



Universiteit
Leiden
The Netherlands

Explanation and determination

Gijsbers, V.A.

Citation

Gijsbers, V. A. (2011, August 28). *Explanation and determination*. Retrieved from <https://hdl.handle.net/1887/17879>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/17879>

Note: To cite this publication please use the final published version (if applicable).

Nederlandse samenvatting

Dit proefschrift handelt over verklaringen. Centraal staat de vraag wat verklaringen precies zijn. Wanneer we als voorbeeld een concreet feit nemen dat we zouden willen verklaren, dan is het over het algemeen heel gemakkelijk om aan te geven of een bepaald verhaal wel of niet een verklaring voor dat feit is; of in ieder geval, of het verhaal een verklaring zou zijn indien het waar was. (De concrete twijfels liggen dus meestal bij de *waarheid* van de potentiële verklaring, maar niet bij haar *verklarende kracht*.) Proberen we echter een theorie te geven die ons vertelt wat alle concrete gevallen van goede verklaringen met elkaar gemeen hebben, dan blijken we met allerlei problemen te maken te krijgen.

In **hoofdstuk 2** van het proefschrift wordt één verzamelingen theorieën over verklaringen kritisch tegen het licht gehouden: de zogenaamde unificatietheorieën. Het basisidee van deze aanpak, dat door verschillende auteurs op verschillende wijzen wordt uitgewerkt, is dat een theorie verklarende kracht heeft dan en alleen dan wanneer deze theorie feiten van verschillende aard unificeert, onder één noemer brengt. Zo zou de theorie van Newton verklaren waarom appels vallen doordat zij niet alleen het vallen van appels, maar ook het bewegen van de planeten en het ontstaan van de getijden beschrijft. De unificerende kracht van Newtons theorie zorgt ervoor dat we het vallen van appels begrijpen.

Het eerste deel van het hoofdstuk bespreekt de verschillende definities van unificatie die gegeven worden door aan de ene kant Kitcher, en aan de andere kant Schurz en Lambert. Het blijkt dat Kitchers definitie onhoudbaar is, aangezien zij leidt tot de valse conclusie dat de meest unificerende theorie een triviale theorie is die elk feit afleidt uit dat feit zelf. Kitcher probeert dit te omzeilen, maar zijn poging faalt. De definitie van Schurz en Lambert lijkt wel houdbaar.

Vervolgens is de vraag of unificatie en verklaring inderdaad zo nauw met elkaar samenhangen als de unificatietheorieën stellen. Het tweede deel van het hoofdstuk beargumenteert dat dit niet zo is. Noch de theorie van Kitcher, noch die van Schurz en Lambert, kan de asymmetrie van verklaringen die in

bekende causale scenario's optreedt reproduceren, tenzij deze asymmetrie er met de hand in wordt gestopt. Unificatie is dus kennelijk niet een *voldoende* voorwaarde voor verklarende kracht: één van de centrale eigenschappen van verklaringen kan er niet uit worden afgeleid. Daarnaast geef ik voorbeelden waaruit blijkt dat unificatie ook geen *noodzakelijke* voorwaarde voor verklarende kracht is: verklaringen kunnen anti-unificerend zijn. De enige reden dat verklaringen vaak gebruik maken van unificerende theorieën, is dat unificerende theorieën over het algemeen een grotere waarschijnlijkheid hebben om waar te zijn dan disunificerende theorieën; maar dit is voor de hoofdvraag van de dissertatie niet van belang.

In **hoofdstuk 3** wordt kritisch gekeken naar het idee dat er speciale soort inferentie is die gebruik maakt van verklaringen, de zogenoemde *inferentie naar de beste verklaring* of IBE (*inference to the best explanation*). De verklarende kracht van een potentiële verklaring zou hierbij een indicatie zijn voor de waarschijnlijkheid van die potentiële verklaring: een beter verklarende theorie heeft ook meer kans om waar te zijn. IBE zou een aanvulling moeten zijn op bestaande vormen van inferentie.

Aan de hand van een bespreking van de drie soorten argumenten die voor deze these gegeven worden laat ik zien dat er geen adequate verdediging van IBE is gegeven. Het eerste soort argument claimt dat eigenschappen die een theorie verklarende kracht geven die theorie ook een hogere waarschijnlijkheid geven. Maar in veel gevallen is dit onduidelijk, en in de andere gevallen is het zelfs simpelweg onwaar: zo zal een theorie die meer precies is daardoor beter verklaren, maar uiteraard ook minder waarschijnlijk zijn. Het tweede soort argument probeert IBE aan te wijzen in specifieke gevallen van succesvol wetenschappelijk onderzoek. Ik bespreek het meest uitgewerkte voorbeeld, Liptons analyse van het werk van Semmelweis, en toon aan dat Semmelweis geen andere inferenties hoefde te doen dan normale probabilistische inferenties. Het derde soort argument probeert aan te tonen dat Bayesiaanse probabilistische inferenties in feite gebruik maken van verklarende kracht; ook hier vinden we echter geen bewijs voor.

De conclusie van het derde hoofdstuk is aldus dat inferentie naar de beste verklaring niet lijkt te bestaan. Dat betekent uiteraard niet dat verklaringen geen rol spelen in de wetenschap. Integendeel: bij Liptons bespreking van Semmelweis hebben we juist gezien dat het een belangrijk deel van de wetenschappelijke methode is dat je op zoek gaat naar verklaringen. De waarheid van die verklaringen wordt echter niet getest door hun verklarende kracht tegen het licht te houden.

Hoofdstuk 4 gaat in op het belangrijkste alternatief voor de unificatietheorieën uit het tweede hoofdstuk: causale theorieën, dat wil zeggen, theorieën die verklaring koppelen aan de relatie van oorzaak en gevolg. Het basale

idee hier is dat oorzaken hun gevolgen verklaren. De (lastige) uitdaging is om dit basale idee in meer detail uit te werken.

Eén van de causale theorieën die op dit moment in de belangstelling staan is de interventietheorie van James Woodward. In dit hoofdstuk verdedig ik die theorie eerst tegen de claim dat zij circulair zou zijn: ik suggereer een manier waarop deze circulariteit omzeild, of in ieder geval onschadelijk gemaakt, kan worden. Daarna is mijn doel om de theorie van Woodward dusdanig te generaliseren dat zij niet alleen causale verklaringen, maar alle soorten verklaringen kan vatten. Hiertoe bespreek ik uitgebreid het fenomeen van wiskundige verklaringen. De beste theorie hierover is van de hand van Steiner, maar deze heeft enkele zwakheden. Ik laat zien hoe deze zwakheden opgelost en de onderliggende intuïties toch behouden kunnen worden door een theorie van wiskundige verklaringen te gebruiken die dezelfde formele structuur heeft als Woodward's causale theorie. Vervolgens herhalen we deze procedure voor verklaringen van natuurwetten door middel van meer omvattende natuurwetten. Aan het eind van het hoofdstuk wordt een theorie van verklaringen gegeven die dezelfde vorm heeft als Woodward's theorie, maar volkomen algemeen is. Deze theorie zal als voorlopige versie fungeren van de determinatietheorie van hoofdstuk 6.

In **hoofdstuk 5** kijken we naar het soort contrasten dat in verklaringen voorkomt. Het is bekend, onder andere uit het werk van Van Fraassen, dat veel verklaringen een ‘contrastklasse’ hebben: een verzameling feiten waartegen het te verklaren feit wordt afgezet. We vragen waarom Adam de appel at, *in plaats van* deze niet te eten. Dat is een andere vraag, met een ander antwoord, dan de vraag waarom Adam, *in plaats van* iemand anders, de appel at.

Er bestaat echter ook een ander soort contrast in verklaringen, dat in de literatuur met het eerste soort contrast wordt verward. Ik noem dit een ‘contrast of parallels’, en een typisch voorbeeld is de vraag waarom Piet wel en Jan geen voldoende voor zijn tentamen heeft gehaald. In het Engels kan deze vraag gesteld worden met dezelfde frase, “rather than”, die gebruikt wordt bij het aanduiden van een contrastklasse. In het Nederlands zien we direct dat “in plaats van” niet gebruikt kan worden: immers, Piets voldoende zorgt er niet voor dat Jan geen voldoende kan halen, en dus is er geen sprake van een “in plaats van”. Wanneer we een contrast van parallellen opstellen, willen we weten welke verschillen ervoor gezorgd hebben dat in het ene geval de ene mogelijkheid (voldoende *in plaats van* onvoldoende), en in het andere geval de andere mogelijkheid (onvoldoende *in plaats van* voldoende) werkelijkheid is geworden.

Wanneer we dit onderscheid consequent hanteren kunnen we laten zien dat de tegenvoorbeelden die zijn ingebracht tegen de theorie dat alle verk-

laringen een contrastklasse bevatten, niet kloppen. Ik beargumenteer dat deze ‘contrastieve theorie’ valt te prefereren boven het alternatief, de ‘conjunctieve theorie’, die juist beweert dat geen enkele verklaring een contrastklasse bevat, en zogenaamde contrastklassen geanalyseerd moeten worden in termen van conjuncties van niet-contrastieve verklaringen.

In de tweede helft van het hoofdstuk laat ik zien dat in een goede verklaring niet alleen het explanandum (het feit dat verklaard moet worden), maar ook het explanans (het feit of de feiten die als verklaring worden gegeven), in een contrastklasse zijn ingebed. Zo kom ik tot de ‘double-contrast theory’ van verklaringen. Er wordt aangetoond dat deze theorie, hoe minimaal zij ook is, reeds de vraagstukken rondom irrelevante informatie kan oplossen: we kunnen ermee begrijpen wanneer het toevoegen van irrelevante informatie een verklaring van zijn geldigheid berooft en wanneer dit niet het geval is, een vraagstuk dat voor andere theorieën van verklaring vaak lastig is.

We komen dan in **hoofdstuk 6** bij de formulering van mijn eigen determinatietheorie. Eerst beargumenteer ik dat het idee van determinatie een centrale rol speelt in ons concept van verklaren: *A* verklaart *B* alleen wanneer *A* er op de een of andere manier voor zorgt dat, determineert dat, *B* het geval is. De combinatie van dit inzicht met de eerdere conclusies over interventies en contrasten leidt tot mijn centrale claim: een verklaring is een waar deductief argument waarin zowel de conclusie als minstens één van de premissen de vorm heeft van een feit ingebed in een contrastklasse, en waar een interventie die zo’n in de premissen genoemd feit verandert in een van zijn contrasten, er toe zou leiden dat het feit in de conclusie verandert in een van zijn contrasten. Deze claim wordt in de rest van het hoofdstuk formeel uitgewerkt, en gedeeltelijk verdedigd.

De belangrijkste verdediging die nodig is, is het aantonen dat verklaringen inderdaad een deductief/deterministisch karakter hebben. Er zijn immers veel schijnbare tegenvoorbeelden: probabilistische en indeterministische verklaringen. In **hoofdstuk 7** laat ik eerst zien dat een grote klasse van ogenschijnlijk niet-deductieve verklaringen gereconstrueerd kan worden als instantiaties van de determinatietheorie, om precies te zijn van een model hiervan dat ik het niet-specifiek deductieve (ND-)model noem. Vervolgens laat ik zien dat het ND-model en dus ook de determinatietheorie in overeenstemming zijn met de verschillende intuïties in het ‘size debate’, het debat over de vraag of je alleen kan verklaren wanneer de premissen de conclusie een hoge waarschijnlijkheid geven. Wanneer we twee plausibele aannames maken, namelijk (a) dat zelfs afgekorte verklaringen in ieder geval een contrastieve premisse van de bijbehorende volledige verklaring moeten noemen, en (b) dat een ND-verklaring kan worden omgevormd tot een betere verklaring door de niet-specifieke premisse meer specifiek te maken; dan volgt

dat de determinatietheorie ook in overeenstemming is met alle intuïties in het ‘change debate’, het debat over de vraag of alleen feiten die de kans op het te verklaren feit doen toenemen verklarend zijn. De conclusie van dit deel van het hoofdstuk is dat de determinatietheorie recht doet aan de rol van waarschijnlijkheid en statistiek in verklaringen van deterministisch geproduceerde gebeurtenissen.

Het is duidelijk dat de determinatietheorie geen recht kan doen aan verklaringen van indeterministische gebeurtenissen. Het laatste deel van het hoofdstuk is er dan ook op gericht aan te tonen dat deze verklaringen geen rechten hebben, aangezien ze helemaal niet bestaan. Daarmee is de grootste horde voor de determinatietheorie genomen.

Eenzelfde soort vraagstukken wordt behandeld in **hoofdstuk 8**, waar ik een verzameling argumenten van Michael Strevens bespreek. Het doel van zijn argumenten is om te laten zien dat zelfs in deterministische gevallen indeterministische verklaringen soms het beste zijn. Hiertoe stelt hij een hiërarchie van verklaringen op: de laag-niveau deterministische verklaring, de statistisch-deterministische verklaring en de indeterministische verklaring, waarbij de laatste de beste is. Ik laat eerst zien dat de tweede soort verklaring niet bestaat, waarmee ook het belangrijkste argument tegen de laag-niveau deterministische verklaring (namelijk dat deze veel irrelevanties bevat aangezien er een veel simpeler deterministische verklaring is) komt te vervallen. Daarna beargumenteer ik dat de laag-niveau deterministische verklaring beter is dan de indeterministische verklaring, omdat (a) kritische gebeurtenissen altijd verklarende kracht hebben; en (b) Strevens’ theorie het plausibele principe dat ik ‘deductive closure’ noem weerspreekt, namelijk het principe dat als we een goede verklaring hebben voor A , en we weten dat A impliceert B , we dan ook een goede verklaring voor B hebben.

De conclusie van deze twee hoofdstukken is dat het goed is vol te houden dat alle verklaringen deterministische verklaringen zijn.

Het laatste hoofdstuk, **hoofdstuk 9**, fungeert als een *capita selecta* van onderwerpen die voor de determinatietheorie van belang zijn. Ten eerste beargumenteer ik dat we kunnen claimen dat alle verklaringen argumenten zijn zonder dat dit de slechte gevolgen heeft die sommige auteurs hieraan hebben toegeschreven. Daarna laat ik zien dat verklaringen geen gebruik hoeven te maken van wetmatigheden, en maak ik in dit verband duidelijk hoe herbeschrijvingen verklarend kunnen zijn. We onderzoeken hoe en onder welke omstandigheden verschillende verklaringen met elkaar kunnen worden gecombineerd tot grotere verklaringen, waarbij blijkt dat dit niet kan wanneer er sprake is van referentiële ambiguïteit. Ik beargumenteer dat de methode die in dit proefschrift gebruikt is niet geproblematiseerd wordt door het verschil tussen verklaren en begrijpen dat de laatste tijd zeer in de belangstelling

staat; en dat de determinatietheorie ons in staat stelt een verband te zien tussen de meer klassieke termen *Erklären* en *Verstehen*. Uiteindelijk maak ik duidelijk onder welke omstandigheden een verklaring die aan de determinatietheorie voldoet geen enkele verklarende kracht heeft: de verklaring heeft alleen dan verklarende kracht wanneer ze de juiste soort counterfactuele uitspraken impliceert, met als extra eis dat deze niet al door de vraag naar de verklaring werden geïmpliceerd. We vinden hier ook aanwijzingen dat de notie van interventie niet geheel subject-onafhankelijk is, zodat ook verklaring als geheel een niet puur objectieve notie zou zijn.

Als geheel heeft dit proefschrift de ambitie om op twee manieren bij te dragen aan de filosofie van verklaring. Ten eerste, door een nieuwe en mogelijk betere theorie van verklaringen te postuleren en te verdedigen. Ten tweede, door licht te laten schijnen op een scala aan gerelateerde onderwerpen, zoals unificatie, inferentie naar de beste verklaring, mathematische verklaring, contrastieve verklaring, en deterministische versus non-deterministische verklaring. Er is nog veel te doen, maar enige vooruitgang is hopelijk geboekt.