geheel van argumenten een slechtere indruk maakt. Ook vanuit het principe van representativiteit kan dit fenomeen begrepen worden. Een argument is representatief voor degene die de argumenten geeft of voor het voorstel dat ingediend wordt. Een slecht argument plaatst het hele voorstel of de persoon die het indient in een negatief frame.

Na 1980

In het voorgaande lijkt het alsof mensen zeer slecht statistisch redeneren en makkelijk te manipuleren zijn. Het feit dat de proefpersonen over het algemeen statistisch geschoolde studenten waren, maakt het beeld er niet vrolijker op. Eén oorzaak is dat Tversky en Kahneman zich geconcentreerd hebben op de vraag: wanneer en waarom gaat het mis?
Na 1980 kwam er veel kritiek op hun onderzoek. Het was te veel op fouten gericht, er kleefden methodologische bezwaren aan en de principes bleken niet goed in staat om te voorspellen wanneer fouten zouden optreden.
In sommige gevallen zijn de experimenten naar mijn smaak gekunsteld en niet erg dagelijks. Denk aan het experiment met de ingenieurs en advocaten. Een nietsvermoedende proefpersoon denkt waarschijnlijk dat hij of zij de informatie uit de beschrijving moet halen en vergeet de scheve verhouding binnen de populatie. Een betere verklaring voor het feit dat mensen 50% kiezen is vermoedelijk dat ze een ander idee van het begrip 'kans' hebben. Voor veel mensen betekent 50%: je weet het

niet, het kan allebei. Zelfs als de kans op een ziekte 1 op 1000 is, dan nog denken sommige mensen dat ze 50% kans hebben om die ziekte krijgen.

Uitkomstbenadering

Een dergelijke foutieve gedachtegang noemt Konold (1989, p. 63) een voorbeeld van de uitkomstbenadering. Volgens Konold menen mensen namelijk dat kansen als doel hebben om een succesvolle voorspelling te geven in een individueel geval. Een voorbeeld vormt de weersvoorspelling van de volgende dag. Als er in de krant staat 'kans op regen 70%', dan gaan mensen ervan uit dat het morgen waarschijnlijk regent. Als het dan niet regent, heeft het KNMI zich volgens velen vergist.
Een voorbeeldje van een student die de uitkomstbenadering hanteert:

Onderzoeker: *Wat zegt het getal, in dit geval 70%, je?*
Student: *... Ik denk dat er een redelijke kans op regen is.*
O: *Stel dat het de volgende dag niet regent, wat kunnen we dan concluderen?*
S: *Misschien vergisten ze zich. Of de regen ging een andere kant op door de wind.*
O: *Als ze nu tien dagen 70% voorspellen, en 3 op de 10 keer is er geen regen, wat vind je dan van de voorspelling?*
S: *Nou, ik denk dat ze het beter hadden kunnen doen. Maar ik denk dat ze hun best gedaan hebben.*

Waarschijnlijk denkt de student dat de voorspelling drie keer niet uitgekomen is. Kennelijk betekent 70% voor deze student niet 'in 7 van de 10 gevallen'. Overigens kan ook de krant verweten worden dat een dergelijk gebruik van kanspercentages misleidend is. Mensen zijn niet geïnteresseerd in de vraag hoe vaak het KNMI gelijk heeft, maar of het morgen regent of niet. Wat die 70% precies betekent wordt ook op de homepage van het KNMI (www.knmi.nl) niet duidelijk.

Wanneer redeneren mensen correct?

Als tegenbeweging op het foutgerichte onderzoek uit de jaren zeventig houden psychologen zich sinds de jaren tachtig meer bezig met de vraag wanneer mensen eerder geneigd zijn om (correct) statistisch te redeneren en of statistische scholing het dagelijks redeneren kan stimuleren en verbeteren. Nisbett, Krantz, Jepson en Kunda (1983) en Fong, Krantz en Nisbett (1986) geven zeer bemoedigende voorbeelden van statistisch redeneren in het dagelijks leven die wiskundig correct zijn. Bovendien laten hun experimenten zien dat statistische training ook het statistisch oplossen van dagelijkse problemen bevordert en verbetert.

Als een leerling een 9,5 haalt terwijl iedereen minstens een 8 heeft, zal geen leerling raar opkijken. Als alle anderen een onvoldoende hebben wordt de 9,5 als een uitzondering gezien. Leerlingen houden dus wel degelijk rekening met de onderliggende verdeling. Evenzo vindt men het niet raar als een oma zegt: ik heb drie kleinkinderen en het zijn allemaal jongens. Mensen vinden het wel vreemd als ze negen kleinkinderen heeft die allemaal jongens zijn. Men weet dat de kans daarop erg klein is.
Uit een klein onderzoekje dat ik zelf in een brugklas heb gedaan, blijkt dat brugklassers wel degelijk gevoel voor steekproeven hebben. Een steekproef van vijf mensen vinden ze wel erg klein als je wilt uitvinden hoe lang Nederlanders gemiddeld per dag televisiekijken. Ze vragen liever honderd mensen.

Uit deze voorbeelden blijkt dat er dagelijkse situaties zijn waarin mensen intuïtief correct redeneren. Ervaring met het onderwerp helpt: mensen met meer ervaring kennen de bijbehorende verdelingen (vaak onbewust) beter dan onervaren mensen. Dit betekent dat volwassenen volgens Nisbett redelijk goed omgaan met statistiek als het gaat over dagelijkse dingen zoals het weer, sport, toetsen en ongelukken.
Men redeneert ook beter als de mogelijke uitkomsten en de kansfactoren duidelijk zijn. In veel sociale contexten is dat niet het geval; die nodigen niet uit tot statistisch redeneren, want er zijn te veel variabelen en onzekerheden. Nisbett en de zijnen hebben verder aangetoond dat mensen niet zomaar generaliseren, maar wel degelijk rekening houden met wat ze generaliseren. In sommige gevallen generaliseren mensen terecht van één voorbeeld, in andere pas na vele voorbeelden. Als het gaat om eigenschappen van elementen zoals de kleur of geleidbaarheid van fosfor, dan generaliseren mensen terecht na één voorbeeld. Als het gaat om karaktereigenschappen van groepen mensen zijn proefpersonen, terecht, veel voorzichtiger. Overigens generaliseren mensen wel sneller bij groepen die ze niet goed kennen.

Wat betekent dit alles voor het onderwijs?

Veel psychologisch onderzoek toont aan dat er weinig overdracht (transfer) plaatsvindt tussen het leren van abstracte regelsystemen en het redeneren bij problemen in het dagelijks leven (Van den Brink, 1993, p. 192). Ook op de bemoedigende resultaten van Nisbett, Fong en anderen is kritiek geuit. De training betrof vaak psychologiestudenten die slechts één uur een kleine training kregen. Over realistische onderwijssituaties is nog niet veel bekend (*op. cit.*, p. 195).
Toch is er hoop voor leraren en opleiders. In het voorgaande zijn we al dingen tegengekomen die helpen voor het onderwijzen van statistisch redeneren.

1. Als duidelijk is wat het onderliggende kansmechanisme is, redeneren kinderen beter (Piaget, Nisbett en anderen). Het verdient dus aanbeveling hier aandacht aan te besteden.

2. Hoe helderder de variabelen, de kansfactoren en mogelijke uitkomsten zijn, des te beter wordt er geredeneerd (Nisbett). Statistische training helpt om deze variabelen, factoren en uitkomsten te definiëren.
3. Als mensen veel weten van een bepaalde context, redeneren ze beter. Bij dagelijkse zaken als het weer, sport, toetsen en ongelukken maken de meeste mensen weinig fouten.
4. Uit voorbeelden uit de vorige paragraaf blijkt dat mensen in veel contexten intuïtief correct redeneren. Deze contexten kunnen we als aanknopingspunten gebruiken om dergelijke kennis te formaliseren. Soms kunnen we een beroep doen op het principe van representativiteit (zie ook de volgende paragraaf).
5. Enkele psychologen hebben laten zien dat wiskundige kennis die in het VWO geleerd is, volledig verdwijnt, tenzij deze kennis in het vervolgonderwijs wordt herhaald en uitgebreid. In dat geval blijkt de kennis vijftig jaar later nog aanwezig te zijn (Van den Brink, 1993, p. 193). Wat de statistiek betreft, is dit gunstig: er zijn weinig studies waar niets aan statistiek gedaan wordt. Zelfs bij sommige talenstudies wordt een statistiekvak aangeboden.

De schatting van het aantal olifanten

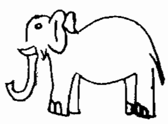
Ik heb een HAVO-brugklas het aantal olifanten op de foto aan het begin van dit artikel laten schatten. De foto had ik uitvergroot tot bijna A4. De strategieën die de leerlingen daarbij gebruikten, kunnen we grofweg in drie categorieën indelen:

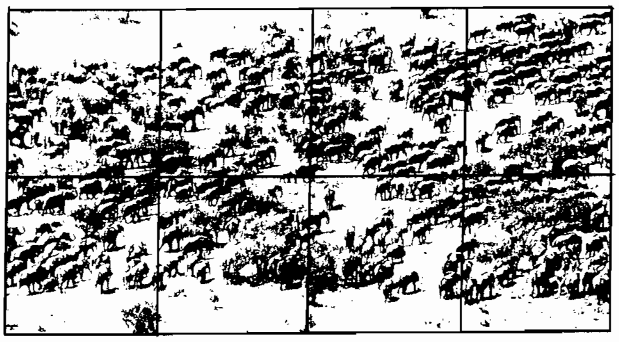
1. Maak groepen en schat het aantal in de groep, en tel alles op.
   Bijvoorbeeld: 30 + 40 + 10 + 25 + 28 + 40 = 173.
2. Maak groepjes van een vast aantal en schat hoeveel groepjes er in het totaal passen.
   Voorbeeld: 12 groepjes van 19 geeft 228.
3. Maak hokjes, tel een gemiddeld hokje en vermenigvuldig met het aantal hokjes: 35 × 8 = 280 (bij 8 hokjes). Een leerling middelde een hokje met de minste en een met de meeste olifanten:
   (19 + 39) : 2 = 29, 29 × 8 = 232.

strategie	1	2	3
aantal leerlingen	3	6	13
gemiddelde schatting	197	218	252

Opvallend aan de schattingen is dat alle leerlingen veel te laag uitkwamen: al tellend kom ik op ongeveer 330 olifanten. Vooral de eerste twee strategieën leidden tot te lage schattingen. De leerlingen die de derde strategie gebruikten, schatten nog steeds te laag, maar deden het aanzienlijk beter. (Dit is significant met p = 0,016.) De hoogste schatting was 310. Hier gold dus duidelijk niet: de meeste stemmen gelden.

A. per groep schatten.
b groep 1. 30 olifanten
groep 2. 40 olifanten
groep 3. 10 olie fanten
groep 4, 25 olifanten
groep 5. 28 olifanten
groep 6. 40 olifanten
173 olifanten





Werk van leerlingen

Laten we teruggaan naar de principes van Tversky en Kahneman. Ik vermoed dat leerlingen zich niet realiseren dat ze slecht '*toegang*' hebben tot een groot aantal. Een kleine groep van vijf of zes hoef je niet eens te tellen, maar bij grotere groepen kun je er flink naast zitten. Een verklaring van de te lage aantallen kan het principe van *verankering* zijn. Denk even terug aan het experiment waarin leerlingen 8! (= 40.320) moesten schatten. Ze kwamen over het algemeen veel te laag uit. Als ze met 1 × 2 × 3 begonnen kwamen ze zelfs lager uit (mediaan 512) dan als ze met 8 × 7 begonnen (mediaan 2.250). Ik vermoed dat bij het schatten van het aantal van een groep olifanten iets vergelijkbaars gebeurt. Leerlingen beginnen te tellen (1, 2, 3, 4, ...) en extrapoleren dan als ze geen zin meer hebben. Omdat ze een tijdje met kleine getallen werken, raken ze verankerd in te lage aantallen, en schat-

ten vervolgens te laag. Het is vermoedelijk heel anders als we eerst zouden vragen: zijn er meer of minder dan 500 olifanten? Ik vermoed dat leerlingen dan veel hoger zouden schatten, omdat ze zich aan het getal 500 gaan 'aanpassen'. Mijn vermoeden is gebaseerd op het experiment met het rad met de getallen 0 tot en met 100 en het percentage Afrikaanse landen in de NAVO.
Bij de derde strategie heb ik de indruk dat de leerlingen het principe van *representativiteit* gebruiken: ze kiezen een 'gemiddeld hokje' en vermenigvuldigen het aantal daarin met het aantal hokjes. Dit is dus een voorbeeld van correct intuïtief statistisch redeneren: een 'gemiddeld hokje' is representatief.

Bovenstaande berust slechts op vermoedens, maar het illustreert hoe je op een cognitief psychologische manier tegen wiskundige opgaven aan kunt kijken.

Arthur Bakker, Freudenthal Instituut

Literatuur

Brink, W. van den (1993). Statistisch redeneren. In: P. Koele & J. van der Pligt (red.) *Beslissen en Oordelen.* Hoofdstuk 8. Meppel: Boom.

Fong, G.T., D.H. Krantz & R.E. Nisbett (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18*, 253-292.

Kahneman, D. & A. Tversky (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.

Kahneman, D. & A. Tversky (1982). On the study of statistical intuitions. *Cognition, 11*, 131-141.

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction, 6*(1), 59-98.

Nisbett, R.E., D.H. Krantz, C. Jepson & Z. Kunda (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90*(4), 339-363.

Nisbett, R.E. & L. Ross (1980). *Human Inference. Strategies and Shortcomings of social judgment.* Englewood Cliffs, N.J.: Prentice-Hall.

Peterson, C.R. & L.R. Beach (1967). Man as a intuitive statistician. *Psychological Bulletin, 68*, 29-46.

Piaget, J. & B. Inhelder (1951/1975). *The origin of the idea of chance in children.* New York: Norton.

Shafir, E. (1993). Choosing versus rejecting: why some options are both better and worse than others. *Memory & Cognition, 21*, 546-556.

Shafir, E., I. Simonson & A. Tversky (1993). Reason-based choice. *Cognition, 49*, 11-36.

Tversky, A. & D. Kahneman (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*, 1124-1131.

Boekbespreking

Wiskundige Methoden Toegepast
Auteur: *J. Grasman*
Uitgever: *Epsilon Uitgaven, Utrecht 1999*
Herziene herdruk van de uitgave uit 1993
ISBN *90-5041-053-7*
Prijs: *f 42,50*
Bedoeld voor: Onderwijs wiskunde universiteit en hogeschool

Studenten aan de Landbouwuniversiteit te Wageningen maken in hun eerste jaar kennis met een 'basispakket' wiskunde. De leerstof daarvoor is in dit deel vastgelegd. Aan de orde komen: Infinitesimaalrekening, Getalrijen en het iteratieproces, eerste orde differentiaalvergelijkingen. Vectoren, matrices en lineaire afbeeldingen, lineaire iteratieve processen, functies van twee variabelen. Uit deze lijst blijkt dat het hier gaat om een combinatie van onderwerpen uit wiskunde A en wiskunde B. Het ingangsniveau sluit goed aan bij het eindniveau van vwo A en B. De behandeling van de onderwerpen gaat uiteraard wiskundig wat dieper op de zaak in. Bij ieder onderdeel komen een aantal toepassingen aan bod. Daarmee biedt het boek een bron aan ideeën voor proefwerkopgaven of praktische opdrachten in de verschillende profielen.

Meetkunde en Fysica
Auteur: *Henk Broer*
Uitgever: *Epsilon Uitgaven, Utrecht 1999*
ISBN *90-5041-054-5*
Prijs: *f 42,50*
Bedoeld voor: Universiteit wiskunde en natuurkunde vanaf het tweede jaar

Regelmatig verschijnen in de Epsilonreeks publicaties die gebaseerd zijn op collegedictaten van de wiskundestudie aan de universiteit. Ook dit deel is in de loop van een aantal jaren ontwikkeld vanuit een college over differentiaalmeetkunde en fysische toepassingen aan de universiteit te Groningen. De auteur geeft een overzicht van de inmiddels klassieke theorie van de differentiaalvormen. Het geheel is gelardeerd met historische verwijzingen en toepassingen in de theoretische fysica. De stof is gedegen maar goed leesbaar opgeschreven, maar zal de nodige inspanningen van de lezer vragen. Voor wie de tijd en de moed heeft, kan lezing een goed beeld geven van de stand van zaken van de wiskunde op dit terrein en van het niveau van de wiskundestudie aan de universiteit.

Sieb Kemme, Lettelbert