

ALL PEER REVIEWED! ALL REPEATABLE! ALL ADVENTURE!

# SCIENCE STORIES

## DATA ZIJN GEEN KENNIS

*Afscheidscollega  
René van Hezewijk*

Nooit meer data sets...IF als Inteeft Factor..  
Weg met Big Data! Nu met meer metafysica!

CREATED WITH **PULP-O-MIZER** COVER MAKER

# DATA ZIJN GEEN KENNIS

---

*Donderdag 8 januari 2015*

## **AFSCHEIDSCOLLEGE**

**prof. dr. René van Hezewijk**  
decaan faculteit Psychologie  
en onderwijswetenschappen  
*van 1 juli 2002 - 1 januari 2015*

**Open Universiteit**  
[www.ou.nl](http://www.ou.nl)



## SAMENVATTING

In de sociale wetenschappen, en in het bijzonder de psychologie, is de laatste jaren veel gesproken over 'data': we verzamelen data, we behoeven datamanagement en datamanagers, integriteit van data, privacykwesties, datasets, big data, syntax files ... begrippen die vanzelf lijken te spreken voor de moderne onderzoeker in de sociale wetenschappen. In mijn afscheidscollege ben ik ingegaan op de vraag hoe we kunnen denken over onderzoekgegevens, en hoe die zich verhouden tot kennis, kenniscroei, theorieën en hypothesen, of zelfs onderzoeksprogramma's waarbinnen theorieën en de daarbij horende data zich bevinden. Beantwoorden data alle

vragen die we vanuit de wetenschap stellen? Zijn alle vragen die we stellen alleen met data te beantwoorden? Impliceert een hypothese meer dan alleen een vraag die noopt tot data verzamelen? Als data als paddenstoelen in het bos zijn te verzamelen en in een datamandje worden opgeslagen, weten we dan iets? In het antwoord op deze vragen ga ik (nog voor één keer) terug naar waar wetenschap voor bedoeld was, naar de aloude discussie over inductie versus deductie, en hoe in de psychologie kennis kan groeien zonder dit te identificeren met 'data verzamelen'. Kortom: data zijn belangrijk, maar staan niet op zichzelf.

*Dit boekje is een uitgebreide versie van het afscheidscollege dat ik omwille van de tijd op 8 januari 2015 in bekorte vorm heb uitgesproken.*

## INLEIDING

Mevrouw de Rector, Mijnheer de voorzitter van het College van bestuur, Hooggeleerde, Zeergeleerde en Doorgeleerde en Lerende collega's, Geachte aanwezigen,

Er is een vraag in de psychologie die me al jaren bezighoudt: wat doen data ertoe? De vraag is niet retorisch bedoeld, het is een serieuze kwestie. De psychologie en andere wetenschappen moeten 'evidence informed' zijn. Maar het lijkt steeds meer alsof dit devies eenzijdig wordt geïnterpreteerd.

De opvattingen over 'evidence' zien we terug in de hype rond 'big data', de koorts rond 'impact scores', het

meervoudig gebruik van 'datasets', in de onverzadigbare queeste naar nog subtielere statistische technieken. Wat deze trends kenmerkt is de theoretische leegte die je aantreft als je probeert na te gaan wat de big data, wat de impact scores, wat data sets, en wat die technieken nu allemaal te betekenen hebben. 'Betekenen' hier in de betekenis van 'logisch impliceren'.

Ik wil in dit laatste college betogen dat data relevant zijn voor een theorie of hypothese, of het zijn geen data. Met andere woorden: data bestaan niet. Ze bestaan alleen in relatie tot een theorie.

Maar eerst een beetje Beckett.

**BECKETT**

*Wachten op Godot* van Samuel Beckett (Beckett, 1954) biedt een illustratie van een speciaal effect op wat mensen denken, en daarom doen. In deze 'tragicomedy in two acts' wachten Estragon en Vladimir op Godot. Punt. Ze wachten.



*Afbeelding: Vladimir, Estragon en Lucky, wachtend op Godot (Beckett, 1954)*

ESTRAGON:

Charming spot. (He turns, advances to front, halts facing auditorium.) Inspiring prospects. (He turns to Vladimir.) Let's go.

VLADIMIR:

We can't.

ESTRAGON:

Why not?

VLADIMIR:

We're waiting for Godot.

ESTRAGON:

(despairingly). Ah! (Pause.) You're sure it was here?

VLADIMIR:

What?

ESTRAGON:

That we were to wait.

VLADIMIR:

He said by the tree. (They look at the tree.) Do you see any others?

ESTRAGON:

What is it?

VLADIMIR:

I don't know. A willow.

ESTRAGON:

Where are the leaves?

VLADIMIR:

It must be dead.

ESTRAGON:

No more weeping.

VLADIMIR:

Or perhaps it's not the season.

ESTRAGON:

Looks to me more like a bush.

VLADIMIR:

A shrub.

ESTRAGON:

A bush.

VLADIMIR:

A—. What are you insinuating? That we've come to the wrong place?

ESTRAGON:

He should be here.

VLADIMIR:

He didn't say for sure he'd come.

ESTRAGON:

And if he doesn't come?

VLADIMIR:

We'll come back tomorrow.

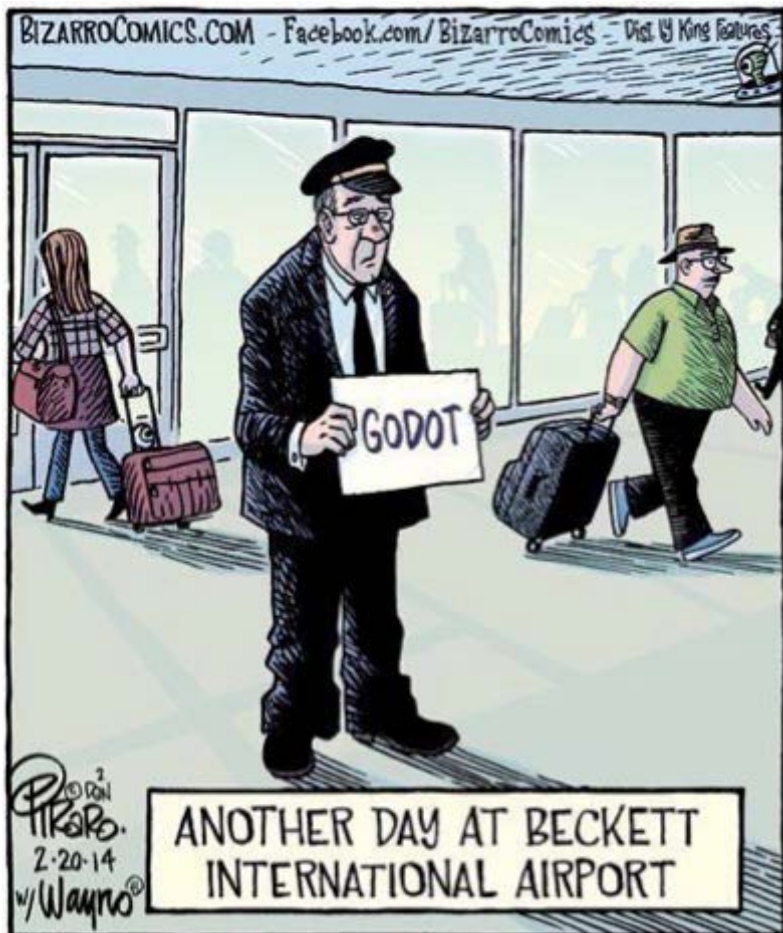
ESTRAGON:

And then the day after tomorrow.

VLADIMIR:

Possibly.

Godot bestaat, hij komt, maar wanneer? Bij welke boom? Hoe ziet Godot eruit? Als de baker, de butcher? De behaviour shaper? Ziet hij eruit als een decaan? Hij komt uit de lucht, Godot. Maar hoe? Vallend, zwevend, per parachute, vliegtuig?



Afbeelding: Another day at Beckett International Airport (Facebook, 2014)





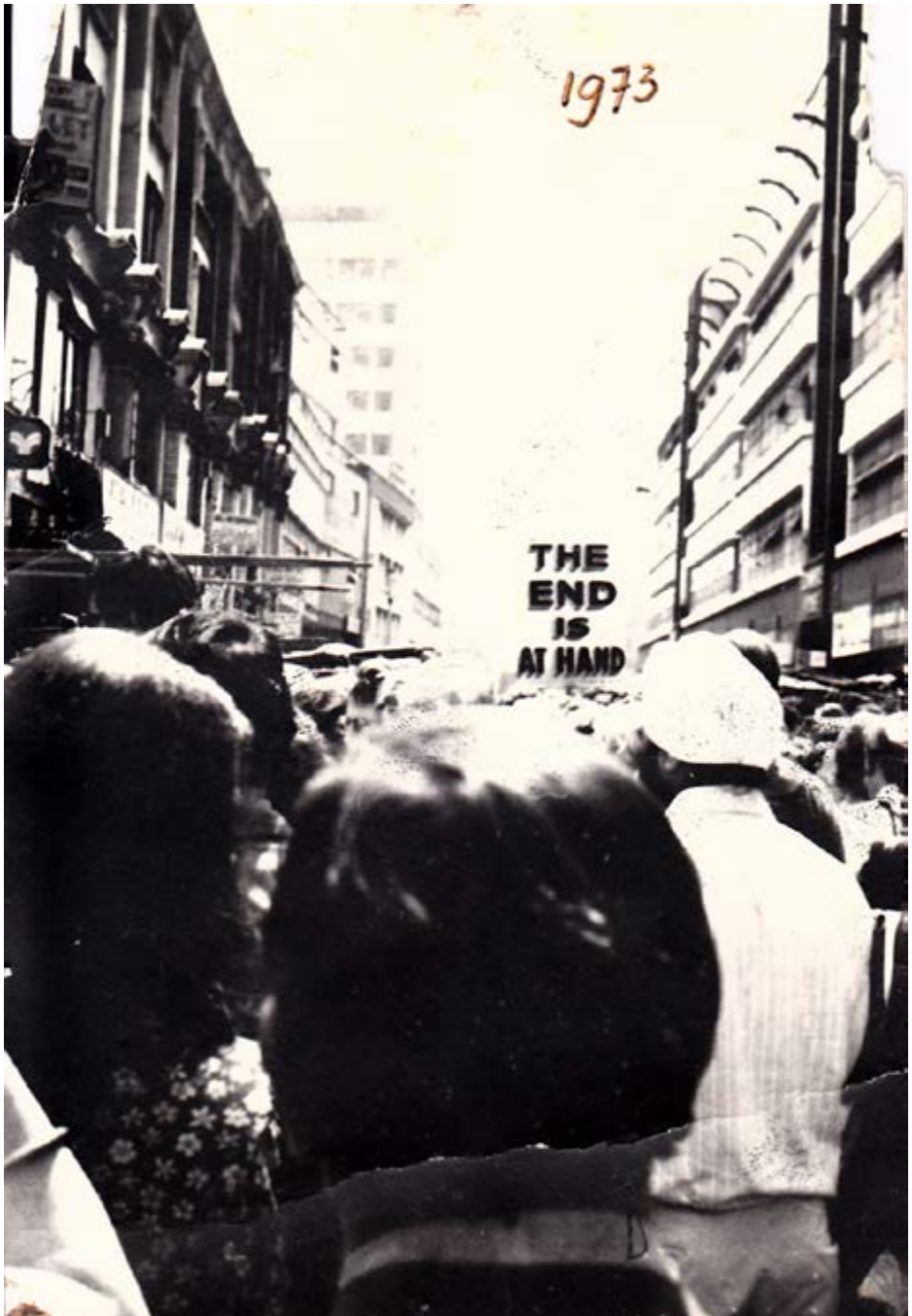
In de tweede akte  
wachten we nog  
steeds. Moeten we  
ons reisbereid houden?

*Afbeelding: Wachten op Godot,  
tweede akte (Beckett, 1954)*



Moeten we onze  
RABO-rekening  
leeghalen?

*Afbeelding: Wachtend op Godot,  
animatie*



*Afbeelding: The end is at hand in 1973*

Je vraagt je af hoe een dergelijk effect ontstaat. Laten we dit het ‘*Godot-effect*’ noemen. Het *Godot-effect* houdt in:  $x$  verwijst naar iets dat we in onze acties betrekken; dus  $x$  bestaat en zal zich te eniger tijd ergens manifesteren. Het *Godot-effect* is een illustratie van de cognitieve kracht of macht die impliciete *existentiële* beweringen over ons hebben. Ik gebruik “*existentieel*”, hier en in het vervolg in de betekenis van de propositiologica, dat wil zeggen als louter beweringen over het bestaan van iets. Ze beweren dat iets bestaat, zonder nadere aanduiding. Estragon en Vladimir hebben nog een aanknopingspunt omdat de verschijning van Godot bij de wilg schijnt plaats te hebben. In andere gevallen ontbreken aanknopingspunten.

Dit effect heeft te maken met imperfecte rationaliteit (J. W. N. Watkins, 1970), of ‘bounded rationality’ zoals Herbert Simon het noemde (Simon, 1982). Omdat het zelden mogelijk is alle benodigde informatie op tijd te verwerven en verwerken om tot een volledig rationeel besluit te komen, hebben mensen veel gemak van heuristieken die hen dagelijks helpen te overleven en die het zelfs beter doen dan de weloverwogen keuzen die we zouden kunnen maken als we alle

informatie en alle tijd hadden. Een ervan is de agency-heuristiek. Het gedrag van wat leeft of beweegt in het universum is beter te voorspellen door ervan uit te gaan dat ‘het’ een wil heeft die oorzaak is van wat gebeurt, dan door de noodzakelijke en voldoende, in tijd en plaats bepaalde voorwaarden voor de ingebouwde mechanismen te zoeken die leiden tot het vertoonde of vertonen gedrag. We zien natuurkundige en fysiologische verschijnselen liever als ‘agents’. We doen alsof ze willen wat ze doen. Mensen, kinderen, zuigelingen, beesten, pc’s, wasmachines, faculteiten .... they’ve got a will of their own, tenzij we het tegendeel merken. We gaan daar ver in en dat is nog handig ook.

Het strafrecht heeft op basis hiervan een uiterst verfijnde theorie van de menselijke wil en het menselijk weten ontwikkeld, waarin de agency-heuristiek is uitgebeend. Boos opzet, dolus malus, kleurloos opzet, voorwaardelijk opzet: het zijn varianten van de wilsvorming die achter een menselijke handeling wordt verondersteld (Strijards, 1983).

De agency-heuristiek gaat veel verder maar is minder evidence informed in vormen van magisch en religieus denken. De ‘agent’ is hier een Alwetende en Almachtige, soms een Wel-

willende. Iets overkomt je of iemand treft het slecht; dat moet wel de wil van de almachtige zijn. Natuurrampen als resultaat van een gewilde en alwetende handeling van de almachtige. Zie je wel....Godot bestaat.

Ik ben een en ander beter gaan begrijpen door de wijze waarop onderzoekgegevens en beweringen in de vorm van theorieën en hypothesen zich tot elkaar verhouden. In wat volgt

ga ik daar op in. Met een beetje propositiologica kijken we naar Poppers falsifieerbaarheidsprincipe, we dalen af naar de observatiegegevens en staan stil bij de mishandeling die wij data aandoen, en die data ons aandoen, met als voorbeelden big data, impact scores, databeheer en data sets. Maar we stijgen eerst op naar wat boven theorieën uitgaat, naar de tegenhanger van het geploeter met data.

## WETENSCHAP EN METAFYSICA

Wetenschap, zo betoogt Karl Popper al in de voorloper van zijn *Logic of scientific discovery* (Popper, 1934, 1959), kan worden onderscheiden van metafysica door als basisprincipe te nemen dat wetenschappelijke theorieën falsifieerbaar moeten zijn. Het is '...Aufgabe der Wissenschaft [...], ein theoretisches Gebäude möglichst weit weg von der Metaphysik zu errichten' (Popper, 1979, p. 393). Er moeten testsituaties (bijvoorbeeld experimenten) denkbaar zijn die een theorie kunnen weerleggen. De test-situatie moet leiden tot observaties die worden uitgedrukt in observa-

tie-uitspraken. Data dus.

Logisch gesproken zijn theorieën netwerken van samenhangende *uni-versele uitspraken* [ $\forall(x)D(x)$ ], en de bijhorende data zijn welomschreven existentiële uitspraken [ $\exists(x)D(x)$ ]. Omdat beide soorten uitspraken falsifieerbaar zijn behoren ze tot de wetenschap en kunnen dus worden afgegrensd van ... ja van wat?

Popper had, onder andere, een stevige achtergrond in de psychologie en pedagogiek. Hij ging die vakken studeren vanwege zijn activiteiten in de Oostenrijkse Schulreform-beweging. Omdat de toenmalige Denkpsycholo-



gie van Karl Bühler de enige discipline in Wenen was die Immanuel Kant serieus nam (Hacohen, 2000) kwam Popper bij Bühler terecht. Hij volgde colleges bij Bühler, die onder zijn studenten en assistenten ook Ludwig Wittgenstein, Paul Lazarsfeld, Konrad Lorenz, Egon Brunswik en Charlotte Bühler telde. Later in zijn studie richt hij zich op kennistheoretische vragen en komt hij via Otto Neurath in aanraking met de Wiener Kreis. Popper is overigens nooit lid geweest van de Wiener Kreis. Hij was geen positivist. Wat Popper in de kantiaanse aanpak van de Würzburger Schule en Bühler aantrok was wat de Würzburgers onderscheidde van de ideeën van de, toen, grote naam in de psychologie, Wilhelm Wundt. Volgens Wundt staan waarnemingen aan de basis van bewuste kennis. Uit de associatie van sensaties vormen zich beelden, en de beelden vormen de grondstof van

het denken. Bühler en de Würzburgers toonden evenwel aan dat denken ook zonder beelden geschiedt, dat denken niet bestaat uit associaties tussen elementaire sensaties, en zelfs dat niet alle sensaties tot het bewustzijn doordringen maar toch hun werk doen. Gedachten zijn juist gehele mentale, vaak talige structuren, '*Gestalten*'. Zij vormen de basis van het denken en ze organiseren en selecteren sensorische gegevens.

Popper schreef er de in de studie verplichte verhandeling over, maar ging toch de weg op van de wetenschapsfilosofie. De denkstijl van vooral Bühler is er in terug te vinden: anti-associationistisch, anti-inductivistisch, anti-elementaristisch, deductivistisch, denken als probleem oplossen, en gedachten als sturend voor de waarneming, de waarneming als corrigerend voor het denken.

## FALSIFIËREN EN VERIFIËREN

Hoewel in de meest recente wetenschapsfilosofische literatuur de aandacht voor de structuur van wetenschappelijke kennis minder groot is dan dertig jaar geleden, geldt nog steeds dat wetenschap niet zoekt naar de essentie van objecten, naar definities, naar classificatiesystemen of zelfs naar alleen maar voorspellingen en interventies. Kennis zit in theorieën, en theorieën zijn netwerken van beweringen over de relaties tussen concepten, welke concepten verwijzen naar entiteiten of gebeurtenissen die worden verondersteld werkelijk te bestaan.

**Theorieën zijn wetenschappelijk als ze empirisch toetsbaar zijn; dat wil zeggen ze impliceren hypothesen die kunnen worden gefalsifieerd door observeerbare gebeurtenissen**

(ook wel feiten genoemd) (Brown, 1977; Chalmers, 1999; Hempel, 1962; Lakatos, 1970; Miller, 1994; Musgrave, 1993; Popper, 1934, 1959, 1963, 1972; Salmon, 1984; J.

W. N. Watkins, 1984). Kennis zit dus niet in data, theorieën zijn meer dan associaties van gegevens. Theorieën geven verbanden in de wereld aan, en verbeteren de verwachtingen over wat er kan of zal gebeuren met gespecificeerde dingen en mensen.

Wie het daarmee eens is, is al een heel eind bij waar ik wil uitkomen. Maar de duivel schuilt in het detail. De (vetgedrukte) zin zegt niet dat theorieën geen niet-observeerbare elementen kunnen bevatten. Het door Popper gearticuleerde onderscheid tussen wat we wetenschappelijk en wat we metafysisch noemen, illustreert dat heel aardig. Onthoud dat 1) theorieën niet-observeerbare elementen kunnen bevatten, en dat 2) Popper het onderscheid maakte tussen wetenschappelijk en metafysisch. Nemen we beide bij elkaar, dan is de consequentie dat sommige implicaties van wetenschappelijke theorieën niet empirisch zijn, en dus niet falsifieerbaar, en dus metafysisch. Maar wel betekenisvol en nuttig. We onderscheiden beide maar scheiden ze niet. Daarom moeten we Popper ook geen positivist noemen.

Om het onderscheid tussen wetenschap en metafysica te kunnen maken introduceerde Popper het falsifieerbaarheidsprincipe. Wat de implicaties van dat principe zijn is door hemzelf, maar ook door bijvoorbeeld Imre Lakatos en John Watkins verder uitgewerkt (J. W. N. Watkins,

1957, 1958, 1960, 1975, 1978). Ik beperk me nu tot Watkins' versie.

Watkins onderscheidt vier typen beweringen (of uitspraken) die een rol spelen in wetenschappelijke theorieën: twee empirische, twee niet-empirische.

### **Empirische uitspraken**

Tot de empirische uitspraken behoren:

- *Observatie-uitspraken* (ook wel *circumscribed existential statements* of *singular statements*); deze horen ook tot de empirische uitspraken. Bijvoorbeeld een bewering als 'Er is op dit moment een roker in de aula'. In de propositiologica worden existentiële uitspraken aangeduid met  $\exists(x)$  dus voor de gelegenheid duiden we die welomschreven existentiële uitspraken aan als volgt:  $\exists(x)p, t$ . Een observatie-uitspraak is empirisch want kan worden gefalsifieerd door een contraire uitspraak ('Er is nu geen roker in de zaal'), maar kan ook worden geverifieerd door nog een keer zo'n uitspraak van iemand anders die de roker ziet. Met dit type uitspraken kun je een universele uitspraak falsifiëren. Voorspellingen die men uit een theorie afleidt hebben de logische vorm van een observatie-uitspraak: 'Er zal omstreeks kwart over vijf een roker staan bij de buitendeur van dit gebouw'. Die uitspraak is te verifiëren. Maar een universele uitspraak kan er niet door worden geverifieerd (in de betekenis van volledig confirmeren). Een universele uitspraak zegt immers iets over oneindig veel gevallen.

- *Universele uitspraken*, zoals 'Alle gedragingen worden herhaald als zij worden beloond' : in wat meer formele termen
  - 1) *Voor alle gedragingen R en voor alle stimuli S geldt dat als R in stimulus conditie S wordt versterkt door V, dan zal R herhaald worden in S.*

Dit type uitspraak zou in de propositielogica de universele quantor hebben:  $\forall(x)$  Universele uitspraken zijn falsifieerbaar. Ze kunnen empirisch of experimenteel worden getoetst als we de technische middelen hebben de stimulussituatie en de versterking (reinforcement) te controleren. Let wel: ze zijn ook falsifieerbaar als we de technische middelen nog niet beheersen. Het gaat om het karakter van dit type uitspraken.

### **Niet-empirische uitspraken**

De niet-empirische uitspraken zijn per definitie metafysische uitspraken.

- *Existentiële uitspraken*: beweringen over het bestaan van iets. 'Er is een X', of 'Er zijn X'en', of zelfs 'Er zijn X'en zonder eigenschap Y'. Formeel:  $\exists(x)$ . Ze zijn belangrijk ook al lijken ze triviaal. Existentiële uitspraken kunnen zonder meer worden afgeleid uit een theorie. Uit voorbeeld (1) hierboven volgt bijvoorbeeld dat 'Er bestaan stimuli'. Daarmee wordt ontkend dat je bewuste keuzes of wilsbesluiten hebt die 'van binnenuit' aanzetten tot gedrag.

Kenmerkend is onder andere:

- De ontkenning van een universele uitspraak is ook een existentiële uitspraak:
- 2)  $\sim[\text{Alle zwanen zijn wit}] \rightarrow [\text{Er is een niet-witte zwaan}]$ .
- Existentiële uitspraken kunnen niet worden gefalsificeerd. Ze kunnen wel worden geverifieerd: aan één zwarte zwaan heb je genoeg. Of 'Er zijn yeti's': vind er twee en we zijn blij want 'Er bestaan yeti's' is bevestigd.





*Afbeelding: Er bestaat minstens een Yeti (Kuifje in Tibet, 1960)*

Dergelijke beweringen kunnen uitnodigen tot zoeken. Als het uit een serieuze theorie afgeleide, existentiële uitspraken zijn hebben ze eigenlijk het karakter van een nog niet voldoende omschreven voorspelling; als we een verklaring zoeken voor de voetsporen in de sneeuw komt wellicht de theorie

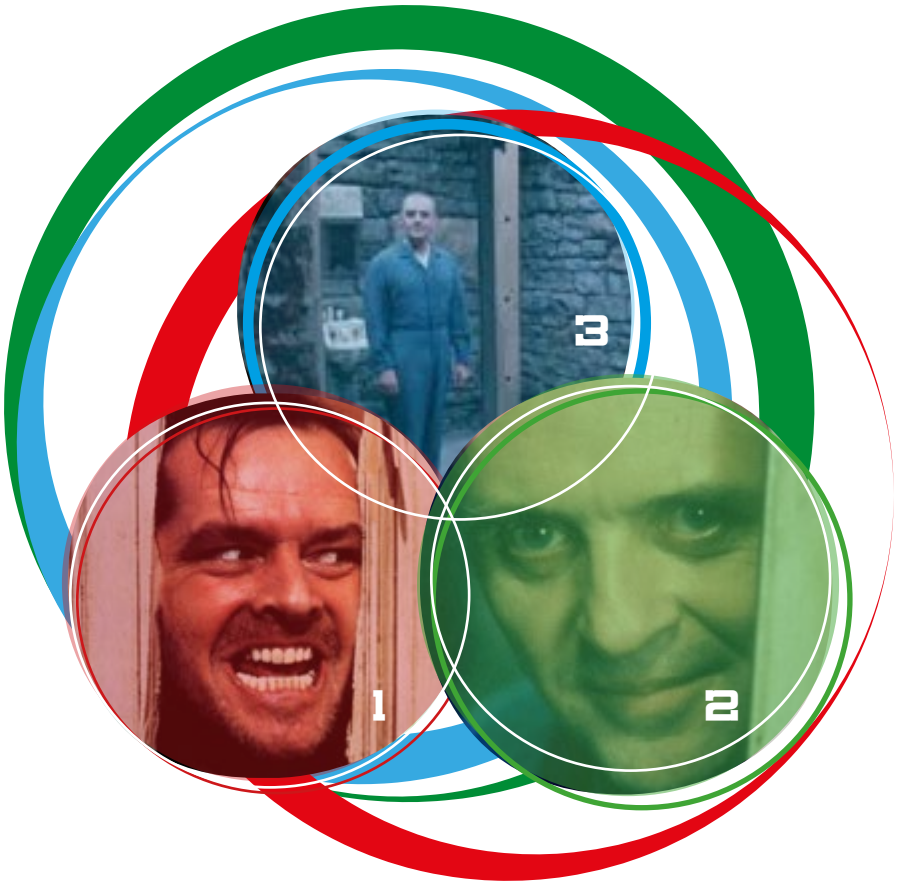
op dat 'Alle yeti's zijn groot en hebben dus grote voeten; ik zie verse sporen van grote voeten; dus voorspel ik dat zich over enkele ogenblikken een yeti manifesteert.' Maar het kunnen ook profetieën zijn, zoals: 'Voorwaar ik zeg u, het einde is nabij'. Daar zit het verschil tussen heuristiek en profetie.



*Afbeelding: The end is at hand in 1980*

Hoe metafysisch existentiële uitspraken ook zijn, ze spelen een rol in de wetenschap. Het zijn voorlopers van welomschreven existentiële uitspraken (observatie-uitspraken). Het kunnen aanvankelijk theoretische constructen zijn die een soort 'theoretische ontologie' realiseren. Als ze zijn afgeleid uit een theorie geven ze aan wat er moet bestaan om de theorie te toetsen. Is de theorie op een ander deel van het netwerk van concepten getoetst en (nog) niet gefalsifieerd, dan versterkt dat impliciet de

claim op het bestaan van de betrokken entiteit, ook al is er geen direct bewijs voor. Het bestaan van Brout-Englert-Higgs-deeltje, beter bekend als higgsboson of Higgs-deeltje, is lange tijd van dien aard geweest, tot het experiment met de grote centrifuge in gang was gezet, waarna het bestaan ervan werd geverifieerd. Kortom: existentiële beweringen kunnen heel krachtig zijn. Het Godot-effect is er op gebaseerd, en speelt ook een rol in films als *The Shining* en *The silence of the lambs*.



*Afbeelding: 1 Jack Nicholson in The Shining*

*2 Anthony Hopkins als dr. Hannibal Lecter*

*3 Anthony Hopkins als dr. Hannibal Lecter*

Er is nog een ander type non-empirische uitspraak:

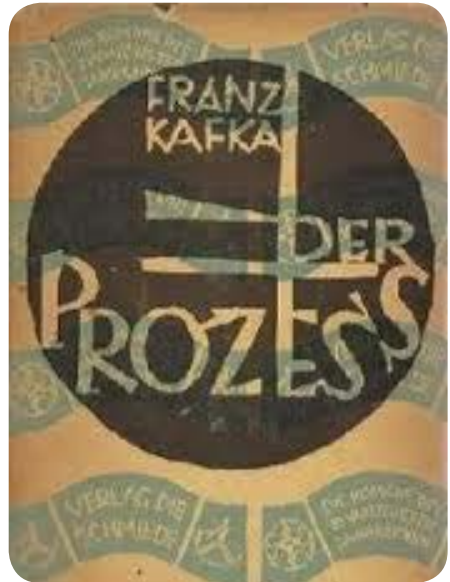
De zogenaamde ‘Voor-*alle-X-is-er-een-of-andere-Y*’ uitspraken. Dat zijn uitspraken die je niet kun verifiëren en niet kunt falsifiëren. Verifiëren kan niet door het universele deel van de uitspraak (je kunt niet elke ooit bestaande of bestaand hebbende *x* op die eigenschap testen); en ze zijn niet te falsifiëren door het tweede, existentiële deel : je kunt volhouden dat er een of andere *y* moet zijn, ook al is die nog niet bekend. Zijn dit me-

tafysische uitspraken? Ja, ze zijn niet empirisch. Zijn dit in de wetenschap belangrijke uitspraken? Ja, ze vormen een heuristiek voor verder onderzoek: ‘Voor elk gedrag is er een of andere versterkende stimulus. De laatste uitspraak is logische gesproken volkomen valide uit het eerdere voorbeeld af te leiden. Of wat dacht u van ‘Voor elke gebeurtenis is er een oorzaak’.

Maar niet alleen in de wetenschap spelen ze een rol. Neem Kafka, *Der Prozess*.



*Afbeelding: Kafka*



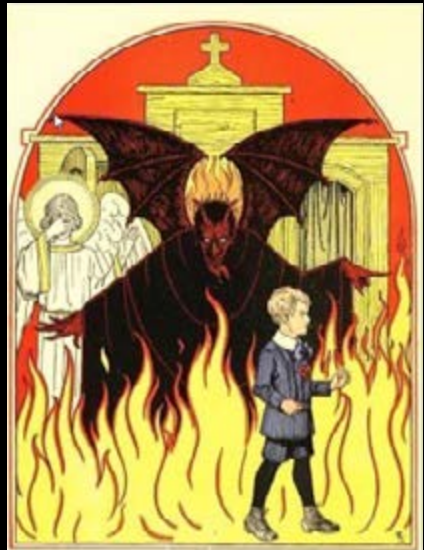
*Afbeelding: Der Prozess*

Een citaat uit *Het Proces*:

Iemand moest kwaad van Josef K. hebben gesproken, want zonder dat hij iets slechts had gedaan werd hij op een ochtend gearresteerd. (p. 5) 'Ik maak dat op uit het feit dat ik aangeklaagd ben, maar niet de minste schuld kan ontdekken op grond waarvan men mij zou kunnen aanklagen', zegt K. tegen de opziener van de bewakers die hem arresteren. (p. 26) (Kafka, 2014)

K. wordt na zijn arrestatie niet opgesloten in een cel, maar raakt verder verstrikt in ondoorzichtige beschuldigingen. Hij heeft iets misdään. Hij is immers gearresteerd. Maar wat? En het wordt alleen maar erger. Wat hij ook doet, het wordt kennelijk door zijn tribunaal als overtreding beschouwd.

Het is de burgerlijke opvoeding ten voeten uit: je doet iets, je krijgt straf, maar je weet niet precies wat er fout was. Het eerste-communieboekje staat nog helder voor de geest. De zonden bleven vaag. Je bent bij voorbaat schuldig. Bij de douane staat het zweet in je handen, zelfs als je geen smokkelwaar bij je hebt. De regel 'Voor elke zonde is er een straf' raakt omgedraaid: 'Voor elke straf is er wel een of andere zonde begaan'. Uiteindelijk erkent ook Jozef K. zijn niet begane zonde en laat toe dat hij wordt veroordeeld en ter dood gebracht.



*Afbeelding: Na een slechte biecht.  
Uit Versteeg, M.C., 1928, p.34*

Als heuristiek speelt deze bewering een vitale rol in de wetenschap. Op zichzelf genomen lijken ze te verwijzen naar een magisch universum – een ‘haunted universe’ (J.W.N. Watkins, 1958). Omdat ze een sterke heuristiek bieden om op een bepaalde manier nieuwe kennis te zoeken, en omdat ze ook nog eens worden geïmpliceerd door succesvolle theorieën, zijn ze belangrijk in de wetenschap. Ze bewijzen niets maar geven aan wat niet het geval kan zijn, welke empirische theorieën niet kunnen voorkomen – bijvoorbeeld dat onvooroorzaakte gebeurtenissen (bijvoorbeeld wonderen) niet bestaan, dat de geschiedenis geen wetmatigheden kent waar individuen geen rol in spelen, dat de bewuste wil een illusie is, enz. In romans zie je ze terug als het Kafka-effect.

Popper maakte dus een werkbaar onderscheid tussen wetenschap en metafysica, maar bracht geen scheiding aan. Hij gaf al in zijn vroegste werken aan dat metafysica een belangrijke

en betekenisvolle relatie had tot de wetenschap. Watkins heeft dat verder uitgewerkt (J.W.N. Watkins, 1958). Elke weerlegbare theorie impliceert (logischerwijze) een metafysische kern die invloed uitoefent op andere theorieën of zelfs andere disciplines en andere onderzoekprogramma’s. Maar het is niet andersom. Metafysica’s zijn niet *ins Blaue hinein* te bedenken. De volgorde is niet ‘eerst de metafysica dan theorie’. ‘Eerst theorie, dan hypothesen, en dat impliceert vervolgens een metafysica’  $[T \rightarrow H \rightarrow M]$  is wel een optie. De tweede stap blijft vaak impliciet. De metafysische kern vormt echter het cement tussen de bouwstenen, de lijnen tussen de knopen van het netwerk. Ze hebben een heuristische werking, en ze suggereren een ontologie die aangeeft waaruit we kunnen denken dat de wereld bestaat, om daarmee indirect het bestaan van ‘andere werelden’ te kunnen ontkennen<sup>1</sup>. Theorieën zijn, met andere woorden, metafysisch geladen.



## MYSTIEK UNIVERSUM

De volgorde  $[T \rightarrow H \rightarrow M]$  is niet de enige optie. Zowel in natuurkunde als in psychologie komt het voor dat er eerst een metafysica is geformuleerd die vervolgens tot succesvolle theorieën heeft geleid (ook al bleken die later niet altijd de vergelijking met andere theorieën en metafysica's te doorstaan). Sigmund Freuds theorie is voor een groot deel een metafysica gebleven, maar daarom niet minder interessant (Panhuysen, 1990). Een aantal problemen leidde tot speculatieve ideeën die soms zelfs tijdelijk de toetsing doorstonden. Maar wat C. G. Jung ermee deed was het proces omdraaien. De volgorde veranderde in  $[M \rightarrow T \rightarrow H]$ . Sommige elementen van Jungs metafysica kwamen weliswaar terecht in een theorie met als elementen extravert-introvert (bijvoorbeeld van Eysenck), en zijn nog steeds te vinden in het diagnostisch systeem van de Big Five. Maar net als in de homeopathie werd de geïmpliceerde metafysische kern zelf beschouwd als een getoetste theorie. Freud zag bijvoorbeeld libido aanvan-

kelijk als een seksuele drift, later als vitale energie. Jung maakte daar psychische energie met een religieuze inhoud van. In de homeopathie zijn enkele verifieerbare, maar niet falsifieerbare implicaties van negentiende-eeuwse natuurwetenschappelijke/medische theorieën over de immuunsysteem-activerende werking van kleine hoeveelheden viraal materiaal verbreed tot een algemene stelling dat 'voor elke ziekte is er een of andere op de ziekte gelijkende stof die bij minimale toediening een helende en/of preventieve werking heeft', of zelfs 'voor elke lichamelijke toestand is er een of andere daarop gelijkende stof die bij inname van een minimale hoeveelheid ervan de lichamelijke werking versterkt'. Bijvoorbeeld, walnoten zouden goed zijn voor je hersenen (want walnoten lijken op hersenen). In een universum dat vol is van 'correspondenties' speelt gelijkenis – 'zo hier zo daar' – de rol die oorzakelijke verbanden spelen in ons wereldbeeld (Lovejoy, 1964; Tillyard, 1943 / 1986; Yates, 1964)<sup>2</sup>.

Wat we hier missen is de toets van de theorie zelf. Via de geïmpliceerde metafysica van een (wel getoetste) theorie wordt een theorie geponeerd

die niet werd getoetst maar als ‘wetenschappelijk bewezen’ wordt verkocht. De ‘metafysische geladenheid’ suggereert nog slechts de waarheid.

## EMPIRISCHE ASPECTEN

Tot zover wat je gene kant van hypothesen en theorieën zou kunnen noemen: non-empirische implicaties van empirische beweringen. Nu de verhouding tot data, deze zijde van theorieën en hypothesen. Theorieën worden in het algemeen gezien als netwerken van universele beweringen. Ze zijn empirisch want falsifieerbaar. Of ze zijn al gefalsificeerd. Falsifiëren is confronteren met observatie-uitspraken. Dat geschiedt mede met behulp van instrumenten die als hulpstukken dienen bij het vinden van de onderzoekgegevens.

Basisuitspraken hebben uiteindelijk het karakter dat we al uit de deden deden, namelijk dat van in tijd en plaats welomschreven existentiële uitspraken. Natuurlijk willen we

er veel van hebben, maar dat is een kwestie van sommeren over de welomschreven existentiële uitspraken. Data is dus altijd meervoud, maar een van de tekenen die wijzen op de obsessie met data is dat er Amerikanen zijn en in hun voetsporen Europeanen – Nederlanders – die ‘data’ als een enkelvoud gebruiken. ‘De data spreekt voor zichzelf’, en vergelijkbare uitspraken. Maar data spreekt niet, en data spreken ook niet, en zeker niet voor zichzelf. Hoe veelvuldig de waarnemingen ook, het blijft niet meer dan een optelsommetje van observaties in het licht van de eeuwigheid. Wat moeten we dan denken van de obsessie met data die we in allerlei vormen terugzien in de sociale wetenschappen?

## VERFIJNDE NIVEAUS VAN BASISUITSPRAKEN

Wat data zijn voor de een zijn ze niet voor een ander. Ik ontleen aan Watkins de volgende tabel (J. W. N. Watkins, 1984).

| Level | Definition                                                                                                                                                                                                           | Falsified by         | Supported by                                                                                              | Verified by          | Falsifying                                                                                                               | Supporting                                                                                                               | Verifying            |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------------------------------------|----------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|----------------------|
| L-0   | perceptual reports of a first-person, here-and-now type (e.g. 'I feel a sharp pain in my left lower wisdom tooth')                                                                                                   | --                   | L-1 statements about behavior or brain events supposed by L-2 statements to correlate with L-0 statements | Other L-0 statements | L-1 statements<br>L-2 statements about behavior or brain events supposed to correlate L-1 statements with L-0 statements | L-1 statements<br>L-2 statements about behavior or brain events supposed to correlate L-1 statements with L-0 statements | Other L-0 statements |
| L-1   | singular statements about observable things or events (e.g. 'The dentist is using his drill on his left lower wisdom tooth' or 'The dentist showed his drill and had it spinning without touching the wisdom tooth') | Other L-1 statements | L-0 statements<br>L-2 statements                                                                          | Other L-1 statements | L-2 statements                                                                                                           | L-2 statements                                                                                                           | --                   |

| Level | Definition                                                                                                                                                                   | Falsified by   | Supported by                                       | Verified by | Falsifying                       | Supporting                                         | Verifying                  |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------------------------------------|-------------|----------------------------------|----------------------------------------------------|----------------------------|
| L-2   | empirical generalizations about regularities displayed by observable things and events (e.g. 'A suggestion of forthcoming pain is followed by experienced (reported) pain'); | L-1 statements | L-1 statements<br>L-3 statements<br>L-4 statements | --          | L-3 statements<br>L-4 statements | L-1 statements<br>L-3 statements<br>L-4 statements | --                         |
| L-3   | exact experimental laws concerning measurable physical magnitudes (e.g. Fechner's Law of relation between magnitude of sensation and physical magnitude);                    | L-2 statements | L-2 statements                                     | --          | L-4 statements                   | L-4 statements                                     |                            |
| L-4   | scientific theories that are universal and exact but postulate unobservable entities (e.g. the Gate Theory of pain suggested by Melzack & Wall).                             | L-3 statements | L-2 statements<br>L-3 statements                   | --          | --                               | It's own metaphysical core                         | It's own metaphysical core |

*Tabel: Niveaus van beweringen en hun relatie tot bewijsmateriaal in de psychologie. Falsification betekent weerlegging in geval van tegenspraak; verification betekent accepteren in het geval van volledige en voltallige overeenstemming; support betekent corroboratie: tentatieve aanvaarding wanneer de weerlegging (nog steeds) faalt. Naar Watkins (1984). Science and scepticism. London: Hutchinson.*

De tabel verfijnt de eerder gegeven argumentatie, en is voor de psychologie belangrijk omdat het onderscheid tussen L-0 en L-1 voortdurend een rol speelt. In ESM onderzoek bijvoorbeeld, waar in onze faculteit veel gebruik van wordt gemaakt, is het resultaat een grote hoeveelheid L-0-uitspraken, dus rapporten van de eerste persoon over hun ervaringen. Tegelijk geven zij echter ook aan onder welke omstandigheden de gerapporteerde ervaringen ontstonden. Dat zijn L-1-uitspraken. Als we dan kijken naar hoe die zich verhouden dan moge duidelijk zijn dat L-0-uitspraken niet te falsifiëren zijn, maar worden ondersteund door L-1-uitspraken die worden gesuggereerd door L-2-uitspraken die iets beweren over de correlatie tussen L-0 en L-1. Met andere woorden: L-0-uitspraken zijn niet onafhankelijk voor zover zij een rol kunnen spelen in wetenschappelijke theorieën. Die ESM-rapportages moeten dus per persoon netjes bij elkaar blijven. Maar ze zijn ook op een andere manier interessant, omdat ze niet alleen een steekproef vormen van proefpersonen, zoals we in de psychologie gewoon zijn, maar ook helpen om een steekproef te krijgen uit de omstandigheden die mensen

meemaken en de waarschijnlijkheid waarmee dat per persoon gebeurt, en omdat ze helpen een steekproef te krijgen uit de reacties van mensen en de waarschijnlijkheid van reactietypen waarmee dat per persoon kan gebeuren. Ik kan er nu niet verder op ingaan, maar ik hoor nu Egon Brunswik instemmend brommen, in zijn graf (Brunswik, 1943, 1947, 1955). De tabel laat nog meer zien dat voor psychologen relevant is maar niet wordt ingezien. L-4-uitspraken, bijvoorbeeld, ondersteunen en verifiëren zelfs hun eigen metafysische kern. Dat wil zeggen dat een pijntheorie als die van Melzack, of de cognitieve-dissonantietheorie van Festinger, of de machtafstandstheorie van Mauk Mulder, of de Theory of Bounded Rationality, enzovoort, theorieën van het niveau L-4 zijn die een metafysica impliceren. Zij ondersteunen L-2-uitspraken – empirische generalisaties – maar kunnen er ook door worden ondermijnd. Dat betekent dat ook empirische generalisaties theoretisch geladen kunnen zijn. Omdat empirische generalisaties de brug vormen naar observaties, zien we hier heel duidelijk dat kennis de verwevenheid van observatie-uitspraken en theoretische uitspraken behelst.

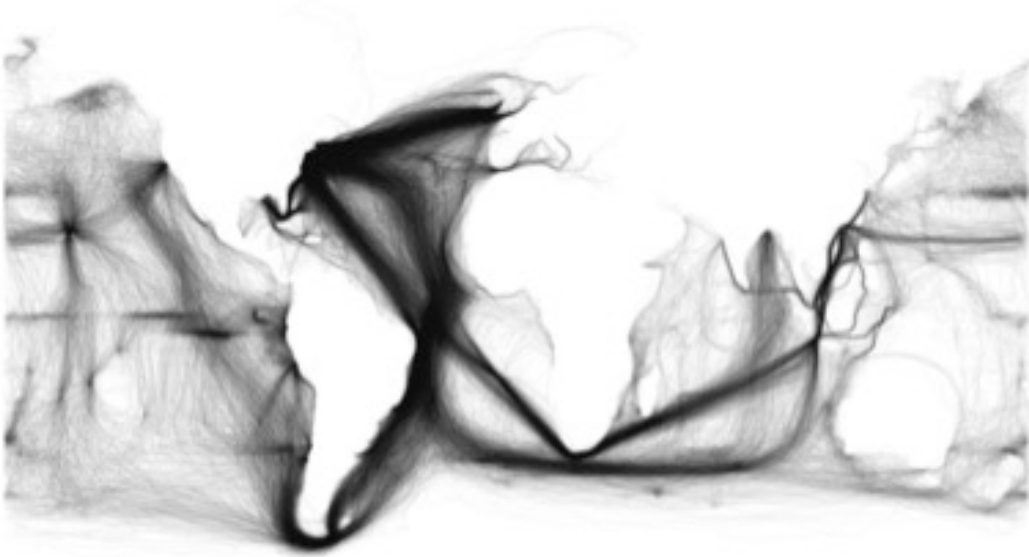
## BIG DATA

'Big data' is de nieuwe hype in de sociale wetenschappen. Big data zijn gegevens die door gigagebruikers van servers, zoals Google, Apple, Amazon, worden geëxtraheerd uit digitale activiteiten van individuen. Vaak weten die individuen niet hoe, of zelfs niet dat, die gegevens worden opgeslagen. Als sociale wetenschappers er toegang toe krijgen (als dat mag van de ethische commissies, want die zijn natuurlijk veel scherper en consequenter dan Google c.s.) dan beschikken zij over grote hoeveelheden gegevens die interessant kunnen zijn, niet alleen door hun aantal, maar ook door het feit dat het unobtrusive measures zijn. Unobtrusive measures zijn meetgegevens van personen die niet weten dat ze aan een meting deelnemen. De veronderstelling is dat daarmee een nadelige invloed op de toets van een theorie wordt vermeden.



*Afbeelding: [www.VentureBeat.com](http://www.VentureBeat.com), 2014*





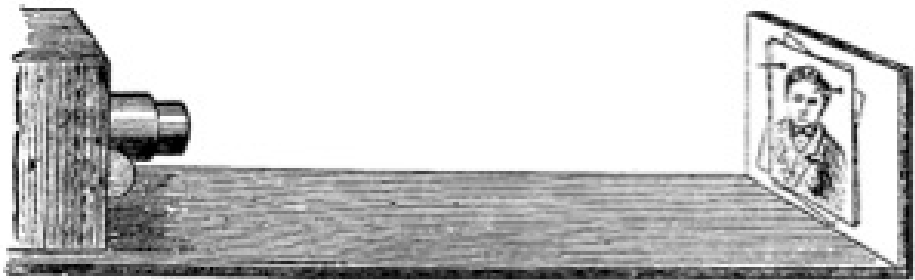
*Afbeelding: Smoke trail van logboeken van schepen in de 18e en 19e eeuw ([www.VentureBeat.com](http://www.VentureBeat.com), 2014)*

Een prachtig voorbeeld is het 'smoke trail' van de logboeken van schepen in de achttiende en negentiende eeuw. Elke reis van een schip is een lijntje, en de lijntjes gezamenlijk vormen de visualisatie. Het geeft een prachtig beeld. Drie dikke bundels die vanuit de Oostkust van de VS naar respectievelijk Groot Brittannië, naar West Afrika (en verder), en naar de Oostkust van Zuid Amerika gaan. De web site waar dit vandaan komt (VentureBeat, 2014) vertelt erbij dat de visualisatie laat zien hoe in de achttiende en negentiende eeuw de machtsbasis verschoof van Europa, in het bijzonder van Groot-Brittannië, naar de Verenigde Staten.

Wat mij fascineert is *hoe* de data, gevisualiseerd, meteen een beeld oproepen van wat er aan de hand was in die tijd. Wat er precies getoond wordt is de gevisualiseerde aggregatie van gebeurtenissen. Wat er gesuggereerd wordt is dat er een patroon in zit. Onze eerste reactie is wat psycholoog Bühler al in het begin van de 20e eeuw noemde een Aha-erlebnis.

Wat we niet beseffen is hoeveel kennis we eigenlijk al – impliciet – in het plaatje stoppen om er de betekenis van te zien. Wat we evenmin beseffen is dat die betekenis per persoon anders kan zijn, juist omdat veel veronderstellingen impliciet blijven.

De wijze van visualiseren is technisch verbeterd, maar doet sterk denken aan wat in de negentiende eeuw de

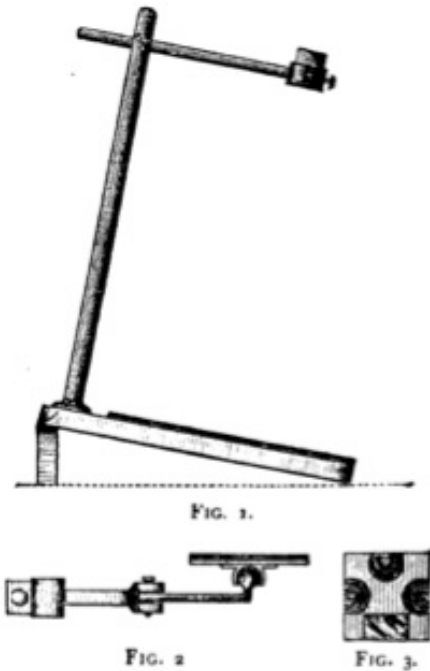


*Afbeelding: De Physiognomic Synthesizer van Sir Francis Galton (1878)*

Het bevestigt wat je altijd al dacht: er gaan veel schepen op en neer, heen en weer, en o ja, natuurlijk klopt het dat in die eeuwen steeds meer schepen vanuit New York dan vanuit Liverpool de wereld overgingen, en dat dat ‘dus’ een verschuiving van macht indiceert.

neef van Darwin, Sir Francis Galton, al deed met zijn Physiognomic Synthesizer. Hij projecteerde de negatieven van bijvoorbeeld criminelen uit de gevangenis over elkaar heen om te komen tot een gegeneraliseerd portret van – jawel – ‘de crimineel’.

Dat resulteerde na vele tussenstappen in het volgende portret.



*Afbeelding: Apparaat waarmee Galton projecties voor de 'physiognomic synthesizer' op de juiste plaats kreeg.*

*Afbeelding: Representatie van het resultaat van een fysiognomische synthese (Galton, 1878)*

De tussenstappen van Galton zijn interessant. In een artikel in *Nature* legt Galton de techniek uit waarmee je zorgt dat de omtrekken van het gezicht op de juiste plaats worden afgebeeld (Galton, 1878).

Maar hoe zijn big data verkregen? Is de gebruikte methode beschikbaar? Hoe vindt selectie plaats? De grote hoeveelheid data maakt het gevaar van sample error en vooral sample bias niet minder (Harford, 2014), eerder groter. De impliciete veronderstelling is dat er patronen zijn om te onthullen. Het data-extractieproces is belangrijker dan de data zelf. Maar veel van de interpretatie wordt overgelaten aan de toeschouwer. Big data enthousiastelingen menen dat correlaties voldoende zijn dat er geen statistische analyse meer nodig is, dat we het einde van de theorie beleven, en wat dies meer zij. 'Complete bollocks, absolute nonsense, volgens David Spiegelhalter van Cambridge University (Harford, 2014, p. 15).

Hoe, en welke conclusie trekt degene die naar de plaatjes kijkt. We hebben tekst nodig om te weten wat we in het smoke trail zien, maar we weten nog niet wat dat betekent. Dat er een machtsverschuiving plaatsvond is er niet uit op te maken omdat de lijnen geen richting hebben, en geen inhoud in termen van macht, er geen periodes of stadia in te zien zijn, enzovoort. Het wordt overgelaten aan de fantasie van de lezer. Tenzij we zoveel grip hebben op de meetwijze, op het waarom van juist deze data-

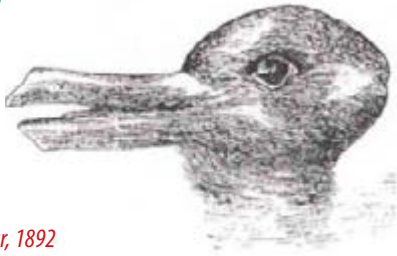
verzameling en het waartoe van de oorspronkelijke visualisatie, de selectie van data, enzovoort, dat er een theoretisch interessant beeld uit is te halen waarmee we verder komen. Er wordt voorbij gegaan aan enerzijds de theoriegeladenheid van data en anderzijds de talloze projectiemogelijkheden. Maar data bestaan niet in het wild, het zijn altijd waarnemingen. We zijn weer terug bij het grote debat van een eeuw of wat geleden tussen de Würzburger Schule en de Wundtschool in Berlijn. Wilhelm Wundt ging ervan uit dat denken (cognitie) voortkwam uit het associëren van bewustzijnselementen die op hun beurt uit sensaties bestonden. Het resultaat kwam in beelden, en denken zou dus geschieden in beelden (Van Hezewijk & De Vries, 1989). De Würzburgers, in het bijzonder Karl Bühler en Narziß Ach, toonden echter aan dat denken niet uitsluitend in beelden geschiedt en wordt gestuurd door een probleem (Aufgabe) of een verwachting, een 'determinierender Tendenz'<sup>3</sup>.

Toegepast op big data (en learning analytics) vraag ik mij dus af of de plaatjes juist voor 'elck wat wils' kunnen inhouden. Je ziet erin wat je erin wilt zien. Voor de marketingwereld is dat misschien interessant, maar levert het ook verklaringen op? Is het

kennis? Zoals Loesje al zei: 'Je moet niet alles geloven wat je denkt.' Net zoals het menselijk oog slechts voor een beperkt spectrum van waarnemingen gevoelig is, en net zoals zelfs wat het oog bereikt nog niet altijd wordt gezien omdat de cognitieve instelling onze interpretatie van de sensorische informatie op diverse niveaus selecteert, net zo geldt dat we selectief ontvankelijk zijn voor de visualisatie van de data. Wat ze ons te zeggen hebben is altijd iets in het licht van een of andere theorie: gegevens zijn gegevens ten opzichte van iets. De haas-eend-figuur – bekend van onder andere Ludwig Wittgenstein, die haar leende van de psycholoog Joseph Jastrow (1900) die haar op zijn beurt weer leende uit *Harper's Weekly* (November 19, 1892, p. 1114). Een maand eerder was zij opgenomen als een humoristisch plaatje in *Fliegende Blätter*, een tijdschrift dat verscheen in München (23 oktober, 1892, p. 147) (Kihlstrom, 2004).



*Afbeelding: Joseph Jastrow die de haas-eend-figuur als eerste in de psychologie introduceerde*



*Afbeelding: Haas-eend-figuur, 1892*

## WHAT'S ON A MAN'S MIND



*Afbeelding: Twee dubbelzinnige figuren*

Dit plaatje en andere dubbelzinnige figuren worden al tijden gebruikt om te laten zien hoezeer de waarneming theoriegeladen is. Dat wil zeggen: een onderzoekgegeven is nooit los te zien van dat waar het een gegeven voor is. Een hypothese laadt, als het ware, de betekenis van de data. Wanneer we nu data beschouwen als onafhankelijk van een theorie (of hypothese), krijgen we meteen de vraag wat wel en niet als data kan of mag worden beschouwd. Net als een ecologische nis een nis is voor een bepaalde soort, met diverse eigenschappen die relevant zijn voor precies die soort, zo zijn data relevant voor een theorie of hypothese, of het zijn geen data!

## *Learning analytics*

Learning analytics lijkt aan een vergelijkbaar euvel. Het doel is veel gegevens vast te leggen over wat leerlingen doen en laten. Uit deze gegevens worden met programma's voor patroonherkenning gedragspatronen geëxtraheerd. Die helpen de docent of de onderwijsinstelling in een vroeg stadium afwijkende en ongewenste gedragspatronen op te sporen. Dat maakt vroegtijdig ingrijpen mogelijk, en het onderwijs kan erdoor worden verbeterd.

There is an increasing interest in using new sources of data for personalizing the learning experience and for performance measurement.

As learners participate in online activities, they leave a clear trail of analytics data that can be mined for insights.

Learning analytics experiments and demonstration projects are currently examining ways to use data for enrichment. Dashboards filter this information so that student progress can be monitored in real time.

As the field of learning analytics matures, the hope is that this information will enable continual improvement of learning outcomes.<sup>4</sup>

(Johnson et al., 2013)

Op zichzelf zou het mooi zijn in een vroeg stadium gedragspatronen te kunnen opsporen die naderend onheil aankondigen. Paradoxaalwijze zijn die patronen beter te vinden wanneer je over veel data beschikt, terwijl de adviezen vooral op individuen zijn gericht en de docent in de klas maar een beperkt aantal studenten heeft. Bovendien moet men gegevens kunnen registreren opdat men met de

software de data kan analyseren. De gegevens worden dan ook uitsluitend geëxtraheerd uit digitale activiteiten.<sup>5</sup> Behalve wat de leerling toch al doet vergt dat ondubbelzinnige invoer van data. Het wordt een hele bureaucratie, dat begrijpt u. Er is toch al niet erg veel 'quality time' in het onderwijs. Maar het gaat me om iets anders. Patronen moeten worden gevisualiseerd, net als de al genoemde smoke



trails, om het algemene karakter van die patronen te doen inzien. Je krijgt mooie plaatjes, en zoals ik al zei: wie weet wat we er allemaal in kunnen en willen zien. Maar verklaren doet het niet. Het werpt wel interessante verklaringsproblemen op. De big data of learning analytics leveren de gecorre-

leerde verbanden; rest de vraag die aanleiding kan zijn tot echt onderzoek: wat is de verklaring ervan? In die zin behoren big data en learning analytics tot de moderne technologie, maar het *is* geen wetenschap, het kan er wel toe leiden.



## IMPACT SCORES

Een ander voorbeeld: impact van wetenschappelijke publicaties en tijdschriften. We passen de data-obsessie ook op onszelf toe, heel compulsief zelfs. De H-index, impact factors, citatiescores, het zijn allemaal maten als indices voor data die worden verzameld, maar waarvan we niet weten wat ze betekenen of hoe ze die berekenen. In New York rekent Thomson Reuter het allemaal uit, maar net als in *Het Proces* weten we niet precies hoe dat werkt (Noorden, 2014). We leren onze studenten over betrouwbaarheid en validiteit en tegenwoordig ook over data-integriteit, maar we weten niet hoe het al die maten op dit

punt vergaat. Het leidt tot vooralsnog beschaafde discussies in het decanenoverleg sociale wetenschappen wanneer moet worden besloten op welke wijze we bibliografische gegevens gaan gebruiken. De psychologen doen het immers anders dan de antropologen. In de faculteiten ontstaan hevige discussies die tot een compromis moeten worden gebracht door de decanen, die vervolgens de visitatiecommissies moeten toepassen, met gevaar voor eigen leven want voor je het weet stormen de antropologen met speren op hen af. Interessant is in dit verband om na te gaan wat al die verzamelde impact-

en citatiedata nu eigenlijk te zeggen hebben. Er is een lange lijst van gesignaleerde interpretaties (Abbott et al., 2010; Bar-Ilan, 2008; Brems, Button, & Munafò, 2013). Was IF oorspronkelijk een hulpmiddel voor universiteitsbibliothecarissen om een keuze te maken voor de tijdschriften die op de rekken moesten komen, in-tussen is het een bij eigen publicaties vermelde score. Het zegt natuurlijk niets over de kwaliteit van een artikel, alleen over een tijdschrift, maar toch kennen onderzoekers er veel waarde aan toe, net als theosofen aan het astraal lichaam. We moeten niet vergeten dat er ook nog eens zo'n zestig 'limitations' en 'biases' zijn vastgesteld die op de onbetrouwbaarheid van de IF lijken te wijzen (Reinstein, Hasselback, Riley, & Sinason, 2011), waaronder het vormen van clubjes

die elkaar citaties bezorgen en tijdschriften in de IF omhoog stuwten. Ik stel dan ook voor om IF voortaan te laten staan voor de Inbreeding Factor, de universitaire of misschien wel universele Inteelt Factor.

Reinstein et al. hebben zestig voorbeelden gevonden van beperkingen en vertekeningen van citatieanalyses (Reinstein et al., 2011). Mede daarvoor is de vraag ook of de H-index wel betrouwbaar is. Judith Bar-Ilan vergeleek drie H-indices van 47 Israëlische auteurs, en zag grote verschillen tussen de H-indices van Web of Science, Scopus en Google Scholar (tabel 1). Dat was volgens de auteur mede te wijten aan de verschillen in de gebruikte citatiescores (tabel 2) (Bar-Ilan, 2008). Sample error of sample bias?

## J. BAR-ILAN: Which h-index? – A comparison of WoS, Scopus and Google Scholar

Table 1. H-index according to WoS, Scopus and Google Scholar for highly cited Israeli researchers for 1996–2006

| Researcher         | Category                         | WoS | Scopus | Google Scholar |
|--------------------|----------------------------------|-----|--------|----------------|
| Alexander, Gideon  | Physics                          | 32  | 30     | 20             |
| Alon, Noga         | Mathematics, Computer Science    | 14  | 17     | 27             |
| Aumann, Robert J.  | Mathematics                      | 8   | 6      | 11             |
| Aurbach, Doron     | Materials Science                | 29  | 29     | 19             |
| Beeri, Catriel     | Computer Science                 | 3   | 3      | 8              |
| Chet, Ilan         | Plant & Animal Science           | 21  | 21     | 20             |
| Ciechanover, Aaron | Biology & Biochemistry           | 33  | 34     | 30             |
| Cohen, Irus R.     | Immunology                       | 29  | 32     | 26             |
| Dagan, Gideon      | Engineering, Ecology/Environment | 13  | 14     | 12             |
| Dekel, Avishai     | Space Sciences                   | 25  | 25     | 24             |
| Dolev, Daniel      | Computer Science                 | 5   | 7      | 18             |
| Duchovni, Ehud     | Physics                          | 32  | 29     | 15             |
| Geiger, Benjamin   | Molecular Biology & Genetics     | 34  | 33     | 31             |
| Gohberg, Israel    | Mathematics                      | 8   | 8      | 11             |
| Goldreich, Oded    | Computer Science                 | 12  | 14     | 32             |
| Harel, David       | Computer Science                 | 9   | 10     | 22             |
| Hershiko, Avram    | Biology & Biochemistry           | 21  | 21     | 21             |
| Hochberg, Yosef    | Mathematics                      | 4   | 4      | 7              |
| Jortner, Joshua    | Chemistry                        | 26  | 26     | 21             |
| Kanner, Joseph     | Agricultural Sciences            | 9   | 9      | 6              |
| Kerem, Batsheva    | Molecular Biology & Genetics     | 20  | 19     | 18             |
| Kotler, Burt P.    | Ecology/Environment              | 10  | 10     | 9              |
| Leviatan, Yehuda   | Computer Science                 | 6   | 5      | 5              |
| Lubotzky, Alex     | Mathematics                      | 5   | 5      | 10             |
| Mechoulam, Raphael | Pharmacology                     | 28  | 29     | 26             |
| Mikenberg, Giora   | Physics                          | 31  | 10     | 4              |
| Moran, Shlomo      | Computer Science                 | 7   | 6      | 11             |
| Netzer, Hagai      | Space Sciences                   | 28  | 28     | 18             |
| Oren, Moshe        | Molecular Biology & Genetics     | 47  | 49     | 38             |
| Peleg, David       | Computer Science                 | 8   | 11     | 21             |
| Piran, Tsvi        | Space Sciences                   | 32  | 32     | 33             |
| Pnueli, Amir       | Computer Science                 | 11  | 8      | 23             |
| Procaccia, Itamar  | Physics                          | 18  | 19     | 18             |
| Shainberg, Isaac   | Ecology/Environment              | 8   | 10     | 9              |
| Shamir, Shlomo     | Computer Science                 | 16  | 17     | 20             |
| Sharir, Micha      | Engineering, Computer Science    | 13  | 15     | 17             |
| Shelah, Saharon    | Mathematics                      | 10  | 10     | 15             |
| Sklan, David       | Agricultural Sciences            | 18  | 20     | 15             |
| Turkel, Eli        | Mathematics                      | 9   | 9      | 11             |
| Wigderson, Avi     | Computer Science                 | 8   | 13     | 23             |

Tabel 1: Vergelijking van drie bronnen voor de H-index van een aantal chemici (Bar-Ilán, 2008)

Table 2. H-index and the number of citations to the h most highly cited publications according to WoS, Scopus and Google Scholar for highly cited Israeli researchers for 1996–2006

| Researcher         | h-WOS | # cits.<br>WOS | h-Scopus | # cits.<br>Scopus | h-GS | # cits.<br>GS |
|--------------------|-------|----------------|----------|-------------------|------|---------------|
| Alexander, Gideon  | 32    | 1,624          | 30       | 1,383             | 20   | 938           |
| Alon, Noga         | 14    | 363            | 17       | 540               | 27   | 1,694         |
| Aumann, Robert J.  | 8     | 101            | 6        | 81                | 11   | 407           |
| Aurbach, Doron     | 29    | 2,062          | 29       | 2,127             | 19   | 1,041         |
| Beerl, Catriel     | 3     | 26             | 3        | 52                | 8    | 347           |
| Chet, Ilan         | 21    | 722            | 21       | 772               | 20   | 804           |
| Ciechanover, Aaron | 33    | 6,084          | 34       | 7,013             | 30   | 5,239         |
| Cohen, Irun R.     | 29    | 2,247          | 32       | 2,738             | 26   | 1,651         |
| Dagan, Gedeon      | 13    | 403            | 14       | 469               | 12   | 469           |
| Dekel, Avishai     | 25    | 1,818          | 25       | 1,801             | 24   | 1,933         |
| Dolev, Daniel      | 5     | 184            | 7        | 221               | 18   | 1,372         |
| Duchovni, Ehud     | 32    | 1,624          | 29       | 1,352             | 15   | 550           |
| Geiger, Benjamin   | 34    | 4,241          | 33       | 3,855             | 31   | 3,439         |
| Golberg, Israel    | 8     | 111            | 8        | 90                | 11   | 243           |
| Goldreich, Oded    | 12    | 444            | 14       | 527               | 32   | 3,569         |
| Harel, David       | 9     | 188            | 10       | 429               | 22   | 3,374         |
| Hershko, Avram     | 21    | 4,228          | 21       | 4,256             | 21   | 3,373         |
| Hochberg, Yosef    | 4     | 143            | 4        | 153               | 7    | 200           |
| Jortner, Joshua    | 26    | 1,935          | 26       | 1,734             | 21   | 1,074         |
| Kanner, Joseph     | 9     | 281            | 9        | 293               | 6    | 153           |
| Kerem, Batsheva    | 20    | 1,108          | 19       | 1,103             | 18   | 992           |
| Kotler, Burt P.    | 10    | 290            | 10       | 224               | 9    | 188           |
| Leviatan, Yehuda   | 6     | 77             | 5        | 62                | 5    | 53            |
| Liabotzky, Alex    | 5     | 47             | 5        | 43                | 10   | 247           |
| Mechoulam, Raphael | 28    | 2,470          | 29       | 2,771             | 26   | 1,943         |
| Mikenberg, Giora   | 31    | 1,545          | 10       | 270               | 4    | 186           |
| Moran, Shlomo      | 7     | 86             | 6        | 88                | 11   | 423           |
| Netzer, Hagai      | 28    | 1,955          | 28       | 1,968             | 18   | 1,157         |
| Oren, Moshe        | 47    | 6,926          | 49       | 7,399             | 38   | 5,460         |
| Peleg, David       | 8     | 113            | 11       | 181               | 21   | 953           |
| Piran, Tsvi        | 32    | 3,200          | 32       | 3,223             | 33   | 3,976         |
| Pnueli, Amir       | 11    | 216            | 8        | 236               | 23   | 2,243         |
| Procaccia, Itamar  | 18    | 582            | 19       | 625               | 18   | 688           |
| Shainberg, Isaac   | 8     | 128            | 10       | 152               | 9    | 154           |
| Shamai, Shlomo     | 16    | 876            | 17       | 1,398             | 20   | 2,232         |
| Sharir, Micha      | 13    | 281            | 15       | 361               | 17   | 646           |
| Shelah, Saharon    | 10    | 139            | 10       | 139               | 15   | 476           |
| Sklan, David       | 18    | 486            | 20       | 538               | 15   | 334           |
| Turkel, Eli        | 9     | 184            | 9        | 221               | 11   | 330           |
| Wigderson, Avi     | 8     | 107            | 13       | 329               | 23   | 1,126         |

Tabel 2: De H-index vertekend als gevolg van onbetrouwbaarheid citatiescores (Bar-Ilan, 2008)

David Spiegelhalter, Winton Professor of the Public Understanding of Risk van Cambridge University, met volgens Google Scholar 64000 citaties, een h-index van 85, en volgens Web of Science 32000 citations, en een h-index van 63, zegt er over:

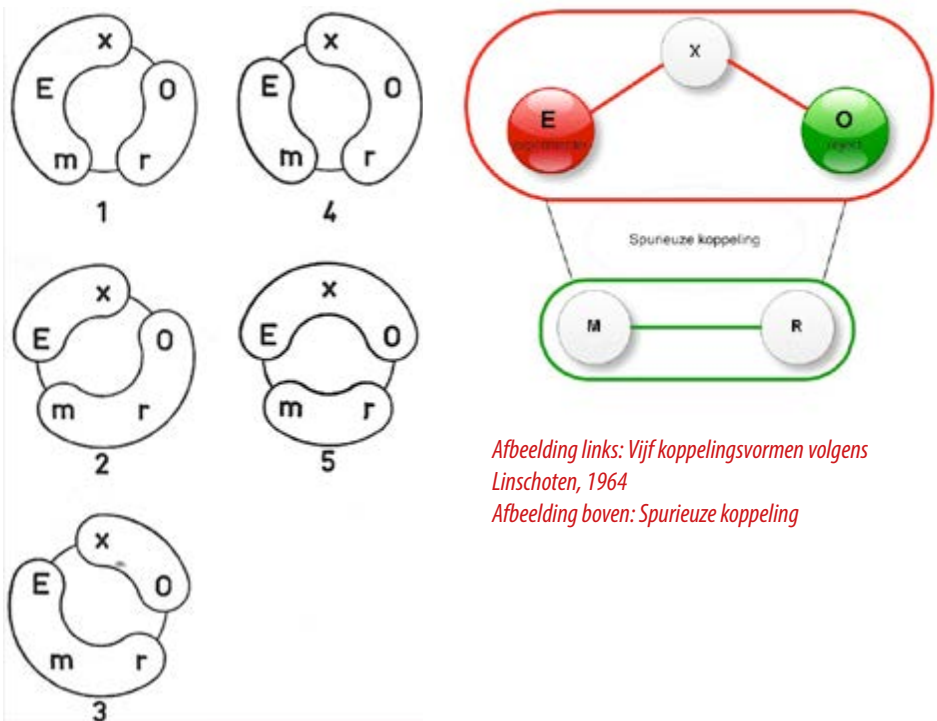
'My own experience shows some of the problems. My highest-cited paper clocks in at over 14,000, and yet it has roughly 150 authors and to be honest I have forgotten what, if anything, I contributed. How would these citations be shared out? Or the WinBUGS User manuals for software: around 5000 citations that do not even appear in WoS.'

Dit roept vooral de vraag op waarom vrijwel iedere onderzoeker erkent dat de gegevens eigenlijk niet deugen maar er compulsief aan meedoet. De wetenschap staat voor een raadsel.

More research needs to be done. Intussen raadpleegt de wetenschappelijke gemeenschap het Web of Science zoals de priesters in de oudheid een orakel raadpleegden.

## DATA SETS

Uit het tot nu gedebiteerde kan niet worden afgeleid dat ik tegen data ben. Integendeel. Ik ben voor precisie, voor betrouwbare en valide metingen in het licht van theorieën. Welke gegevens dat zijn wordt onder andere bepaald door welke gegevens verschil maken voor de hypothesen of theorieën die je toetst. Dat wordt niet alleen bepaald door de gegevens die je überhaupt kunt krijgen. Vooral de onafhankelijkheid van de wensen en voorkeuren van de onderzoeker en de proefpersonen is van belang, maar dat legitimeert niet het idee dat onderzoekgegevens los staan van theorie of hypothese. Linschoten heeft ooit eens in zijn prachtboek *Idolen van de psycholoog* de ‘spurieuze koppeling’ geïntroduceerd als een relatie tussen het systeem van onderzoeker/onderzoek en het systeem van onderzochte (Linschoten, 1964).



*Afbeelding links: Vijf koppelvormen volgens Linschoten, 1964*

*Afbeelding boven: Spurieuze koppeling*

In het vijfde geval ging het erom dat de onderzoeker en zijn hypothese zodanig verweven was met het onderzochte en de proefpersoon dat in feite de onderzoeker de uitslag bepaalt. In de wetenschap willen we dat niet. Maar je ziet dat maar al te vaak, en – sterker nog – het lijkt het wetenschapsbeleid van de overheid. Het idee van de vorige staatssecretaris dat wetenschap meer aan valorisatie moet doen en daarop moet worden afgerekend, de topsectoren van de wetenschap-industrie, bedrijven die onderzoek kopen (we noemen dat derde geldstroom), universiteiten die als ondernemers in online kennis MOOCs verkopen, het lijkt in mijn ogen verdacht veel op spurieuze koppelingen. Data als los spul gaan zien helpt daar niet bij. Voor je het weet staat er een miljardair op de stoep die nadat hij zijn voetbalclub heeft verkocht nu besluit in data te gaan.

Ik betoog dat data zich altijd verhouden tot een theorie, en dat een theorie een metafysica impliceert of er door wordt geïmpliceerd. In de psychologie zijn we gaandeweg geobsedeerd geraakt door data, niet door de vraag hoe we meer kunnen begrijpen van de wereld. Michael Billig wees er al op hoe slecht we eigenlijk schrijven, en vooral van werkwoorden zelfstandige naamwoorden maken en daar-

door anonimiseren (Billig, 2013). *Wij* verdinglijken alles, maar we noemen dat verdinglijking. Mensen vervreemden van elkaar maar het heet dat zij een vervreemdingsproces ondergaan. Meer mensen plegen misdrijven, maar we spreken van stijgende criminaliteit. Willen wordt opzet, hevig willen wordt strong efficacy, eten wordt eetgedag vertonen, bewegen wordt beweeggedrag vertonen, en nog even en we hebben theorieën die ‘gedraaggedrag’ verklaren. We ontbladeren als het ware de mensen waar het in de sociale wetenschappen over gaat. Ik zie daarin een verband met het obsessief compulsief gedrag rond data, de fascinatie voor data als losse te verzamelen dingen. Jantje zag eens data hangen, o! als eieren zo groot. ’t Scheen, dat Jantje wou gaan plukken, Schoon de methodoloog ’t hem verbood, maar daaraan had hij een broertje dood (vrij naar Hieronymus van Alphen).

De obsessie heeft vreemde consequenties. Een ervan behelst datasets. Op een gegeven moment verzamelen we data en brengen die onder in een grote tabel. We hebben dan een dataset; een mooie dataset, heet het zelfs. Een mandje met data, en dat mandje kun je los van een idee of theorie zo maar statistisch verwerken. SPSS – de dominante statisti-



sche gereedschapskist in de sociale wetenschappen – poetst alle resterende schroom weg: alles mag wat SPSS biedt, is het devies. Je geeft Jantje, inmiddels een master student, een dataset, en die hoeft er alleen maar wat significante verbanden uit te plukken, door simpelweg de receptuur van SPSS te gehoorzamen en wat toetsen in te drukken. Hoe meer hoe mooier. Wat de data precies betekenen, wat de variabelen allemaal voorstellen, is niet zo interessant, als de bijhorende syntaxfile maar okay is en goed wordt bewaard, want die hebben we straks hard nodig voor de dataopslag in DANS Easy, of wat voor repository ook. Alles wordt beschreven, zelfs de metadata. Maar wat het zegt?

Ik voorspel dat het doel van DANS Easy, namelijk beter toegankelijk maken voor controle op de integriteit van data, niet zal worden bereikt. Pas wanneer we een vermoeden van boos opzet tot fraude hebben, zullen we die dataset raadplegen en alles narekenen. Het andere beoogde doel, replicatie mogelijk maken, is vrijwel zinloos. Exacte herhaling – anders dan ter controle van de integriteit van de onderzoeker – is nu juist wat replicatie *niet* moet zijn. Als theorieën *algemene* beweringen zijn dan willen we juist replicatie onder *verschillende*

omstandigheden, met andere, nieuwe proefpersonen, en met andere responsopties. Met ‘andere’ bedoel ik zowel ‘vers’, als ‘van een andere verschijningsvorm’. DANS Easy zal echter vooral een Efteling worden voor ‘datafitters’. Maar zoals Richard Feynman – de beroemde natuurkundige – al eens in 1998 liet weten, ‘[datafitting]...doesn’t count...[b]ecause it doesn’t make sense to calculate [probabilities] after the event. [To test a] hypotheses ... [you] cannot do it from the same data that [gave] the clue.’ (Feynman, 1998, p. 34).

Als data echter theoriegeladen zijn, en dat zijn ze, dan bestaan losse data in datasets niet. Niet alleen zijn ze verworven vanuit het perspectief van een te toetsen theorie, maar ook vergen ze meetinstrumenten met bepaalde eigenschappen. Theorieën zijn geen meetinstrumenten. Een theorie geeft aan dat voor alle  $x$  geldt dat als een bepaalde conditie aanwezig is, er bepaalde consequenties zullen zijn te vinden. Om die condities op hun aanwezigheid te controleren, en om de aanwezigheid van de consequenties vast te stellen, moeten we meetinstrumenten hebben die dat kunnen. Die zijn onafhankelijk van de ‘waarheid’ van de theorie waar we het over hebben, maar ze kunnen wel een antwoord helpen vinden op

de toetsvraag. Dat is een voordeel want als er iets mis gaat – als we de consequenties niet aantreffen – kan dat liggen aan het meetinstrument, of aan de theorie, of beide. Dit is realistisch, overigens, en zelfs rationeel. Dat betekent dat we met twee zaken rekening moeten houden in de data discussie.

Eerst Bayes: stel we hebben heel fatsoenlijk een hypothese afgeleid uit een redelijk goed geëxpliciteerde theorie. De hypothese  $H$  is gesteld in waarschijnlijkheidstermen en die willen we toetsen zodat we indirect iets meer kunnen zeggen over de vraag waarom we voorlopig op theorie  $T$  kunnen vertrouwen – ook weer in termen van waarschijnlijkheden. We bekijken of we voor  $H$  misschien een meetinstrument  $M$  hebben dat observaties  $e$  op een bepaald niveau oplevert die over de waarschijnlijkheid van  $H$  iets zeggen, die ertoe doen. De crux zit hem in dat ‘ertoe doen’. Wat sociale wetenschappers met het  $H_0$  gedoe doen, is, strikt genomen, aangeven wat de waarschijnlijkheid van de gevonden data is, gegeven  $H_0$ . Met andere woorden: de  $p$ -waarde geeft aan hoe waarschijnlijk het is dat we deze gegevens vinden als de nulhypothese waar is. Zo althans is de juiste interpretatie ook al weten weinig gebruikers van het  $H_0$ -ritueel dat. In veel gevallen worden echter

de waarschijnlijkheid en gegevenheid omgedraaid: wat is de waarschijnlijkheid van de nulhypothese gegeven de onderzoeksgegevens? We vergeeten daardoor hoezeer de waarschijnlijkheid om deze gegevens te vinden afhankelijk is van de waarheid van de hypothese zelf en van de kans dat de meetinstrumenten de gegevens juist meten, niet van de nulhypothese. De  $p$ -waarde zegt helemaal niets over wat de eigenlijke hypothese ( $H_1$ ) waard is. Als de hypothese verandert, verandert de waarschijnlijkheid van  $H_0$  niet, gegeven de data tenzij de data veranderen; als de meetinstrumenten veranderen in hun ‘belang’ voor de hypothese, pas dan verandert er iets aan de waarde van de gegevens. Alleen al daarom moeten we data, hypothese, meetinstrumenten en de theorie waar zij uit komen in hun samenhang als een geheel zien. En dat significantieniveau vergeten. Het andere aspect heeft te maken met wat Watkins er over zegt in *Science and scepticism* (J.W.N. Watkins, 1984). Hij onderscheidt daar de vijf niveaus in de uitspraken (beweringen) die we in de wetenschap kunnen doen. Ik herhaal hier de eerder gepresenteerde tabel met de voorbeelden uit de psychologie, en heb zijn voorbeelden uit de fysica in cursief weergegeven.

| Level | Definition                                                                                                                                                                                                    | Falsified by         | Supported by                                                                                              | Verified by          | Falsifying                                                                                                               | Supporting                                                                                                               | Verifying            |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------------------------------------|----------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|----------------------|
| L-0   | perceptual reports of a first-person, here-and-now type (e.g. 'I feel a sharp pain in my left lower wisdom tooth') (e.g. 'In my visual field there is now a silvery crescent against a dark blue background') | --                   | L-1 statements about behavior or brain events supposed by L-2 statements to correlate with L-0 statements | Other L-0 statements | L-1 statements<br>L-2 statements about behavior or brain events supposed to correlate L-1 statements with L-0 statements | L-1 statements<br>L-2 statements about behavior or brain events supposed to correlate L-1 statements with L-0 statements | Other L-0 statements |
| L-1   | singular statements about observable things or events (e.g. 'The dentist is using his drill on the left lower wisdom tooth') (e.g. 'There is a new moon tonight')                                             | Other L-1 statements | L-0 statements<br>L-2 statements                                                                          | Other L-1 statements | L-2 statements                                                                                                           | L-2 statements                                                                                                           | --                   |
| L-2   | empirical generalizations about regularities in observable things and events (e.g. 'A suggestion of forthcoming pain is followed by experienced pain') (e.g. 'A new moon is followed by spring tide')         | L-1 statements       | L-1 statements<br>L-3 statements<br>L-4 statements                                                        | --                   | L-3 statements<br>L-4 statements                                                                                         | L-1 statements<br>L-3 statements<br>L-4 statements                                                                       | --                   |

| Level | Definition                                                                                                                                                                                           | Falsified by   | Supported by                     | Verified by | Falsifying     | Supporting                 | Verifying                   |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------------------|-------------|----------------|----------------------------|-----------------------------|
| L-3   | exact experimental laws concerning measurable magnitudes (e.g. Fechner's Law of relation between magnitude of sensation and physical magnitude) (e.g. <i>The gas law of Charles and Gay-Lussac</i> ) | L-2 statements | L-2 statements                   | --          | L-4 statements | L-4 statements             |                             |
| L-4   | scientific theories that are universal and exact, and postulate unobservable entities (e.g. the Gate Theory of pain by Melzack & Wall). (e.g. <i>The Faraday-Maxwell theory of fields of force</i> ) | L-3 statements | L-2 statements<br>L-3 statements | --          | --             | It's own metaphysical core | It's own meta-physical core |

*Tabel herzien: Niveaus van beweringen en hun relatie tot bewijsmateriaal in de psychologie. Falsification betekent weerlegging in geval van tegenspraak; verification betekent accepteren in het geval van volledige en voltallige overeenstemming; support betekent corroboratie: tentatieve aanvaarding wanneer de weerlegging (nog steeds) faalt. Naar Watkins (1984). *Science and scepticism*. London: Hutchinson.*

Laatst ontstond in onze onderzoekcommissie een interessante discussie. Ik kwam daar met de boude bewering dat data geen loslopende en robuuste feiten zijn maar samenhangen met een inhoudelijke theorie, en dat je geen 'mooie datasets' hebt die voor andere dan de oorspronkelijke doeleinden, en dan nog slechts eenmalig, bruikbaar zijn. Waarop een hoog gewaardeerde en zeergeleerde collega het voorbeeld gaf van DNA: tien jaar geleden was lichamenlijk materiaal verzameld en bewaard – in de koelkast neem ik aan. Tien jaar geleden kon men daarmee het DNA – meettechnisch gesproken – tot op een bepaald niveau van precisie analyseren. Een jaar geleden was de precisie toegenomen en konden veel verfijnder meetgegevens worden ontlokt aan de stalen speeksel.

Kijkend naar deze tabel is wel iets te zeggen over het verschil tussen het gebruik van de verfijnder DNA meetinstrumenten op 'oud' speeksel, en het hergebruik van vragenlijstgegevens. De eerste zijn stalen waarop verder niemand invloed heeft gehad, ook de leverancier van het speeksel niet. Als er al invloed uitgaat van de mentale of lichamelijke toestand van de producent, dan nog mogen we ervan uitgaan dat er geen bewuste invloed mogelijk is op de DNA-eigenschappen van/in het speeksel. Een speekselbank is echter geen dataset; het is de eerste, technisch verouderde analyse die we de (oude) dataset noemen. De tweede analyse op het 'oude' speeksel is een nieuwe dataset. Hoe zit dat bij vragenlijsten? Vragenlijsten zijn eigenlijk een massale vorm van introspectie. We vragen niets aan het speeksel of aan het DNA, we meten dat; maar vragenlijsten vragen naar 'perceptual reports of a first-person, here-and-how-type'. Participanten interpreteren elk woord op hun eigen wijze. We vragen in de ogen van de onderzoeker heel specifiek, maar voor de participanten is er sprake van een heel andere specificiteit. Het resultaat is in feite een hele reeks specifieke interpretaties van de vraag die de onderzoeker stelde. Daar kun je wel gemiddelden op bepalen, maar dat zegt net zo veel als het portret van 'de' crimineel dat uit de physiognomic synthesizer van Galton rolde. Nog afgezien van het feit dat vragen ook dwingen tot het bedenken van een antwoord. Op de vraag: 'Hoe hard slaat u uw partner wanneer u hem/haar slaat?' dwing je mensen tot een te middelen score met een Likert-schaal van 'heel hard' naar 'niet'. Een antwoord, welk ook, zou in

het gemiddelde verdrinken. Met andere woorden wat en hoe we het ook vragen, er zijn twee theorieën impliciet in de vraag, en dat is een theorie van de onderzoeker, namelijk dat mensen hun partner soms slaan, of misschien wel dat sommige mensen hun partner altijd slaan. Dat is de vraag van de onderzoeker; maar ook de proefpersoon heeft zijn theorie: Wat wil die onderzoeker eigenlijk weten? Bij DNA heb je dat niet.

We zitten hier al op een veel hoger niveau van uitspraken dan met dat speeksel dat ten grondslag ligt aan de meetuitkomst. De laborant levert de rapportage van de analyse in het lab nadat hij het resultaat van een onafhankelijke meting afleest van de print-out van het apparaat dat een speekselstaaltje analyseert. Het apparaat realiseert een stabiele, ingebakken meettheorie op basis van frequencying, zoals dat heet. Degene die de vragenlijst beantwoordt is echter geen apparaat met een print-out in de vorm van een ingevulde vragenlijst, maar iemand met eigen ideeën. Hij of zij heeft geen meettheorie, en als er al ideeën zijn dan zijn ze weinig stabiel, naar ik vrees. Zijn antwoorden zijn eigenlijk op niveau L-2, of L-1, of er sterk door gekleurd. Slechts een enkele keer, bijvoorbeeld in perceptie-onderzoek naar binoculair diepte-zien, zijn ze van niveau L-0. Zelfs hier is het van belang theorieën te ontwikkelen en verfijnen die laten zien welke reacties 'reflectief' en welke 'reflexief' zijn.

Kijken we naar de volgende kolommen dan zien we dat L-2-uitspraken voor falsificatie afhankelijk zijn van L-1-uitspraken, en voor hun support van L-1-uitspraken of 'hogere' uitspraken L-3 en L-4. Falsificatie zou van andere L-0-uitspraken kunnen komen. Dat kan ik me bij het lezen van de print-out van een DNA-analyserend apparaat heel goed voorstellen. Bij de rapportage over kiespijn niet: dat voel je alleen zelf, er is geen falsificatie op dat niveau denkbaar. Precies daar zie je ook waarom Ronald Melzack zijn theorie nodig had (Melzack & Wall, 1967). Melzack werkte met patiënten met fantoompijn waarvoor indertijd geen medische verklaring was. Het goede, de 'support', komt in de psychologie dus van boven, van theorieën die L-0-uitspraken een plaats geven. 'Support' betekent corroboratie, dat wil zeggen een niet-complete bevestiging, steun voor een uitspraak, een uitspraak op het ene niveau die consistent is met een bewering op een ander niveau. L-2 uitspraken worden niet (definitief en volledig) geverifieerd, ze falsifiëren wel L-3 en L-4-uitspraken, en bieden steun

aan L-1, L-3 en L-4-uitspraken. Ze verifiëren niets. Dat is nog even afgezien van de mate waarin de beweringen op alle niveaus, ook op niveau L-0, zeker (kunnen) zijn. Ze lijken onweerlegbaar waar voor de waarnemer van de maan of de patiënt van de tandarts, maar we weten uit de psychologie ook hoe bedrieglijk de waarneming kan zijn. Daartoe moeten we onafhankelijke metingen, in op selectie gerichte experimenten, inrichten. We hebben theorieën nodig om vast te stellen of de lichamelijke signalen consistent zijn met pijn of angst, of welke emotie ook. Ongeacht de uitslag blijft de ervaren pijn niet de pijn die de onderzoeker geleden heeft, maar een die de pijn niet kan zijn<sup>6</sup>.

## CONCLUSIES

**Kluger Hans – het paard van die naam – wist er alles van. Hans kon volgens zijn baas Wilhelm von Osten tellen en rekenen, zelfs worteltrekken en machtsverheffen.**

Tot Oskar Pfungst – student van psycholoog Carl Stumpf – in 1911 ontdekte dat Hans slechts reageerde op onbewust gegeven signalen van degene die hem het rekenprobleem voorlegde. We moeten er dus als psycholoog op verdacht zijn de proef-

persoon niet de clues te geven die we willen horen, we moeten er als proefpersoon op verdacht zijn niet onszelf de clues in te fluisteren als we denken onze eigen psycholoog te kunnen zijn, zelfs als we een vragenlijst van een ander krijgen voorgelegd.





*Afbeelding: Vlnr Wilhelm Von Osten, Kluger Hans, Oskar Pfungst*

De recente hype op controle en integer databeheer en dergelijke is begrijpelijk. Er is een stapeltje van voorbeelden van hoe wetenschappers, in biologie, in geneeskunde, in natuurkunde, in sociale psychologie, het vertrouwen hebben beschaamd. Er is een vertrouwenscrisis in de wetenschap. Minder vertrouwen in onderzoekers, nauwelijks vertrouwen in de dataproducenten, alleen in papieren maatregelen en opgeslagen data. Dus nog een schepje meer controle. Er was vertrouwen, nu fluctueert het vertrouwen in het ritme van het gezegde: vertrouwen komt te voet en gaat te paard. De werkelijk relevante vraag is: waarom fluctueert het? Ik denk: omdat we denken dat we met veel en nog meer data ook meer overzicht krijgen, zoals Reuters met zijn IF. Daardoor missen we historie en theorie die aangeeft wat wel en niet onzin is, wat wel en niet data zijn die verschil kunnen maken, hoe data zich wel en niet tot theorieën kunnen verhouden. Wetenschappers zijn orakels geworden. *Nature en Science* bieden entertainment, big data wordt big business, impact scores zijn inteelfactoren. Maar levert dat kennis op? Het overzicht kan niet komen van nog meer publicaties. Er is al veel te veel om te kunnen overzien. Het overzicht kan ook niet komen van

nog meer controle van data. Niemand heeft immers de tijd voor controle, en tijdschriften publiceren replicaties niet, tenzij er sensationele redenen voor zijn. Het overzicht kan evenmin komen van nog meer regels: hoe meer regels hoe meer overtredingen, weten we uit de criminologie: wie een regel stelt geeft tegelijk aan waar de gaten zitten waarlangs je aan de regel ontsnapt. Het overzicht kan alleen nog komen van een gedegen analyse op de implicaties van de vele theorieën en hypothesen die er al zijn. Overzicht krijg je door te abstraheren, niet door onderin te specificeren, niet door te extraheren, maar door na te denken. Daarom heb ik er nog eens op gewezen hoe non-empirisch, dat wil zeggen, metafysisch ons werk nu eenmaal is. De *logische implicatie* van wat we met theorieën en hypothesen en toetsende observatie-uitspraken doen, is geen mystiek maar biedt het cement dat meer inzicht in het geheel oplevert. Want, zoals Bob Rieber al eens zei: *"If you don't know where you're going, you're always gonna get there"*. Maar dat is geen wetenschap. Data zijn relevant voor een theorie of hypothese, of het zijn geen data. Het Godot- en Kafka-effect ontstaat pas bij ontbreken van een theorie, wanneer 'losse' data verkeerde suggesties wekken.



*Afbeelding: The end is at hand (2010)*



## DANK

Dames en heren, het eind is in zicht. Ik zet binnenkort mijn weetmuts af, maar wil nu een dankwoord uitspreken. Ik heb veel te danken aan veel mensen in de OU. Van de zeshonderd zou ik er zeker driehonderd willen noemen, maar dat kan niet. Ik moet me beperken. Ik dank om te beginnen het College van bestuur en het Universitair Management Team voor de prettige samenwerking. Ik heb het idee dat we er langzamerhand een stevige, echte universiteit van hebben gemaakt: klein van schaal, groot van daden, subtieler van onderwijs en onderzoek, en zakelijker wanneer we het niet eens zijn. Sander en Anja, Ronald, Koen en Huub, Evert en Gerard: veel succes en behoudt de vaart!

Sander en Anja moet ik trouwens nog even iets bekennen nu het toch geen kwaad meer kan: ik heb een keer in een onbewaakt ogenblik jullie namen in een frase uitgesproken als Lasagna; dat was geen boos opzet en al helemaal geen vorm van noodzakelijkheid-weten. Slip of the tongue, en of het freudiaans is? Nee, Freud is dood.

Ik dank ook de leden van stuurgroepen, raden, commissies en colleges (IPO, Maestro Kompas, Raad van decanen, AIBB, ad-hoc commissie dubbel-loopse belasting, Commissie De

Paardensprong, College van promoties, College voor beroep voor de examens, College van feesten en partijen,... enz.) voor wat ik ervan heb geleerd.

Ik roem het landelijke Decanenoverleg Sociale Wetenschappen en de Kamer Psychologie als belangrijke organen voor samenwerking tussen de faculteiten en opleidingen.

De laatste jaren kreeg het overleg een overwegend collegiaal karakter, wat jullie zal helpen komende stormen te doorstaan.

Ik wil in elk geval alle mensen van de voormalige faculteit Psychologie noemen. Het was een groot plezier om met jullie samen te werken, en voor jullie te werken. Het was een genoegen te zien hoe de productie, zowel in onderwijs als onderzoek, voor een financieel zo achtergestelde faculteit zo enorm en van hoge kwaliteit was. Ook de andere leden van de sinds een jaar grotere faculteit Psychologie en onderwijswetenschappen dank ik zeer voor de aanvankelijk aarzelende later werkelijk beoefende samenwerking. Ik was vast enige tijd de vijand die hielp de oude ploegen bij elkaar te houden, maar ik denk dat het nu wel een beetje over is, en ik hoop dat jullie met de nieuwe decaan tot de aloude grote en grootse prestaties komen.

In het bijzonder dank ik Harrie Adriaens en Jos van den Broek die in verschillende fasen hebben bijgedragen aan de financiële en projectmatige aspecten van de facultaire bedrijfsvoering. Maar natuurlijk ook omdat we veel hebben gelachen en over resp. muziek en de bellettrie konden praten. Ik heb veel titels meegereggen. Die ga ik nu beluisteren en lezen.

Oscar, Marie-Louise en Richard wil ik ook noemen. Jullie maakten in verschillende situaties de personele zaken van de faculteit robuust, en werkten ieder op hun eigen manier betrouwbaar en vertrouwbaar, als ik dat laatste woord zo mag gebruiken. De onderwijsondersteuners van psychologie, Nanda, Marjan, Susan, Ellis: jullie hebben altijd hard gewerkt, en veel bijgedragen aan het succes van de faculteit. De medewerkers van AV-Media, Sandra en Annette, en de redacteur, Jan, die steeds weer met prachtige uitgaven kwamen wil ik er ook in betrekken voor wat zij voor onze facultaire presentatie deden. Het was prachtig en dat is het nog steeds.

Mijn naaste collega's Karen, Jacques, Lilian, Marie-José, het was de laatste jaren plezierig om met jullie te werken. Ik hoop dat ik jullie voldoende professionele ruimte heb gelaten, maar ik dank jullie voor de inzet om er een mooie faculteit van te maken. Dat laatste geldt ook voor de wat later aangehaakte collega's van het Welten-instituut. Jullie namen de professionele ruimte, zelfs in de spelling, maar altijd met de hoop op goede samenwerking. Saskia, Paul, Marcus,

Rob, Peter, Wim, Theo, Els, Maarten, Joseph, voorwaar ik zeg u: er komen goede tijden.

Wie ik zeker niet wil vergeten is mijn voorganger, Jasper von Grumbkow, die een gespreid bed voor mij had bereid toen ik begon. Hij was een jaar of negen de meest naaste collega die ik ooit kende, en een wijs raadgever. Heel veel dank aan Monique Verschuren. Monique is een gouden secretaresse: altijd opgeruimd, zowel qua humeur als qua stofzuigen, en meer dan wie ook in grote delen van de faculteit de sociale spil waar het werkzame deel van het leven om draait. Als er weer iets was te organiseren – hoe moeilijk ook, zoals een bijeenkomst van een Benoemingsadvies-Commissie, een oratie, een af-

scheidssymposium, een bijeenkomst van commissie zus of college zo, het lukte haar altijd. Ik hoop van harte dat het met en voor de nieuwe decaan op dit niveau kan blijven.

Ik dank ook nogmaals de vier sprekers op het symposium: Edward, Hank, Ep en jarige Paul. Het was ontzettend leuk en leerzaam. Precies wat ik wilde: jullie variatie op enkele thema's uit de psychologie die in de faculteit wordt beoefend, zonder de psychologie over boord te zetten.

Tot slot Emma. Ze heeft geduldig gewacht op dit moment, en mij al die jaren gesteund met scherpe inzichten in de mensen die zij van een afstand soms beter doorzag dan ik van nabij.



**Dames en heren,  
the end is at hand.  
Ik heb nu echt gezegd.**

## BIBLIOGRAFIE

- Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Noorden, R. v. (2010). Metrics: Do metrics matter? *Nature*, 465, 860-862. doi: 10.1038/465860a
- Bar-Ilan, J. (2008). Which h-index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257-271.
- Beckett, S. (1954). *Waiting for Godot*. London: Faber & Faber.
- Billig, M. (2013). *Learn to write badly; How to succeed in the social sciences*. Cambridge: Cambridge University Press.
- Brembs, B., Button, K., & Munafò, M. (2013). Deep impact: unintended consequences of journal rank. *Frontiers in Human Neuroscience*, 7. doi: 10.3389/fnhum.2013.00291
- Brown, H. I. (1977). *Perception, Theory and commitment; The new philosophy of science*. Chicago: Precedent Publishing.
- Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, 50(3), 255-272.
- Brunswik, E. (1947). *Perception and the representative design of psychological experiments*. Berkeley: University of Los Angeles Press.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62(3), 193-217.
- Chalmers, A. (1999). *Wat heet wetenschaps (What is the thing called science)*. Amsterdam: Boom.
- Feynman, R. P. (1998). *The meaning of it all: thoughts of a citizen scientist*. Boston, MA: Da Capo Press.
- Galton, F. (1878). Composite portraits. *Nature*(May 23), 97-100.
- Hacohen, M. H. (2000). *Karl Popper; The formative years 1902-1945*. Cambridge: Cambridge University Press.
- Harford, T. (2014). Big data: are we making a big mistake. *Significance*, 11(5), 14-19.
- Hempel, C. G. (1962). Explanation in science and history. In R. G. Colodny (Ed.), *Frontiers of science and philosophy* (pp. 7-33). London: Alan & Unwin.
- Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013). *NMC Horizon Report: 2013 Higher Education Edition*. Austin, Texas: The New Media Consortium.
- Kafka, F. (2014). *Het Proces* (e-book edition ed.). Amsterdam: Atheneum-Polak & Van Genneep.
- Kihlstrom, J. F. (2004). Joseph Jastrow and his duck—or is it a rabbit. *of Trends in the Cognitive Sciences (TICS)*, last visited, 31(07), 2007.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-196). Cambridge: Cambridge University Press.

- Linschoten, J. (1964). *Idolen van de psycholoog* (2nd ed.). Utrecht: Bijleveld.
- Lovejoy, A. O. (1964). *The great chain of being; A study of the history of ideas; The William James Lectures 1933*. Cambridge MA: Harvard University Press.
- Melzack, R., & Wall, P. D. (1967). Pain mechanisms: a new theory. *Survey of Anesthesiology*, 11(2), 89-90.
- Miller, D. (1994). *Critical Rationalism: a restatement and defence*. Chicago: Open Court.
- Musgrave, A. (1993). *Common sense, science and scepticism; A historical introduction to the theory of knowledge*. Cambridge: Cambridge University Press.
- Noorden, R. v. (2014). Transparency promised for vilified impact factor; Thomson Reuters vows to be clearer about how science's most misused metric is calculated. *Nature News*(29 July 2014). doi: doi:10.1038/nature.2014.15642
- Panhuysen, G. (1990). *Het ei van Freud; Over de bijdrage van medische heuristieken aan geboorte en groei van de psychoanalyse*. (Dissertatie), Utrecht University.
- Popper, K. R. (1934). *Logik der Forschung* (Vol. 9). Wien: Julius Springer.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Popper, K. R. (1963). *Conjectures and refutations; The growth of scientific knowledge*. London: Routledge and Kegan Paul.
- Popper, K. R. (1972). *Objective knowledge; An evolutionary approach*. London: Oxford University Press.
- Popper, K. R. (1979). *Die beiden Grundprobleme der Erkenntnistheorie (The two fundamental problems of the theory of knowledge)*. Tübingen: Mohr.
- Reinstein, A., Hasselback, J. R., Riley, M. E., & Sinason, D. H. (2011). Pitfalls of Using Citation Indices for Making Academic Accounting Promotion, Tenure, Teaching Load, and Merit Pay Decisions. *Issues in Accounting Education*, 26(1), 99-131.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton NJ: Princeton University Press.
- Simon, H. A. (1982). *Models of bounded rationality*. Cambridge, MA: MIT Press.
- Strijards, G. A. M. (1983). *Facetten van dwaling in het strafrecht*. Arnhem: Gouda Quint.
- SURF. (2013). Learning analytics in het hoger onderwijs: mogelijkheden en aandachtspunten. Utrecht: SURF.
- Tillyard, E. M. W. (1943 / 1986). *The Elisabethan world picture*. Harmondsworth: Pelican Books.
- Van Hezewijk, R., & De Vries, R. (1989). Denken in beelden, denken in taal? Het debat over mentale verbeelding. In C. Brown, P. Hagoort & T. Meijering (Eds.), *Vensters op de geest* (pp. 82-115). Utrecht: Grafiet.
- VentureBeat. (2014). Data is beautiful is a hidden gem for gorgeous data visualizations. from <http://venturebeat.com/2014/08/09/data-is-beautiful-is-a-hidden-gem-for-gorgeous-data-visualizations/>
- Watkins, J. W. N. (1957). Between analytic and empirical. *Philosophy*, 32, 112-131.
- Watkins, J. W. N. (1958). Confirmable and influential metaphysics. *Mind*, 67, 345-365.



- Watkins, J. W. N. (1960). When Are Statements Empirical? *The British Journal for the Philosophy of Science*, 10(40), 287-308.
- Watkins, J. W. N. (1970). Imperfect rationality. *Explanation in the behavioural sciences*, 167-217.
- Watkins, J. W. N. (1975). Metaphysics and the advancement of science. *British Journal for the Philosophy of Science*, 26, 91-121.
- Watkins, J. W. N. (1978). Minimal presuppositions and maximal metaphysics. *Mind*, 87, 195-209.
- Watkins, J. W. N. (1984). *Science and scepticism*. London: Hutchinson.
- Yates, F. A. (1964). *Giordano Bruno and the hermetic tradition*. London: Routledge and Keegan Paul.



## NOTEN

- 1 In latere artikelen heeft Watkins de aanvankelijke ideeën aangescherpt. In zijn vroege artikelen staat metafysica nog half buiten de wetenschap, in latere artikelen laat hij zien dat er met enig logisch handwerk – quasi-Ramseyfication\* genoemd – uit een theorie de metafysische kern kan worden geëxtraheerd. He komt er op neer dat de feitelijke inhoud van de theorie verder reikt dan de empirische inhoud. De empirische inhoud is het geheel van beweringen die toetsbaar – falsifieerbaar – zijn: de universele uitspraken en de basisuitspraken of observatie-uitspraken. De inhoud die verder strekt is bijvoorbeeld het geheel van existentiële beweringen die door de theorie worden geïmpliceerd. Die zijn niet falsifieerbaar – wel verifieerbaar. Deze aanpak suggereert misschien te veel dat alleen vanuit een theorie een metafysica emergeert. Maar zoals gezegd komt het omgekeerde  $[M \rightarrow T \rightarrow H]$  ook voor. Niet in het minst in ons soort wetenschappen.

Ramseyfication is afgeleid van een procedure die door F.P. Ramsey is ontwikkeld met als doel een theorie te ontdoen van alle beweringen die geen empirische inhoud hebben. Watkins draait het dus om, om te laten zien dat een theorie meer heeft dan alleen empirische inhoud: je hebt immers eerst de volledige theorie nodig alvorens je die kunt ontdoen van de niet-empirische inhoud. Maar die is geringer van inhoud. Vandaar 'quasi-Ramseyfication'.



- 2 Wie een amulet bij zich draagt gelooft in termen van dat mystieke universum; astrologie en numerologie passen in dat zelfde wereldbeeld, oud-psychologische termen als 'zwartgallig', 'droogkloot', 'etterbak' stammen uit het mensbeeld dat aansloot bij dat neo-platoonse wereldbeeld.



- 3 Tegenwoordig noemen we dat de *prime*. De prime kan nu met ongekend precieze timing voorafgaand aan een andere prikkel worden aangeboden, en wel zo snel dat er geen bewuste waarneming van de prime is, maar er wel een effect van kan worden vastgesteld. Maar het idee is al meer dan een eeuw oud.



- 4 Opmerkelijk genoeg bevat dit rapport geen enkel plaatje behalve enkele foto's van leerlingen op omslag en binnenomslag.



- 5 Surf begrijpt het probleem gelukkig wel: 'Learning analytics signaleert, maar verklaart niet' Belangrijk om te realiseren is dat de data binnen learning analytics informatie geven die voorheen niet beschikbaar was voor de student, docent en instelling. De data vertellen echter niet alles over het leergedrag van studenten of over de kwaliteit van de gebruikte didactische materialen. Learning analytics signaleert bijvoorbeeld dat cursusmateriaal weinig gedownload wordt, maar geeft (nog) geen antwoord op waarom dat zo is. Ook hebben 'onzichtbare' leeractiviteiten buiten de digitale leeromgeving invloed op studieprestaties en onderwijskwaliteit. Daarnaast is er een aantal belangrijke aandachtspunten bij de ontwikkeling van toekomstige initiatieven op het gebied van learning analytics (SURF, 2013, p. 3).



- 6 Fernando Pessoa, *Autopsychografie*, vertaling August Willems, 1978

De dichter wendt slechts voor.  
Hij veinst zo door en door  
Dat hij zelfs voorwendt pijn te zijn  
Zijn werkelijk gevoelde pijn.

En zij die lezen wat hij schreef,  
Voelen in de gelezen pijn  
Niet de twee die hij geleden heeft,  
Maar een die de hunne niet kan zijn.

En zo rijdt in het rond,  
Tot vermaak van onze rede,  
Die opwindtrein, in dichtermond  
Ook wel 'het hart' geheten.



COLOFON

*Auteur*

prof. dr. René van Hezewijk

*Lay-out*

Sandra Daems en Annette Bouwels

*Redactie*

Jan Hendriks

© 2015 Open Universiteit, Heerlen

Behoudens uitzonderingen door de Wet gesteld mag zonder schriftelijke toestemming van de rechthebbende(n) op het auteursrecht niets uit deze uitgave worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of anderszins, hetgeen ook van toepassing is op de gehele of gedeeltelijke bewerking.

Save exceptions stated by law no part of this publication may be reproduced in any form, by print, photoprint, microfilm or other means, included a complete or partial transcription, without the prior written permission of the publisher.

ISBN 978 94 918 2578 1