Erasmus Student Journal of Philosophy

#9 | December 2015

ESJP
#9 | 2015

Editorial

This edition, not only papers written for MA courses, but also a number of BA theses were nominated for publication in the Erasmus Student Journal of Philosophy. This is a trend we most certainly welcome. We consider it our primary task to enable excellent students from both the BA and MA programs to publish their work, and since BA theses tend to reflect the individual research interests of their authors, they add additional variety to the already impressive range of topics published in the ESJP so far. I know for certain that students who publish their BA thesis in our journal gain valuable experience during the editorial process. This experience translates itself to a set of skills that can subsequently be applied when writing MA papers. It is for this reason I would like to encourage all BA students, also from the Double Degree program, to make sure their thesis gets nominated for pubication in the ESJP, should a high grade be obtained.

As the philosophy of mathematics has been unrepresented so far, we are proud to now be able to feature a paper (that was originally submitted as a BA thesis) on gödelian mathematical intuition. This issue also contains work retracing the historical reasons for the mathematization of economics. Finally, there is a paper discussing Daniel Dennett's model of consciousness. I would like to ask professors to keep nominating excellent work, and students to write papers with the possibility of publication in the ESJP in mind.

We are doing our best to keep up with the recent developments at our faculty. Sadaf Soloukey is the first editor to join the board that is enrolled in the Double Degree BA program. She is also a student of medicine. If all Double Degree students are as talented as she is, then we certainly hope to be able to welcome more editors from this program in the near future. Daniël Zevenhuizen also joined the team at a later stage in the editorial process. I look forward to seeing him develop his remarkable skills over the course of the next editions. As for myself; I was pleasantly surprised when I was approached by Huub Brouwer to return as editor-in-chief while I was still studying at Osaka University. I am grateful to be given another chance at working with a fantastic team of editors for at least one more edition. I would like to thank the members of the Advisory Board for entrusting me with this wonderful position once more. Dyonne Hoogendoorn enthusiastically took up the position of secretary. I feel blessed to have benefitted from her hard work during this edition.

Unfortunately, the time has come to say goodbye to longtime editor Thijs Heijmeskamp. He has been a member of the editorial board for more than four years. His presence and valuable experience will be missed. I am thrilled to be able to say that Thijs is leaving us a parting gift in the form of an article published in this very issue. The article on Daniel Dennett was written by him. On behalf of the entire editorial team, I would like to wish him the best of luck with all of his future endaevors.

Special thanks go out to all who were involved in making this issue. I am especially grateful to the professors and PhD students who continue to invest their valuable time into writing reviews. It is no exaggeration to say that without them, the ESJP is not possible. The authors of this issue have also shown themselves willing to implement a great number of comments over the course of many weeks. I understand that refining one's written work while having a job or being a full-time student is extremely demanding. I would therefore like to thank them for their hard work. Finally, a word of thanks to our faithful readers. I hope you will enjoy reading these pages.

Dennis Prooi

Editor-in-chief

About the Erasmus Student Journal of Philosophy

The Erasmus Student Journal of Philosophy (ESJP) is a double-blind peer-reviewed student journal that publishes the best philosophical papers written by students from the Faculty of Philosophy, Erasmus University Rotterdam. Its aims are to further enrich the philosophical environment in which Rotterdam's philosophy students develop their thinking and bring their best work to the attention of a wider intellectual audience. A new issue of the ESJP appears on our website every July and December.

To ensure the highest possible quality, the ESJP only accepts papers that (a) have been written for a course that is part of the Faculty of Philosophy's curriculum and (b) nominated for publication in the ESJP by the teacher of that course. Each paper that is published in the ESJP is subjected to a double-blind peer review process in which at least one other teacher and two student editors act as referees.

The ESJP encourages students to keep in mind the possibility of publishing their course papers in our journal, and to write papers that appeal to a wider intellectual audience.

More information about the ESJP can be found on our website:

www.eur.nl/fw/esjp

Contact: esjp@fwb.eur.nl

ISSN: 2212-9677

Editorial Board

Dennis Prooi (Editor-in-chief)
Dyonne Hoogendoorn (Secretary)
Fabian van Dijk
Willem-Rutger van Dijk
Matthijs Geleijnse
Thijs Heijmeskamp
Floris van der Pol
Stefan Schwarz
Rui Francisco de Souza Só Maia
Sadaf Soloukey
(Guest Editor)
Daniël Zevenhuizen
(Guest Editor)

Advisory Board

Huub Brouwer
Jasper van den Herik
Julien Kloeg
Myrthe van Nus

Supervisory Board

dr. B. Leeuwenburgh
prof. dr. F.A. Muller
prof. dr. J.J. Vromen

All work in this issue of the Erasmus Student Journal of Philosophy is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit http://creativecommons.org/licenses/by-nc/3.0/

Disclaimer

Although the editors of the Erasmus Student Journal of Philosophy have taken the utmost care in reviewing the papers in this issue, we cannot exclude the possibility that they contain inaccuracies or violate the proper use of academic referencing or copyright in general. The responsibility for these matters therefore remains with the authors of these papers and third parties that choose to make use of them entirely. In no event can the editorial board of the Erasmus Student Journal of Philosophy or the Faculty of Philosophy of the Erasmus University Rotterdam be held accountable for the contents of these papers.

ESJP
#9 | 2015

In this issue

In 'De Houdbaarheid van Kurt Gödels Wiskundige Intuïtie', Hugo Hogenbirk argues that the notion of gödelian mathematical intuition should not be rejected straight out of hand as one that is too 'exotic'. Hogenbirk attempts to demonstrate the viability of the notion by defending it against attacks from Paul Benacerraf and Charles Chihara.

In 'Economics, Complexity and the Disenchantment of the Social World', Sam van Dijck shows how economics became a mathematical science in spite of concerns raised by economists such as Alfred Marshall, Friedrich Hayek and John Maynard Keynes. These influential figures questioned the ability of mathematics to predict an economic reality that is highly complex. Van Dijck suggests that the mathematization of economics occurred because of an overriding demand for predictive power which in turn might be the result of a process Max Weber called 'disenchantment'.

In 'Dennett's Drafters', Thijs Heijmeskamp argues that Daniel Dennett's work 'Consciousness Explained' fails to provide an explanation of consciousness. Dennett's model, Heijmeskamp claims, is based on a wrong characterization of our phenomenology. As a consequence, Dennett has to rely on the mysterious notion of the 'probe' in order to be able to provide an account of phenomenal experience.

ESJP
#9 | 2015

Table of contents

ESJP
#9 | 2015

De Houdbaarheid van Kurt Gödels Wiskundige Intuïtie

Hugo Hogenbirk

Inleiding

Wiskunde heeft filosofen al eeuwenlang beziggehouden en nog steeds bestaat er een levendig onderzoek naar dit onderwerp binnen de filosofie. De ontologie van de wiskunde is een belangrijk onderdeel van deze filosofie van de wiskunde. Binnen de ontologie van de wiskunde vindt men vragen als: 'wat is een getal?', 'over wat of welke dingen gaat wiskunde?' en 'gaat wiskunde over dingen?' Een mogelijk antwoord op deze vragen wordt gegeven door het platonisme. Het platonisme stelt dat wiskunde gaat over objecten die abstract en onafhankelijk van het menselijk denken zijn. Bovendien zijn het deze objecten die de wiskunde waar maken. Eén van de grootste problemen waar een wiskundig platonist mee te maken heeft, is dat hij niet kan verklaren hoe het kan dat wij kennis van deze abstracte objecten hebben. Voor dit epistemologische probleem moet de platonist een oplossing zien te formuleren. Een mogelijke oplossing voor dit probleem vinden we bij de logicus Kurt Gödel. Het is deze oplossing, Gödels notie van 'wiskundige intuïtie' (voorts: wiskundige intuïtie), die het onderwerp is van dit paper.

Dit paper zal zich richten op de houdbaarheid van wiskundige intuïtie zoals ingezet door Kurt Gödel, namelijk als een epistemologisch vermogen om ons inzicht in wiskunde en haar axioma's mee te verschaffen. Dit paper bestaat uit vijf onderdelen. Ten eerste zal ik uiteenzetten wat wiskundige intuïtie bij Gödel inhoudt en hoe deze positie zich verhoudt tot het wiskundig platonisme. Gödel achtte deze stroming binnen de ontologie van de wiskunde zeer nauw verbonden met wiskundige intuïtie. Ten tweede zal ik twee kritieken op Gödels opvatting over wiskundige intuïtie behandelen. De auteurs van deze twee kritieken, Paul Benacerraf en Charles Chihara, zetten argumenten in die in hun ogen niet alleen de plausibiliteit maar ook de houdbaarheid van wiskundige intuïtie aantasten. Hoewel er andere kritieken mogelijk zijn, wil ik me in dit paper tot deze twee kritieken beperken. Ten derde zal Gödels eigen argument voor wiskundige intuïtie uiteen worden gezet. Dit argument sluit aan op de praktijk van het ontdekken van nieuwe axioma's in de wiskunde, geïllustreerd aan de hand van de continuümhypothese. Ten vierde beschouw ik een mogelijk filosofisch gevolg van Gödels eerste incompleetheidsstelling. Dit gevolg heeft te maken met de aard van wiskundige kennis. In dit verband zal ik tevens de kenleer van Immanuel Kant en David Hume bespreken. Ten slotte beschouw ik een argument voor wiskundige intuïtie op basis van de functie die wiskundige afbeeldingen zouden moeten innemen in de wiskundige praktijk.

Een aantal van de argumenten die ik doorheen dit paper zal geven zijn directe argumenten *voor* wiskundige intuïtie. Het doel van dit paper is echter niet de lezer te overtuigen van de correctheid of het bestaan van wiskundige intuïtie. Ik wil niet pleiten voor de *geldigheid*, maar voor de *houdbaarheid* van wiskundige intuïtie.

§1 Waar hebben we het over?

1.1 Platonisme

Om de functie en werking van Gödels wiskundige intuïtie te begrijpen is het van belang om eerst zijn positie binnen de ontologie van de wiskunde te bezien. Deze positie, die van het wiskundig platonisme, definieert Gödel als volgt.

> [Platonism is] the view that mathematics describes a non-sensual reality, which exists independently both of the acts and [of] the dispositions of the human mind and is only perceived, and probably perceived very incompletely, by the human mind. (Gödel, 1995a, p. 323)

Wat we hier zien is dat bij Gödel platonisme uit een aantal claims bestaat. Ten eerste, wiskundige objecten (de objecten die worden beschreven door de wiskunde) bestaan en zijn *non-sensual*. Daarnaast bestaan deze objecten onafhankelijk van onze eigen geest. Ten slotte zien we dat Gödel hier vooruit blikt op het probleem waar dit paper over gaat, namelijk de toegang tot deze objecten door middel van de menselijke geest. In het artikel uit de *Stanford Encyclopedia of Philosophy* (SEP) over wiskundig platonisme zien we de drie belangrijke delen van de platonistische these ook op een dergelijke manier opgenoemd:

> Mathematical platonism can be defined as the conjunction of the following three theses:
>
> Existence.
> There are mathematical objects.
>
> Abstractness.
> Mathematical objects are abstract.
>
> Independence.
> Mathematical objects are independent of intelligent agents and their language, thought, and practices. (Linnebo, 2013a)

Wat ontbreekt in deze definitie, maar wat we wel terugzien in Gödels definitie, is een nadere omschrijving van de manier waarop er kennis moet worden verkregen van deze objecten. Echter, scherp bezien is dit geen onderdeel van de definitie van wiskundig platonisme, aangezien er nog gediscussieerd wordt over wat deze toegang tot wiskundige objecten zou moeten inhouden. Vandaar dat we ons zullen houden aan de definitie van het artikel uit de SEP.

De eerste van de drie claims van de wiskundig platonist behoeft nauwelijks verdere uitleg. Wiskunde gaat over wiskundige objecten en deze wiskundige objecten bestaan. De tweede claim kan uiteengezet worden in een tweetal andere eigenschappen. De eerste is dat iets pas abstract kan zijn als het niet ruimte-tijdelijk is, en de tweede is dat iets pas abstract kan zijn als het niet causaal effectief is (Linnebo, 2013a). Dat wil zeggen, de objecten waar de wiskundig platonist het over heeft hebben geen causale effecten in de wereld en hebben geen omvang, locatie, moment en tijdsduur. De derde claim van de wiskundig platonist is de onafhankelijkheid van onze geest. Dit betekent dat de abstracte, wiskundige, objecten bestaan ongeacht of er mensen zijn of niet. Met andere woorden, als er niemand iets zou weten over wiskunde (als er bijvoorbeeld niemand zou zijn die inziet dat één en één twee maakt) maken één en één alsnog twee, ondanks ons gebrekkig inzicht daarin.

Aangezien wiskundig platonisme het bestaan van een merkwaardig soort objecten veronderstelt, doemt de epistemologische vraag op hoe het kan dat wij kennis van deze objecten hebben. De vraag naar hoe we kennis van wiskundige objecten hebben blijkt een moeilijkheid te zijn waar de platonist heden ten dage nog steeds mee worstelt (Linnebo, 2013a). Het is als antwoord op dit probleem dat Gödels wiskundige intuïtie zijn intrede doet.

1.2 Gödels wiskundige intuïtie

Een nadeel van Gödels wiskundige intuïtie is dat Gödel zelf de positie weinig uitwerkt en uitlegt. Een van de meest expliciete verwoordingen van zijn positie vinden we in het herziene artikel *What is Cantor's continuum problem*. In het in 1964 toegevoegde addendum lezen we:

> But, despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e. in mathematical intuition, than in sense perception, which induces us to build up physical theories and to expect that future

> sense perceptions will agree with them, and, moreover, to believe that a question not decidable now has meaning and may be decided in the future. (Gödel, 1995b, p. 268)

Hieruit kunnen we opmaken dat we wiskundige intuïtie moeten zien als een extra zintuig. Zoals fysieke objecten zich door middel van zintuigen aan ons opdringen en ons kennis over zich laten verschaffen, op eenzelfde manier verkrijgen we kennis van wiskundige objecten. Direct hierna benadrukt Gödel dat we niet te maken hebben met een zintuig in de klassieke zin van het woord:

> It should be noted that mathematical intuition need not be conceived of as a faculty giving an immediate knowledge of the objects concerned. Rather it seems that, as in the case of physical experience, we form our ideas also of those objects on the basis of something else which is immediately given. Only this something else here is not, or not primarily, the sensations. (…) but, as opposed to the sensations, their presence in us may be due to another kind of relationship between ourselves and reality. (Gödel, 1995b, p. 268)

Hier ontwaren we een probleem waar één van de commentaren zich ook op zal richten. Gödel heeft moeite om duidelijk over te brengen wat wiskundige intuïtie precies is. Het is een vermogen van ons denken dat ongeveer functioneert als een zintuig maar er geen is. Wiskundige intuïtie zorgt ervoor dat wij een relatie hebben met de wiskundige realiteit. Het is precies de verdere beschrijving van deze relatie die bij Gödel onuitgewerkt blijft. Desondanks zijn er wel een aantal eigenschappen gegeven, bijvoorbeeld dat deze relatie analoog loopt aan zintuiglijke perceptie en dat het een band legt tussen ons denken en objecten. In dit paper zullen we zien dat de beperkte hoeveelheid informatie die Gödel verschaft een aantal concrete uitwerkingen kan opleveren, die ons ook weer meer zullen vertellen over de aard van wiskundige intuïtie.

Hiermee weten we ongeveer wat Gödel in gedachten heeft als hij spreekt over wiskundige intuïtie. Op deze positie van Gödel zijn verscheidene kritieken gekomen. Vaak zijn deze kritieken, zoals we zullen zien, zeer sterk, waarbij de plausibiliteit van wiskundige intuïtie wordt afgedaan als zeer klein en deze wordt weggezet als een (te) exotische filosofische positie.

§2 Aanvallen op de houdbaarheid

2.1 Paul Benacerraf: kennisleer en wiskundige intuïtie

Zoals we hebben gezien geeft Gödel met zijn wiskundige intuïtie een verklaring voor onze kennis van wiskundige objecten. De eerste kritiek die we zullen bespreken speelt in op onze reguliere ideeën over kennis en hoe Gödels wiskundige intuïtie niet op deze aansluit. Paul Benacerraf maakt dit punt in zijn artikel *Mathematical Truth* (1973). Benacerraf beargumenteert dat onze beste theorieën van kennis eisen dat tussen de persoon die ergens kennis over heeft en het ding waar hij kennis over heeft enig causaal verband bestaat. Zo schrijft hij:

> (...) the principal defect of the standard [platonic/Godelian] account is that it appears to violate the requirement that our account of mathematical truth be susceptible to integration into our over-all account of knowledge. (Benacerraf, 1973, p. 670)

> I favor a causal account of knowledge on which for X to know that S is true requires some causal relation to obtain between X and the referents of the names, predicates, and quantifiers of S. (Benacerraf, 1973, p. 671)

Benacerraf stelt dat kennis van een object veronderstelt dat er een causaal verband bestaat tussen de kenner en dat object. Dit onderbouwt hij door aan te geven dat een dergelijke theorie van kennis goed aansluit bij onze intuïties. Hij laat dit zien door middel van een analyse van de redenen waarom iemand ergens geen kennis van heeft (Benacerraf, 1973, pp. 671-672). Als we weten dat iets waar is, en we mogen aannemen dat iemand in staat is zaken uit elkaar af te leiden, hoe verklaren we het dan als iemand gebrek heeft aan kennis van een bepaalde zaak? Benacerraf stelt dat we dit doen door duidelijk te maken dat de betreffende persoon niet in causaal contact staat met het object waar de kennis over zou gaan. Oftewel, kennis is afhankelijk van een causaal verband tussen persoon en kenobject.

Nu moet duidelijk worden gemaakt dat er geen sprake is van een causaal verband tussen de platonist en de wiskundige objecten waar hij meent kennis van te hebben. Hierop zegt Benacerraf:

> If, for example, numbers are the kinds of entities they are normally taken to be, then the connection between the truth conditions for the statements of number theory and any relevant events connected with the people who are supposed to have mathematical knowledge cannot be made out. It will be impossible to account for how anyone knows any properly number-theoretical propositions. (Benacerraf, 1973, p. 673)

Met andere woorden, causaliteit gaat niet samen met een opvatting die stelt dat getallen platoonse objecten zijn. De eerder gegeven definitie van wiskundig platonisme stelt dat wiskundige objecten niet tijdelijk, ruimtelijk of causaal zijn. Stellen dat wiskundige objecten wel causaal effectief zijn, zal voor de gödeliaan die zich tegen de kritiek van Benacerraf wil verdedigen dus geen optie zijn.

Verderop zal duidelijk worden gemaakt hoe de gödeliaan deze doodsteek kan ontwijken. De aanname die door Benacerraf wordt gemaakt waar ik tegenargumenten voor zal formuleren vinden we in het volgende citaat terug:

> (...) the principal defect of the standard [platonic/Gödelian] account is that it appears to violate the requirement that our account of mathematical truth be susceptible to integration into our over-all account of knowledge. (Benacerraf, 1973, p. 670)

Als de gödeliaan aannemelijk kan maken dat van wiskundige kennis niet geëist mag worden dat ze hetzelfde is als andere soorten van kennis, dan gaat de zojuist genoemde kritiek van Benacerraf niet langer op. Zoals verderop duidelijk zal worden kan dit door het speciale karakter van de wiskunde te bewijzen.

2.2 Charles Chihara: verklarende kracht

Charles Chihara valt de houdbaarheid van wiskundige intuïtie op een ander manier aan. Hoewel de meeste van zijn argumenten gericht zijn tegen Gödels platonisme, vinden we in Chihara's artikel (1982) ook argumenten gericht tegen Gödels wiskundige intuïtie.

Gödels analogie tussen visuele perceptie en wiskundige intuïtie zet Chihara ertoe aan om de betrouwbaarheid van deze twee te vergelijken. Chihara identificeert een probleem dat zich voordoet bij deze vergelijking die nadelig uitpakt voor wiskundige intuïtie. Empirische waarnemingen kunnen namelijk worden gemeten. We kunnen tests herhalen en verduidelijken wie de waarnemingen heeft gehad. Dit in tegenstelling tot kennis die het effect is van wiskundige intuïtie:

> What sort of data about mathematical experience do we have that is comparable to the scientist's Brownian movement data? What is this experience of axioms forcing themselves upon us as being true to which Gödel appeals? And how many people have had these experiences? What sort of people? And under what conditions? Surely, there is something suspicious about an argument for the existence of sets that rests upon data of so unspecified and vague a nature, where even the most elementary sorts of controls and tests have not been run. It is like appealing to experiences vaguely described as 'mystical experiences' to justify belief in the existence of God. (Chihara, 1982, p. 215)

Chihara laat in de laatste zin zien dat hij wiskundige intuïtie onovertuigend vindt. Het gebrek aan de testbaarheid van wiskundige intuïtie doet hem stellen dat wiskundige intuïtie eigenlijk niet kan fungeren als rechtvaardiging of verklaring van wiskundige kennis. Even verderop schrijft hij:

> But here again, the 'explanation' offered is so vague and imprecise as to be practically worthless: all we are told about how the 'external objects' explain the phenomena is that mathematicians are 'in some kind of contact' with these objects. What empirical scientist would be impressed by an explanation this flabby? (Chihara, 1982, p. 217)

Wederom claimt Chihara dat wiskundige intuïtie het grote probleem heeft geen werkelijke, substantiële uitleg te zijn. Wiskundige intuïtie als uitleg levert immers geen informatie en roept de verdenking op een ad-hocoplossing te zijn voor problemen waar de wiskundig platonist mee zit. Chihara constateert een gebrek aan concretisering, bij zowel de toetsbaarheid als bij de uitleg over de werking van wiskundige intuïtie. Zoals Maddy het verwoordt als het gaat over voorgaand citaat: 'And finally, he [Chihara, HH] questions whether Gödel's intuition offers any explanation at all.' (Maddy, 1989, p. 1135)

Chihara's argument is samen te vatten met een analogie. Als we willen verklaren dat opium slaapverwekkend is, schieten we niks op met de verklaring dat dit komt door de slaapverwekkendheid van opium. Voor Chihara is het gebrek aan concrete effecten op de wereld van wiskundige intuïtie (zoals bijvoorbeeld testbare effecten, of effecten op de wiskundige praktijk) de reden dat wiskundige intuïtie niet als verklaring kan fungeren. Net zoals bij de slaapverwekkendheid van opium het geval is, komt het aannemen van wiskundige intuïtie op niets meer neer dan het geven van een naam aan een probleem en dan zeggen dat je deze hebt verklaard.

De gödeliaan zal in reactie hierop moeten laten zien dat er redenen zijn om aan te nemen dat er zoiets als wiskundige intuïtie bestaat, bijvoorbeeld door middel van het leveren van een concreet effect van wiskundige intuïtie. Hiermee kan hij laten zien dat wiskundige intuïtie meer is dan een ad-hocoplossing voor het epistemologische probleem waar hij als platonist nu eenmaal een oplossing voor dient aan te leveren. Ik zal betogen dat de gödeliaan in staat is om dergelijke concrete effecten van wiskundige intuïtie te leveren. Dit zal tot gevolg hebben dat Chihara's aanval, die steunt op het gebrek aan extern, concreet bewijs voor wiskundige intuïtie, weerlegd kan worden.

§3 De continuümhypothese

In *What is Cantor's continuüm problem?* vinden we een argument voor wiskundige intuïtie dat inhaakt op de continuümhypothese. Het argument dat Gödel geeft is dat de continuümhypothese laat zien dat alleen wiskundige intuïtie in staat is om te verklaren dat er onafhankelijke stellingen zoals de continuümhypothese kunnen bestaan. Hoewel Gödel dit inzet als een direct argument voor wiskundige intuïtie zal er op basis hiervan een ander argument worden geformuleerd dat ingaat tegen Chihara. Immers, als er een verband bestaat tussen een bepaalde wiskundige praktijk (waar Gödel wiskundige intuïtie noodzakelijk voor achtte), dan kan wiskundige intuïtie niet langer een gebrek aan concrete effecten worden aangerekend.

3.1. Wat is de continuümhypothese?

Gödel introduceert de continuümhypothese op de volgende wijze:

> The problem is to find out which one of the ℵ's [aleph numbers, HH] is the number of points of a straight line or (which is the same) of any other continuum (of any number of dimensions) in a Euclidean space. Cantor, after having proved that this number is greater than ℵ 0, conjectured that it is ℵ 1. An equivalent proposition is this: Any infinite subset of the continuum has the power either of the set of integers or of the whole continuum. This is Cantor's continuum hypothesis. (Gödel, 1995b, p. 256)

De continuümhypothese is de claim dat de eerstvolgende zogeheten graad van oneindigheid na die van de verzameling ℕ van de natuurlijke getallen (0, 1, 2, 3, ….) die van de verzameling ℝ van de reële getallen is (-2, -√2, 3, π, 5.82). Dat wil zeggen, tussen de verzameling ℕ van de natuurlijke getallen en de verzameling ℝ van de reële getallen zijn geen oneindige verzamelingen die groter zijn dan ℕ en kleiner dan ℝ. Wat wil het echter zeggen dat de continuümhypothese stelt dat de ene oneindige verzameling groter is dan de andere oneindige verzameling? De manier waarop dit wordt nagegaan is door middel van het zoeken naar een één-op-één relatie tussen de twee oneindige verzamelingen. Als ik in staat ben een dergelijke relatie te vinden, dan zijn de twee verzamelingen even groot. Als daarentegen kan worden bewezen dat een dergelijke relatie niet bestaat, dan weten we dat de verzamelingen niet even groot zijn.

Nagaan of een één-op-één relatie werkelijk één-op-één is gaat op de volgende manier. We hebben een één-op-één correspondentie tussen (de elementen van) een verzameling V en (de elementen van) een verzameling

W als er een functie een functie F van V naar W is, zodanig dat F aan verschillende elementen van V verschillende elementen van W toevoegt en zodanig dat er bij elk element w van W precies één element v in V is met F(v) = w. Een één-op-één correspondentie F tussen V en W associeert dus met elk element v in V precies een element F(v) in W zodanig dat er bij elk element w in W ook precies een element v in V is met F(v) = w. Anders gezegd, we geven een functie die elk element uit de ene verzameling koppelt aan precies één element uit de andere verzameling en vice versa. We zullen twee voorbeelden geven van het redeneren met deze begrippen, een met eindige verzamelingen en een andere met oneindige verzamelingen.

Neem de verzameling X = {1, 2, 3} en de verzameling Y = {1, 2, 3, 4, 5}. Als we een functie willen maken van X naar Y die alle elementen van Y beslaat, dan gaat dit niet lukken. Immers X heeft niet genoeg elementen om alle elementen van Y te beslaan. X heeft drie elementen, en deze drie kunnen nooit vijf elementen bereiken. Als we nu een functie van Y naar X willen maken die alle elementen van X beslaat zonder dubbelen lukt dit ook niet. Er zijn immers vijf elementen in Y die verdeeld moeten worden over drie of minder elementen van X. We zien dat er geen één-op-één relatie kan worden gemaakt tussen X en Y, wat overeen komt met de intuïtie dat een verzameling met drie elementen niet even groot is als een verzameling met vijf elementen.

Zoals we hierboven hebben gezien kunnen we dus bij een eindige verzameling eenvoudigweg de elementen tellen en tot de conclusie komen dat de ene verzameling er meer heeft dan de andere en dus groter is. Bij oneindige verzamelingen werkt dit iets anders; hier kan een verzameling een strikte deelverzameling (een verzameling die bestaat uit slechts een deel van de elementen van een andere verzameling maar niet uit alle elementen van die andere verzameling) maar nog steeds even groot zijn. Neem bijvoorbeeld de verzameling ℕ van de natuurlijke getallen en de verzameling $\mathbb{N}_{oneven}$ van de oneven natuurlijke getallen. In tegenstelling tot wat wellicht valt te verwachten, kunnen we wel degelijk een één-op-één relatie geven tussen de positieve oneven getallen en ℕ, ondanks dat $\mathbb{N}_{oneven}$ een strikte deelverzameling is van ℕ (immers $\mathbb{N}_{oneven}$ is ℕ zonder de positieve even getallen). Inzien dat de twee verzamelingen even groot zijn doen we door middel van onderstaande afbeelding en hieruit leiden we de functie af die de één-op-één correspondentie zal geven (De Swart, 1989, p. 28):

1	3	5	7	9	11	13	15	17	19	21	...	$\mathbb{N}_{oneven}$
|	|	|	|	|	|	|	|	|	|	|	|	
0	1	2	3	4	5	6	7	8	9	10	...	ℕ

Van de positieve oneven getallen naar ℕ en terug

Zoals we zien, kun je vanuit de oneven getallen bij alle natuurlijke getallen terechtkomen en vice versa. De betreffende één-op-één correspondentie F van $\mathbb{N}_{oneven}$ naar ℕ wordt gegeven door F(x) = (x - 1)/2. Immers, voor elk natuurlijk oneven getal krijgen we precies één natuurlijk getal terug, wel specifiek het natuurlijke getal zoals hierboven gekoppeld aan het oneven natuurlijke getal. Omgekeerd is de functie G van ℕ naar $\mathbb{N}_{oneven}$, gedefinieerd door G(y) = 2y + 1, een één-op-één correspondentie tussen ℕ en $\mathbb{N}_{oneven}$. Als we een natuurlijk getal invoeren krijgen we daarvan een natuurlijk oneven getal, precies gekoppeld zoals hierboven. De verzameling van de oneven natuurlijke getallen en de verzameling van alle natuurlijke getallen is dus even groot.

Ik zal niet bewijzen[1] dat de verzameling ℝ groter is dan de verzameling ℕ. Dat ℝ groter is dan ℕ is door Cantor in 1873 ontdekt (Ferreirós, 2012) en dat deed hem redelijk snel hierna de continuümhypothese formuleren. Immers, nu dat bekend is dat ℝ groter is dan ℕ, kunnen we ons afvragen of er oneindige verzamelingen zijn die groter zijn dan ℕ, maar kleiner dan ℝ. Gödel verwoordt de vraag anders, namelijk: 'Hoeveel punten liggen er op een lijn?' (Gödel, 1995b, p. 256). Het aantal punten op een lijn is gelijk aan het aantal elementen van ℝ en we weten dat dit aantal groter is dan het aantal elementen van ℕ. We weten ook dat de laagste graad van oneindigheid die van ℕ is. Als we dus de graden van oneindigheid zouden tellen, vertaalt Gödels vraag zich naar: noem de graad van oneindigheid van ℕ 0, is die van ℝ dan 1?

3.2. De onbewijsbaarheid van de continuümhypothese

Toen Gödel de eerste versie van zijn artikel schreef, was er al geruime tijd gewerkt aan het oplossen van de continuümhypothese. In zijn artikel beschrijft Gödel hoe er nog niks bereikt is ten aanzien van de continuümhypothese en dat er wellicht ook geen oplossing binnen het huidige axiomasysteem van de verzamelingenleer gevonden kan worden. Later bewezen hijzelf en Cohen (Brown, 1999/2008, p.182) dat de continuüm-hypothese inderdaad onafhankelijk is van de huidige axioma's. Voor de verzamelingenleer (waar we de wiskunde naar kunnen terugbrengen) was er een axiomasysteem, ZF(C), opgericht, waar de wiskundige gemeenschap het redelijk eens over heeft kunnen worden. De continuümhypothese blijkt echter onafhankelijk te zijn van de axioma's van ZF(C). Dat wil zeggen, noch de waarheid noch de onwaarheid van de continuümhypothese kan bewezen worden op basis van dit axiomasysteem. Dit terwijl, zoals Gödel in zijn artikel uitlegt, er geen enkele reden is om aan te nemen dat we te maken hebben met een slecht geformuleerde vraag:

> So the analysis of the phrase 'how many' unambiguously leads to a definite meaning for the question stated in the second line of this paper. (Gödel, 1995b, p. 256)

In zijn paper (geschreven toen er nog geen zekerheid bestond over de onafhankelijkheid van de continuümhypothese) stelt Gödel voor om op zoek te gaan naar nieuwe axioma's die ons in staat stellen de continuümhypothese te beslissen. Aangezien een correct gestelde wiskundige vraag, zoals de vraag naar het aantal punten op een lijn, nu eenmaal om een antwoord vraagt, hebben we bij het opstellen van de axioma's van de verzamelingenleer ergens een axioma overgeslagen waarmee we deze vraag kunnen beantwoorden. Immers, zonder dit extra axioma kan er nooit een antwoord worden geformuleerd. Het is als doorsteek op deze opmerkingen dat Gödels argument voor wiskundige intuïtie gezien kan worden.

3.3 Wiskundige intuïtie en de zoektocht naar axioma's

In zijn artikel eindigt Gödel met een bijvoegsel, waarin we een uiteenzetting zien van wiskundige intuïtie. Hij zoekt hierbij aansluiting op de rest van zijn artikel. Wiskundige intuïtie is volgens Gödel de enige manier om te verklaren hoe het kan dat de continuümhypothese onafhankelijk is van onze axioma's maar dat we er uiteindelijk toch een antwoord van mogen verwachten. Gödel karakteriseert de zoektocht naar nieuwe axioma's als volgt.

> Hence its [de continuumhypothese, HH] undecidability from the axioms being assumed today can only mean that these axioms do not contain a complete description of that [the mathematical] reality. (Gödel, 1995b, p. 260)

> However, the question of the objective existence of the objects of mathematical intuition (which, incidentally, is an exact replica of the question of the objective existence of the outer world) is not decisive for the problem under discussion here. The mere psychological fact of the existence of an intuition which is sufficiently clear to produce the axioms of set theory and an open series of extensions of them suffices to give meaning to the question of the truth or falsity of propositions like Cantor's continuum hypothesis. (Gödel, 1995b, p. 268)

De manier om te verklaren dat we in staat zijn tot het vergaren van nieuwe axioma's is dus wiskundige intuïtie. Hoe anders, stelt Gödel, kunnen we incomplete kennis van de wiskunde hebben? Vergelijk dit bijvoorbeeld met de positie van de logicist en met de aan zijn positie gelieerde wiskundige praktijk. Logicisme is de theorie dat de wiskunde terug valt te brengen tot enkel logische wetten (Horsten, 2015) (ik zal verderop laten zien dat deze positie zeer problematisch is, maar voor nu kijk ik naar de effecten op de wiskundige praktijk van logicisme). Het is onduidelijk hoe een logicist omgaat met het gegeven dat de continuümhypothese onbewijsbaar is. Immers, als bewezen is dat de continuümhypothese onafhankelijk is van ZF(C) dan zal de logicist eenvoudigweg moeten accepteren dat de continuümhypothese nooit kan worden beslist. Iets breder getrokken zal de bezigheid van het vinden van axioma's voor een logicist sowieso een onzinnige bezigheid zijn. Immers, de enige axioma's die een logicist in zijn theorie toelaat zijn die van de logica. Daarmee is dus al een compleet beeld van de axioma's gegeven.

Er zijn twee manieren waarop het argument van Gödel ingezet kan worden ten voordele van wiskundige intuïtie. Ten eerste is het een direct argument. Dit wil zeggen dat als we ons willen aansluiten bij het idee van het blijven zoeken van nieuwe axioma's, we ofwel wiskundige intuïtie moeten aannemen, ofwel een nieuw(e) theorie/tegenargument naar voren moeten schuiven. De zoektocht naar nieuwe axioma's is echter ook op een andere manier een argument. Niet alleen maakt de zoektocht naar nieuwe axioma's wiskundige intuïtie plausibel, het maakt wiskundige intuïtie ook houdbaar. Chihara's argument richt zich op het gebrek aan concrete eigenschappen en effecten die toegeschreven kunnen worden aan wiskundige intuïtie. We hebben in dit hoofdstuk echter gezien dat wiskundige intuïtie wel degelijk geconcretiseerd kan worden. Naast dat het een oplossing is voor een ontologisch vraagstuk, kan wiskundige intuïtie een argumentatieve rol spelen bij de keuze tussen bepaalde wiskundige praktijken. De vraag of we moeten zoeken naar nieuwe axioma's voor de verzamelingenleer is afhankelijk van ons antwoord op de vraag hoe we wiskundige kennis vergaren en is daarmee afhankelijk van hoe we ons verhouden tot Gödels wiskundige intuïtie.

§4 De eerste onvolledigheidsstelling en wiskundige kennis

De door Gödel ontdekte eerste onvolledigheidsstelling (een stelling over de onvolledigheid van de getaltheorie die we weldra uiteen zullen zetten) geeft iemand die wiskundige intuïtie wil verdedigen houvast. Uit de eerste onvolledigheidsstelling kan worden afgeleid dat wiskunde een zeer bijzondere/afwijkende vorm van kennis is. Dit helpt de gödeliaan omdat het argument van Benacerraf op de veronderstelling leunt dat van wiskundige kennis mag worden verwacht dat het op eenzelfde manier kan worden verklaard als andere kennis. Als wiskundige kennis aantoonbaar afwijkend is, is deze veronderstelling niet langer gerechtvaardigd.

4.1 Gödels onvolledigheidsstelling

Op basis van de uitleg van H.C.M. De Swart (1993, pp. 465-475) over de onvolledigheidsstellingen zal ik beknopt een overzicht geven van de conclusies van dit theorema. Ik zal mij niet bezighouden met het bewijs dat Gödel hiervoor formuleert, maar enkel naar de conclusies en mogelijke gevolgen voor de aard van wiskundige kennis kijken.

De eerste onvolledigheidsstelling stelt dat voor elk formeel systeem van de getaltheorie dat aan bepaalde minimale eigenschappen voldoet, geldt dat er ware uitspraken zijn die niet te bewijzen vallen binnen dat formele systeem (De Swart, 1993, p. 465). Dit wil zeggen dat indien we een sterk genoeg formeel systeem maken waarin we de getaltheorie modelleren, we voor dit systeem altijd een waarheid kunnen vinden die niet te bewijzen valt binnen dat systeem. Voor elk systeem dat ook maar de kans wil hebben volledig te zijn aangaande de waarheden van de getaltheorie, geldt dat het aan bepaalde eigenschappen moet voldoen (Smith, 2007/2013, pp. 49-52). Dit wil op zijn beurt zeggen dat formele systemen van de getaltheorie altijd onvolledig zullen zijn. Oftewel, het zal nooit mogelijk zijn alle waarheden van de getaltheorie binnen één formeel systeem te bewijzen.

4.2 Hume en Kant: wiskunde als een bijzondere soort kennis

In *An enquiry concerning human understanding* beschrijft Hume dat er twee vormen van kennis kunnen zijn, namelijk *matters of fact en relations of ideas* (2008, p. 18). *Matters of fact* zijn afkomstig van de zintuigen of het geheugen (ofwel, vanuit *impressions* of door *impressions* gevormde *ideas*) en de *matters of fact* zijn de verschillende dingen die we weten door middel van afleiding uit de indrukken die wij hebben. Dit is de eerste oorsprong van kennis en deze kennis voegt informatie toe. Naarmate ik meer *impressions* heb, ontrafel ik meer *matters of fact*. Hiertegenover staan de *relations of ideas*. De *relations of ideas* zijn kennis die ik op doe door waarheden te vinden die per definitie waar zijn. Als ik een *idea* van een eenhoorn heb, dan kan ik daar de ware uitspraak over doen, op basis van *relations of ideas*, dat deze eenhoorn wit of niet wit is.

Als we nu een paar decennia vooruit springen zien we dat Kant in de *Kritiek van de Zuivere Rede* een vergelijkbare onderverdeling maakt van de kennis. Ten eerste is er synthetische a-posteriori kennis. Dit is kennis die wordt opgedaan door middel van de zintuigen (a-posteriori) en die informatie toevoegt (synthetisch). Ten tweede is er a-priori analytische kennis. Dit is kennis die geen informatie toevoegt over de wereld (analytisch) en kenbaar is zonder het gebruik van de zintuigen (a-priori) (Kant, 2004, pp. 93-104). In tegenstelling tot Hume stelt Kant dat er naast deze twee nog een andere vorm van kennis is, namelijk de synthetische a-priori kennis. Deze kennis voegt wel informatie toe maar vereist niet het gebruik van de zintuigen om te worden verworven.

Deze soort van kennis is alles behalve standaard, want de door Hume gegeven onderverdeling lijkt compleet. Hoe kan het dat er kennis is die mij iets vertelt over de wereld zonder dat deze verkregen is met behulp van de zintuigen? Uitgerekend wiskunde is het voorbeeld dat Kant geeft van een dergelijke, bijzondere soort van kennis. Na duidelijk te hebben gemaakt dat de zintuigen niet nodig zijn voor het vergaren van wiskundige kennis (wiskunde is dus a-priori) zegt hij:

> Alle wiskundige oordelen zijn synthetisch. Deze stelling lijkt de aandacht van de analytici van de menselijke rede tot dusver te zijn ontgaan, en zelfs geheel tegengesteld aan al hun vermoedens te zijn, hoewel ze onbetwistbaar is en de gevolgen ervan zeer belangrijk zijn. Want omdat men van mening was dat alle gevolgtrekkingen van de wiskundigen verlopen volgens de wet van de tegenspraak (dat vereist de aard van alle apodictische zekerheid), kwam men tot de overtuiging dat ook de grondbeginselen van de wiskunde volgens de wet van de tegenspraak wordt gekend. (Kant, 2004, pp. 104-105)

Ik zal niet ingaan op de argumenten van Kant betreffende het synthetisch zijn van de wiskunde. Wel zal ik laten zien dat we een argument hiervoor kunnen formuleren op basis van de zojuist geformuleerde onvolledigheidsstelling. Belangrijk is om hier te beseffen wat het synthetisch-zijn van kennis betekent. Zoals Kant zegt is de wiskunde niet analytisch omdat ze niet gebaseerd is op de wet van de non-contradictie. Anders geformuleerd: wiskundige waarheden zijn geen logische waarheden, dat wil zeggen, het zijn geen waarheden die per definitie waar zijn. Ook belangrijk is om stil te staan bij wat voor een bijzondere vorm van kennis dit is. Het betreft hier namelijk kennis die informatie oplevert over de wereld zonder dat er enige *impression* vanuit de wereld nodig is om deze te genereren.

4.3 Synthetische wiskunde en de onvolledigheidsstelling

We hebben gezien dat formele systemen niet in staat zijn om het geheel van de getaltheorie te bewijzen. Bij elk formeel systeem is er een waarheid van de getaltheorie die in dat formele systeem onbewijsbaar is. Dit maakt duidelijk dat wiskunde niet analytisch kan zijn in de zin die zojuist is beschreven. Immers, logische waarheden (waarheden die per definitie waar zijn) kunnen wel degelijk in een formeel systeem worden beschreven, namelijk in een formeel systeem van de logica. Omdat Kant het begrip analytisch niet op formele wijze definieert, is dit wellicht niet direct duidelijk. Bij De Swart vinden we de volgende alternatieve definitie van analytisch:

> In more modern terminology, following roughly a 'Fregean' account of analyticity, one would define a proposition A to be analytic iff either
> (i) A is an instance of a logically valid formula; e.g. 'No unmarried man is married' has the logical form -∃ x(-P(x) & P(x)), which is a valid formula, or
> (ii) A is reducible to an instance of a logically valid formula by substitution of synonyms for synonyms; e.g., 'No bachelor is married'. (De Swart, 1993, p. 359)

De Kantiaanse definitie is niet identiek maar heeft *prima facie* wel dezelfde extensie als de Fregeaanse definitie zoals gegeven door De Swart. Analytische waarheden zijn waarheden waarbij enkel naar de logische vorm van de claim gekeken hoeft te worden. Het voordeel van deze moderne formulering is dat we direct kunnen inzien dat de eerste onvolledigheidsstelling het onmogelijk maakt dat getaltheorie analytisch is. Want als een waarheid, of een verzameling waarheden, analytisch is, dan moet deze te reduceren zijn tot logische waarheden. Als deze waarheden logische waarheden zijn, dan kan er een formeel systeem worden geleverd voor deze waarheden. Sterker nog, de volledigheidsstelling , die stelt dat alle geldige uitspraken ook bewijsbaar zijn in dat formele systeem geldt voor de predicaten en propositielogica (De Swart, 1993, pp. 410-411). Als voor de logica de volledigheidsstelling geldt en de wiskunde te reduceren zou zijn tot logica, dan is het dus zo dat er een formeel systeem van de getaltheorie bestaat dat niet onvolledig is. We hebben echter al met behulp van de eerste on-volledigheidsstelling geconcludeerd dat een dergelijk formeel systeem niet bestaat. De wiskunde is dus niet te reduceren tot de logica. Dit betekent dat de wiskunde niet analytisch maar synthetisch is.

Er is ook verdedigd dat wiskundige kennis a-posteriori zou zijn. Alhoewel ik hier niet verder op in zal gaan, is het interessant om te weten dat Gödel deze positie als absurd beschouwt en verwerpt (Gödel, 1995a,

p. 312). Er zijn tevens voor de hand liggende argumenten die de positie problematisch maken. Het is bijvoorbeeld niet nodig een zintuig aan te wenden om te weten dat één plus één gelijk is aan twee. Daarnaast doen de perfect vormen van de meetkunde (zoals de cirkel) zich nooit aan onze zintuigen voor.

4.4 Benacerrafs argument geproblematiseerd

Zoals we hebben gezien kan op basis van de eerste onvolledigheidsstelling en de onderverdeling van soorten kennis zoals gevonden bij Kant en Hume worden geconcludeerd dat wiskunde een zeer bijzondere vorm van kennis is. Als we nu terugkijken naar het argument dat wordt gegeven door Benacerraf, dan zijn er een aantal vooronderstellingen die we kunnen aanvechten:

> i) Wiskundige kennis moet passen binnen onze verdere, algemene opvattingen over kennis.
>
> ii) Ons idee over kennis leunt op de causale interactie die wij met de objecten van onze kennis hebben. Als er geen sprake kan zijn van causale interactie dan past kennis over de objecten niet binnen ons algemene beeld over kennis.
>
> iii) Als we wiskundige intuïtie gebruiken ter verklaring van wiskundige kennis, dan is deze kennis niet gebaseerd op causaal effectieve objecten.

We hebben gezien dat iii) correct is, omdat wiskundige objecten binnen Gödels verklaring van wiskundige kennis geen causale effecten hebben. Er zijn kritieken uit meerdere hoeken geweest op punt ii). Waar het echter om gaat is dat op basis van wat we net hebben ontdekt over wiskundige kennis i) in een nieuw, kwaad, daglicht komt te staan.

Dat wiskunde synthetische a-priori kennis is, maakt duidelijk dat wiskunde een bijzondere vorm van kennis is. Op het eerste gezicht zijn we wellicht geneigd om een filosoof gelijk te geven als hij kennis onderverdeelt in kennis die per definitie waar is en kennis waarvoor we indrukken vanuit de wereld moeten verkrijgen. Ik heb echter aan de hand van de onvolledigheidsstelling laten zien dat er goede redenen zijn om aan te nemen dat wiskunde buiten dit intuïtieve beeld over kennis valt.

Wiskunde is een bijzondere vorm van kennis. We kunnen vooral zeggen wat wiskundige kennis niet is. Bijvoorbeeld, ze is niet zoals kennis van de logica, en niet zoals kennis van empirisch objecten. Aangezien wiskunde een bijzondere vorm van kennis is, is er geen enkele reden om aan te nemen dat wiskundige kennis past binnen ons algemene beeld van kennis. Dit betekent dat één van de belangrijke veronderstellingen die Benacerraf maakt, namelijk i), op losse schroeven staat. Benacerraf slaagt er zodoende niet in om een vernietigende kritiek te formuleren die de onhoudbaarheid van wiskundige intuïtie aan zou tonen.

§5 Wiskundige afbeeldingen; illustraties of bewijzen?

In het artikel *Proofs and Pictures* (1997) van James Robert Brown treffen we een verdediging aan van het gebruik van wiskundige afbeeldingen en illustraties als bewijzen voor wiskundige stellingen, naast de formele bewijzen die moderne wiskundigen gewend zijn te gebruiken. Hier koppelt hij een epistemologie van de wiskunde aan, die sterk overeen komt met Gödels wiskundige intuïtie. Dit helpt de gödeliaan op twee manieren. Ten eerste kan deze koppeling tussen wiskundige intuïtie en de wiskundige praktijk van het gebruik van afbeeldingen als bewijzen, een direct argument zijn voor wiskundige intuïtie als we het eens zijn met Brown over de functie van afbeeldingen binnen de wiskunde. Het is tevens een argument voor de houdbaarheid van wiskundige intuïtie omdat deze koppeling tegen Chihara's argument ingaat. Er zijn concrete effecten van de Gödels epistemologie te geven, namelijk hoe we ons willen verhouden tot de praktijk van afbeeldingen als bewijzen. Dit laat zien dat Chihara's kritiek dat er geen concretiseringen bestaan van gödeliaanse epistemologie, niet geldig is.

5.1 Overtuigende afbeeldingen

Ik begin met twee voorbeelden die Brown geeft van afbeeldingen die als bewijs kunnen fungeren. Bij het eerste voorbeeld is ook het formele bewijs erbij genomen.

Theorem: $1 + 3 + 5 + \ldots + (2n - 1) = n^2$
[Picture] Proof:

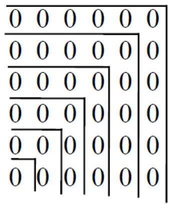

Figure 3

This picture proof should be contrasted with a traditional proof by mathematical induction which would run as follows:

Proof (traditional): We must show that the formula of the theorem holds for 1 (the basis step), and also that, if it holds for n then it also holds for n+ 1 (the inductive step).
Basis: $((2 * 1) - 1) = 1^2$
Inductive: Suppose $1 + 3 + 5 + \ldots + (2n - 1) = n^2$ holds as far as n. Now we add the next term in the series, $2(n + 1) - 1$, to each side:

$$1 + 3 + 5 + \ldots + (2n - 1) + 2(n + 1) - 1 = n^2 + 2(n + 1) - 1$$

Simplifying the right hand side, we get:

$$n^2 + 2(n + 1) - 1 = n^2 + 2n + 2 - 1 = n^2 + 2n + 1 = (n + 1)^2$$

This last term has exactly the form we want. And so the theorem is proven. (Brown, 1997, pp. 169-170)

Het eerste bewijs is geen formeel bewijs. Toch zal menig lezer duidelijk zijn dat het theorema waar is op basis van de afbeelding. Door het blijven plaatsen van nieuwe rijen en kolommen, kan elke keer weer een vierkant gevormd worden waarvan de oppervlakte gelijk is aan de gegeven som.

Het tweede voorbeeld dat Brown geeft is een continue functie (eenvoudig gezegd een functie die je zou kunnen tekenen zonder je potlood van het papier te hoeven halen) die, als hij op enig moment onder en op een ander moment boven de nullijn ligt, dan ook op enig punt de nullijn zal raken. Het formele bewijs hiervoor is uitgebreid (Brown, 1997, pp. 162-163) en zal ik niet geven. De volgende afbeelding kan echter als picture proof dienen.

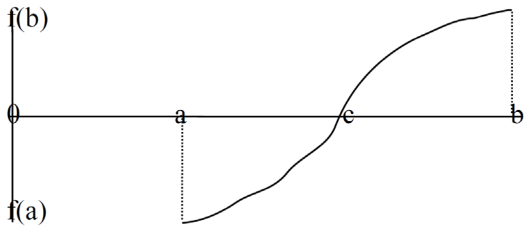

Fig 1. The intermediate zero theorem (Brown, 1997, p. 163)

5.2 Afbeeldingen als bewijzen

Na deze voorbeelden vraagt Brown zich af wat de uitvinding van het formele bewijs in het algemeen heeft opgeleverd. Hij geeft hiervoor drie opties:

1) Het bewijs bewijst een theorema waarvan we daarvoor niet wisten of deze waar was. Brown geeft aan dat we deze optie moeten verwerpen. Iedereen die immers de bovenstaande afbeelding ziet, weet dat de beschreven som gelijk is aan n^2 en dat de functie de nullijn raakt (Brown, 1997, p. 164).

2) Het bewijs legt uit wat het theorema precies betekent en waarom deze waar is. Als we echter naar de afbeelding kijken kunnen we ook zien waarom de som gelijk is aan n^2. Het formele bewijs legt inderdaad uit waarom het theorema waar is, maar het visuele bewijs doet dat eveneens.

3) Het theorema waar we op uit komen, geldt als argument voor de premissen waar het bewijs gebruik van maakt. Het theorema was al gekend, het enige wat dit bewijs bewees is dat de premissen van het bewijs correct waren, aangezien ze op een gekende waarheid uitkomen.

Brown beargumenteert dat optie drie als volgt de beste optie is: stel je voor dat er een bewijs bestaat waaruit blijkt; als een continue lijn een punt heeft boven de nul en een punt onder de nul, dan raakt de lijn nergens de nullijn. Dit zou waarschijnlijk niet ons geloof beschadigen dat continue lijnen die op enig punt boven de nullijn en op een ander punt onder de nullijn liggen, de nullijn raken. Waarschijnlijk zouden we geneigd zijn om te kijken naar wat er mis is met ons bewijs. Brown kiest dus voor de derde optie:

> It is pretty clear that, of our three options, the final one is the best. (The second option, explanation, is compatible with the third, confirmation, but seems much less plausible.) The consequence of adopting (III) is highly significant for our view of pictures. We can draw the moral quickly: on this view pictures are crucial. They provide the known to be true consequences that we use for testing the hypothesis of arithmetization. Trying to get along without them would be like trying to do theoretical physics without the benefit of experiments to test conjectures. (Brown, 1997, p. 165)

Brown verwerpt een mogelijke tegenwerping die stelt dat er nog een vierde optie is. Deze luidt dat we met twee verschillende bewijsmethodes hetzelfde theorema hebben bewezen. Dit is echter niets meer dan een herformulering is van Browns positie, immers:

> I could rephrase in as saying: the two proof-techniques arrive at the same result. One of these (the picture) is prima facie reliable. The other (the analytic proof) is questionable, but our confidence in it is greatly enhanced by the fact that it agrees with the reliable method. (Brown, 1997, p. 166)

Voor de verbinding met en de onderbouwing van wiskundige intuïtie is het echter niet noodzakelijk dat Browns argumenten worden geaccepteerd. Wat belangrijk is, is dat er een onbesliste discussie aangaande wiskundige praktijk is, waarvan we zullen zien dat zij verband houdt met wiskundige intuïtie.

5.4 De relatie tot wiskundige intuïtie

Dat afbeeldingen in de wiskunde als bewijs zouden kunnen dienen, is op het eerste gezicht niet verbonden met wiskundige intuïtie. Brown vraagt zich op enig moment echter af hoe het kan dat wiskundige afbeeldingen fungeren als bewijs. Hij zegt hierover:

> Consequently, I claim, some 'pictures' are not really pictures, but rather are windows to Plato's heaven. (Brown, 1997, p. 174)

> As telescopes help the unaided eye, so some diagrams are instruments (rather than representations) which help the unaided mind's eye. (Brown, 1997, p. 174)

Het gebruik van afbeeldingen als bewijzen sluit uitstekend aan bij het verkrijgen van wiskundige kennis met behulp van ons 'geestesoog', die op een imperfecte manier de platoonse wiskundige objecten aanschouwt. Dit sluit op zijn beurt goed aan bij hoe ik tot nog toe Gödels wiskundige intuïtie heb begrepen. Het lijkt erop dat het gebruik van afbeeldingen als bewijsmethode in de wiskunde alleen kan worden begrepen als wiskundige intuïtie die een scherp beeld krijgt van de objecten die ze probeert te schouwen. Brown voelt zich genoodzaakt tot het gebruik van dergelijke metaforen om de werking van de afbeeldingen als bewijzen te beschrijven en verklaren.

Dit is op twee manieren gunstig voor de houdbaarheid van wiskundige intuïtie. Ten eerste is er een direct argument. Als iemand ervan overtuigd is dat afbeeldingen als wiskundig bewijs kunnen dienen, en het klopt dat dit enkel mogelijk is als we het bestaan van wiskundige intuïtie onderkennen, dan heeft deze persoon een goede reden om wiskundige intuïtie aan te nemen. Een tweede, indirect argument voor de voorstander van wiskundige intuïtie is dat de vraag naar wiskundige intuïtie niet enkel een vraag naar de metafysische fundamenten van de wiskunde kan oplossen. Deze verdenking wordt dankzij Browns artikel weggenomen. Wiskundige intuïtie is een kwestie die niet enkel verband houdt met de ontologie van de wiskunde. Ze houdt ook verband met de geldigheid en waardering van specifieke bewijsvormen binnen de wiskunde. Wiskundige intuïtie blijkt een onderwerp te zijn dat gerelateerd is aan de wiskundige praktijk. De vraag naar wiskundige intuïtie bevindt zich in het goede gezelschap van de vraag naar de geldigheid van afbeeldingen als bewijs.

Conclusie

Ik heb laten zien dat Gödels wiskundige intuïtie een sterker en houdbaarder begrip is dan op het eerste gezicht lijkt. Volgens Chihara roept wiskundige intuïtie herinneringen op aan mystieke ervaringen en lijkt het zonder concrete gevolgen enkel als oplossing te dienen voor één specifiek probleem. Benacerraf merkt op dat wiskundige intuïtie niet strookt met de manier waarop wij gewend zijn kennis te bezien en daarmee lijkt wiskundige intuïtie contra-intuïtief. De gödeliaan heeft de middelen om deze aanvallen te weren. De kwestie aangaande wat we moeten verwachten van concrete wiskundige problemen zoals de continuümhypothese, blijkt afhankelijk te zijn van ons antwoord op de vraag of wiskundige intuïtie bestaat. De eerste onvolledigheidsstelling geeft ons aanleiding om te veronderstellen dat wiskunde een andere vorm van kennis is dan onze alledaagse kennis. Ten slotte vinden we in de vraag naar het gebruik van wiskundige afbeeldingen niet alleen een direct argument voor wiskundige intuïtie, maar zien we tevens de praktische consequenties die een dergelijke theorie zou hebben. Aan de hand van deze argumenten heb ik laten zien dat wiskundige intuïtie houdbaar is.

Hoewel de gödeliaan zeker nog op onduidelijkheid en gebrek aan direct bewijs kan worden aangevallen, kan hem niet worden verweten een te exotische, onhoudbare theorie aan te hangen.

Dankbetuigingen

Graag wil ik hierbij prof. dr. Harrie de Swart en dr. Paul Schuurman bedanken voor de begeleiding tijdens het schrijven van de oorspronkelijke bachelorthese.

Hugo Hogenbirk (1992) studeerde Wijsbegeerte aan de Erasmus Universiteit Rotterdam en studeert Informatica aan de Universiteit Utrecht. Momenteel verricht hij onderzoek binnen de filosofie van spel. Zijn wetenschappelijke interessegebieden zijn de epistemologie, de metafysica en de filosofie van de wiskunde.

'De houdbaarheid van Kurt Gödels wiskundige intuïtie' was oorspronkelijk geschreven als bachelorscriptie aan de Erasmus Universiteit Rotterdam voor de studie Wijsbegeerte, onder begeleiding van prof. dr. Harrie de Swart en dr. Paul Schuurman.

Noten

1. Zie hiervoor bijvoorbeeld De Swart, 1993, pp. 258-276.

Literatuur

Benacerraf, P. (1973). Mathematical Truth. The Journal of Philosophy, 70(19), 661-679.

Brown, J. (1997). Proofs and pictures. The British Journal for the Philosophy of Science, 48, 161-180.

(1999/2008). Philosophy of Mathematics. A Contemporary Introduction to the World of Proofs and Pictures (2nd ed.). New York, NY: Routledge.

Burton, D. (2011). The History of Mathematics An Introduction (7th ed.). New York, NY: The Macgraw-Hill.

Chihara, C. (1982). A Godelian Thesis Regarding Mathematical Objects: Do They Exist? and Can We Perceive Them? The Philosophical Review, 91(2), 211-227.

De Swart, H. (1989). Filosofie van de wiskunde. Leiden: Uitgeverij Martinus Nijhoff.

(1993). Logic: Mathematics, language, computer science, and philosophy, i-577. Bern: Peter Lang.

Ferreirós, J. (2012). The Early Development of Set Theory. In E. Zalta (Ed.) The Stanford Encyclopedia of Philosophy (Winter 2012 Edition). Retrieved from http://plato.stanford.edu/archives/win2012/entries/settheory-early/.

Gödel, K. (1995a). Some basic theorems on the foundations of mathematics and their implications. In S. Feferman (Ed.), Collected works (3), 304-323. New York: Oxford University press.

(1995b). What is Cantor's continuum problem?. In S. Feferman (Ed.), Collected works (2), 254-270. New York: Oxford University press.

Horsten, L. (2015) Philosophy of Mathematics. In E. Zalta (Ed.) The Stanford Encyclopedia of Philosophy (Spring 2015 Edition). Retrieved from http://plato.stanford.edu/archives/spr2015/entries/philosophy-mathematics/.

Hume, D. (2008). An Enquiry concerning Human Understanding (2nd ed.). Oxford: Oxford University press.

Kant, I. (2004). Kritiek van de zuivere rede. Amsterdam: Uitgeverij Boom.

Linnebo, Ø. (2013a). Platonism in the Philosophy of Mathematics Supplement. In E. Zalta (Ed.) The Stanford Encyclopedia of Philosophy (Winter 2013 Edition). Retrieved from http://plato.stanford.edu/archives/win2013/entries/platonism-mathematics/supplement.html.

(2013b). Platonism in the Philosophy of Mathematics. In E. Zalta (Ed.) The Stanford Encyclopedia of Philosophy (Winter 2013 Edition). Retrieved from http://plato.stanford.edu/archives/win2013/entries/platonism-mathematics/.

Maddy, P. (1989). The Roots of Contemporary Platonism. The Journal of Symbolic Logic, 54(4), 1121-1144.

Smith, P. (2007/2013). An introduction to Gödel's Theorems (2nd ed.). Cambridge: Cambridge University Press.

Parsons, C. (1995). Platonism and Mathematical Intuition in Kurt Gödel's Thought. The Bulletin of Symbolic Logic, 1(1), 44-74.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit http://creativecommons.org/licenses/by-nc/3.0/

ESJP
#9 | 2015

Economics, Complexity and the Disenchantment of the Social World

Sam van Dijck

The fact that we predict eclipses does not, therefore, provide a valid reason for expecting that we can predict revolutions. (Popper, 2002, p. 340)

Introduction

Even though he is widely considered to be the founding father of the economic discipline, Adam Smith would have a hard time finding a job at an economics department or getting his ideas published in any of the major economics journals, had he lived today. One of the central reasons for this is that neoclassical economics, which dominates the discipline today, and economics at the time of the great political economists such as Adam Smith and David Ricardo are differentiated by one single characteristic more than any other: their use of mathematics[1] (Schabas, 1989). While today's economics is best characterized as a thoroughly mathematical science, the writings of the classical economists were almost entirely discursive (Lawson, 2012; Hodgson, 2013).

This mathematical condition of modern economics has quite recently become the subject of heated debate and strong criticism in light of the economic crisis that has hit us in 2008 and still lingers on today. Many (i.e., Friedman, 1999, p. 137; Krugman, 2009a, 2009b) have argued that, caught up in more and more complex models, economics itself had become detached from its appropriate subject matter: real world economic problems. The economic science failed to make sense of our reality, and instead got lost in a different reality of their own making consisting of models and equations.

Such criticisms regarding the role of mathematics in economics and its inability to capture our economic reality are not, however, just something of the past seven years. Even though the role of mathematics has evolved to one of absolute dominance since the end of the 19th century, many have voiced criticisms towards this development. And the list of those critical of the mathematization of economics does not just name quirky heterodox economists at the margins of the discipline but also includes some of the most famous and important economists of the 19th and 20th century.

In this paper I will take a closer look at some of the concerns and warnings about the role of mathematics in economics put forward by Alfred Marshall, Friedrich Hayek and John Maynard Keynes. Specifically these economists have been selected because each of them has had a significant and constituting influence on the economic discipline, and because taken together they represent a substantial part of the diversity of the economic discipline at their time and today (i.e. Keynes' argument for the occasional government intervention versus Hayek's laissez-faire economy).[2] This paper will focus on the concerns they voiced regarding mathematics in economics, which were born out of their shared conviction of the complexity of economic reality. They argue that the world is too complex and varied for mathematics to be able to capture it, and that this thus poses limits to its use. They do not deny that mathematics can be useful but one must know its place and restrictions. In this respect they stand in sharp contrast to economic thinkers such as William Stanley Jevons, Irving Fisher, Paul Samuelson, Kenneth Arrow and Gerard Debreu, some of the founders of neoclassical mathematical economics, who believed that it was in fact possible to capture economic reality in mathematics and that it should therefore be adopted as the main engine of enquiry in economic science.

After discussing concerns regarding the role of mathematics in economics I will continue by addressing the question of why, seeing that both sides had respectable and important economists in their ranks, mathematics evolved to the dominant position it has today *in spite of* the concerns that were expressed. Why were these concerns put aside? As an answer to this question I argue that there existed and exists in economics a high demand for predictive power, which can be illustrated by Milton Friedman's famous and influential statement that the performance of a theory is to be judged only by its ability to predict (Friedman, 1966, p. 4). What was demanded of an economic theory was that it provided the ability to predict events in the social world. It is my claim that this demand for predictive power is and was the result of the attempt to bring the world under the mastery of "man": a process that Max Weber called "disenchantment" (Weber, 1958). I then argue that the reason the concerns voiced in the direction of mathematics were not able to stop its triumphal march to dominance is that these concerns were founded on ideas of real world complexity which were incompatible with this process of disenchantment and the demand for prediction that it entailed.

Economics, mathematics and the complexity of economic reality

Let me begin by discussing an opposition between David Ricardo and Thomas Robert Malthus. Though Ricardo is one of those classical economists whose work is almost entirely discursive and who did not specifically advocate the use of mathematics, he did uphold the view that (simple) models could somehow be representative of a wide set of varied and complex phenomena (Ricardo, 1817). In this aspect he is the opposite of one of his fellow classical economists, Thomas Robert Malthus, who held the view that economic reality was far too complex and varied to be captured in models. Malthus therefore concluded that simple formal models were at most of highly limited use (Malthus, 1820).

I introduce this contradistinction here because it is to some extent illustrative of the arguments and concerns that I will address in what follows. I do not mean it to refer to a distinction between inductivism and deductivism, and I do not argue that there exists a one to one relation between an inductivist position and positions preaching caution regarding mathematics in economics nor between the deductivist position and those promoting math (Hodgson, 2013). Rather, the distinction between Malthus and Ricardo is to be illustrative of two opposing positions regarding the possibility of capturing economic reality in mathematics that will be addressed in what follows.

So on the one hand there is the 'Ricardian' and so far victorious side of the story that believed that it is in fact possible to, putting it crudely, capture reality in a model. An important figure on the list of those subscribing to this line of thought is William Stanley Jevons, a famous 19th century economist who, by introducing his theory of marginal utility, was one of the first to provide an anchor for calculus and mechanical analogies in economics (Schabas, 1989; Hodgson, 2013). By arguing that it was possible to assign a number to the utility (pain or pleasure) that a person receives from a certain outcome, it became possible to introduce utility in mathematical economic models. And with the idea that utility was something that existed in reality, or was at least an analogy for something that existed in reality, came the idea that it was possible to capture economic reality in mathematics and mechanical analogies and the conviction that studying these would provide us with information about real world economic phenomena (Schabas, 1989).

Jevons' framework and ideas about mathematics being the primary engine of enquiry for economics were developed further in the late 19th and early 20th century by several important economists (such as Irving Fisher). However, it was not until the late 1950's that these ideas about the formalization of economics found their culmination in the works of economists like Kenneth Arrow, Gerard Debreu and Paul Samuelson, who established the foundations for post-war mathematical economic theory (Blaug, 2002, 2003). That this line of economists believed that mathematics was truly able to capture economic reality is nicely illustrated by the following statement made by Paul Samuelson in one of his early papers at Harvard:

> Mathematical methods properly employed, far from making economic theory more abstract, actually serve as a powerful liberating device enabling the entertainment and analysis of ever more realistic and complicated hypothesis. (Samuelson, 1939)

So for the young Samuelson, as for many of his colleagues at the time, it was clear enough: by using mathematics we are able to make our hypotheses match more and more closely to the way things actually are. It was believed (and is still believed by many today) that the models that mathematical economics produces would be able to capture something useful about economic reality when applied to it (Lawson, 1997; Hodgson, 2013). And so it was thought of not only as possible, but also as better to use mathematics to express oneself in economics. For most economists, to state a theory in terms of a formal model is an unambiguous improvement (Chick, 2001).

But this idea of mathematics being able to capture reality and thus being the ultimate engine to enquire about, it is not something that was without criticism. There were quite some economists who, like Malthus, emphasized that economic reality was too complex to be captured in mathematics, and that models are thus at best of highly limited use. Let us first turn to Alfred Marshall in that respect.

Marshall saw the limits to highly general 'pure theory' of the type found in the works of Ricardo and Jevons and though they were produced quite some time after his death, he would surely have found the same limits in the type of pure theory found in the works of economists like Arrow, Debreu and Samuelson (Marchionatti, 2004; Hodgson, 2013). In his letter to Arthur Bowley, the LSE's first professor of statistics, he wrote:

> I had a growing feeling in the later years of my work at the subject that a good mathematical theorem dealing with economic hypotheses was very unlikely to be good economics. (Pigou, 1925, p. 427)

Though some have also emphasized Marshall's role in the rise of mathematical economics (e.g. Schabas, 1989) it seems indubitable that he himself did not believe in a purely mathematical economic science and continuously emphasized the danger of formalization causing economic theory to stray away from real-life relevance (Marchionatti, 2004).

This does not mean that Marshall opposed the use of mathematics altogether. He simply held that it should not be the economists' principal engine of enquiry. One can, and according to Marshall should, use mathematics as a shorthand language for thinking (Schabas, 1989). But afterwards, he argued, one should always translate back to real world examples and "burn the mathematics" (Pigou, 1925, p. 427).

Now why is it that mathematics should not, according to Marshall, be our primary engine of enquiry? Why should we burn the mathematics and use it only as a shorthand language for thought? The reason for this is precisely that he believed mathematics was unable to capture economic reality (Marchionatti, 2004). Marshall thought that the economic world was far too complex to be represented in mathematics and that if we were to use mathematics as our primary engine of inquiry, we would not be enquiring about our actual world, but rather about some abstract and incomplete version of it (Marchionatti, 2004). According to Marshall, excessive reliance on mathematics would, in the words of Arthur Pigou:

> (…) lead us astray in the pursuit of intellectual toys, imaginary problems not conforming to the conditions of real life, and further, might distort our sense of proportion by causing us to neglect factors that could not easily be worked up in the mathematical machine. (Pigou, 1925, p. 84)

These factors that could not be easily worked up in mathematical machines are precisely those factors that characterize the complexity of economic reality. And it is economic reality rather than any abstract mathematical version of it that should, Marshall believed, be the primary subject of economics (Marchionatti, 2004).

Marshall's recognition of the complex character of real world economic systems is signaled by his invocation of biology rather than physics (Hodgson, 1993; Hodgson, 2013, p. 8). In his *Principles of Economics* Marshall wrote:

> Economics, like biology, deals with a matter of which the inner nature and constitution, as well as the outer form, are constantly changing. (Marshall, 1961, p. 772)

In an article in *The Economic Journal* in 1898 Marshall expressed a similar view. There, he wrote that:

> 'Progress' or 'evolution,' industrial and social, is not mere increase and decrease. It is organic growth, chastened and confined and occasionally reversed by decay of innumerable factors, each of which influences and is influenced by those around it; and every such mutual influence varies with the stages which the respective factors have already reached in their growth. (Marshall, 1898, pp. 42-43)

So instead of emphasizing the parallelism with physics like for example Fisher did, Marshall emphasized much more the similarities that economics shared with biology. And the resulting recognition of the complexity of the world led Marshall to be cautious of mathematics in economics. That is, it led Marshall to be cautious of using mathematics as a primary engine of inquiry. For him, the complexities of economic reality were not arguments against the use of mathematics in economics in general, and he did recognize that mathematics could be of great importance (Schabas, 1989). But these complexities did, according to Marshall, constitute limits to the employment of mathematics in economics (Marchionatti, 2004).

A similar emphasis on the complexity of economic reality and appreciation of economics' similarities with biology rather than physics we find in the works of Friedrich Hayek. Hayek continuously emphasized the difference between the social sciences and physics and criticized what he called scientism: the desire of economics to be like physics. He argued that mathematical economics would never be able to achieve the kind of completeness that physics could achieve:

> While in the physical sciences it is generally assumed, probably with good reason, that any important factor which determines the observed events will itself be directly observable and measurable, in the study of such complex phenomena as the market, which depend on the actions of many individuals, all the circumstances which will determine the outcome of a process, for reasons which I shall explain later, will hardly ever be fully known or measurable. (Hayek, 1989, p. 2)

The reason for that state of affairs, Hayek argues, is the fact that:

> (…) the social sciences, like much of biology but unlike most fields of the physical sciences, have to deal with structures of essential complexity. (Hayek, 1989, p. 4)

Like Marshall, this conviction of the complexity of economic reality led Hayek to have a resistance towards formal modeling. He did not oppose formal modeling in general, but he emphasized that there are limits to its use. He stated that as we advance towards more and more complex situations:

> (…) we find more and more frequently that we can in fact ascertain only some but not all the particular circumstances which determine the outcome of a given process. (Hayek, 1989, p. 7)

The desire to capture reality in mathematical models, Hayek argued, causes economics to neglect factors that cannot be easily incorporated in the mathematical machine. In Hayek's words:

> In the social sciences often that is treated as important which happens to be accessible to measurement. (Hayek, 1989, p. 2)

Much is thus left out when a model is created, and the more complex the situation that is modeled, the less the model is like that actual situation. Already in 1937 this was a criticism that Hayek voiced against much of the economic work that was produced at that time:

> More recent work has been freer from this fault [of mixing up the a priori and the empirical] – but only at the price of leaving more and more obscure what sort of relevance their arguments had to the phenomena of the real world. (Hayek, 1937, p. 54)

So we see an emphasis on the fact that important factors are left out in any model of a complex economic reality and the fact that this poses limits to the use of mathematics in economics. This emphasis is something that we also find in some of the works of John Maynard Keynes, whose views on the matter were deeply rooted in Marshall's (Marchionatti, 2009). Keynes' work as an economist was essentially an attempt to cope with the complexity of the economic world and the organic interdependence of variables, founded on a conception of economics as a science of social complexity (Marchionatti, 2009). In his 1939 article *On Professor Tinbergen's Method* he, like Hayek and Marshall, makes clear that if a model is to be representative of economic reality not only significant but rather all causes must be accounted for in the model and that furthermore all these factors must be measurable so as to be able to account for them in a precise mathematical way. It is highly unlikely that all the factors that affect a certain outcome are in fact measurable and even more unlikely that we know all factors that are involved (Marchionatti, 2004, 2009). Keynes states that:

> If we were dealing with (…) independent atomic factors and between them completely comprehensive, acting with fluctuating relative strength on material constant and homogeneous through time, we might be able to use the method of multiple correlation with some confidence for disentangling the laws of their action. (Keynes, 1973,p. 286)

But we are not. Due to the nature of economic material, a complete and exact generalization is not possible, and any model we create will thus be incomplete (Marchionatti, 2004). Additionally this means that neither a mathematical theory nor a prediction can be confirmed by data from economic reality:

> If the method cannot prove or disprove a qualitative theory and if it cannot give a quantitative guide to the future, is it worthwhile? (Keynes, 1939, p. 566)

With an almost audible 'sigh' Keynes then concludes that he has:

> (…) a feeling that Prof. Tinbergen may agree with him, but that his reaction will be to engage another ten computers and drown his sorrow in arithmetic. (Keynes, 1939, p. 568).

Because of the complexity of economic reality, and because of the uncertainty that results from such complexity, the use of mathematical models is highly limited according to Keynes. Like Marshall and Hayek, Keynes believed that due to this complexity, mathematical models always rested on (implicit) a priori assumptions (such as *ceteris paribus*). The more of these assumptions are involved, the more incomplete and less like reality a model becomes. This was one of the main problems that Keynes noticed in the output of mathematical economics of his time:

> Too large a proportion of recent 'mathematical' economics are mere concoctions, as imprecise as the initial assumptions they rest on, which allow the author to lose sight of the complexities and interdependencies of the real world in a maze of pretentious and unhelpful symbols. (Keynes 1936, p. 298)

Mathematics was, according to Keynes, at best only a small part of the economic discipline. For him, capturing, understanding and interpreting complex economic reality also involved:

> (…) the amalgam of logic and intuition and the wide knowledge of facts, most of which are not precise. (Keynes, 1972, p. 158)

What I have tried to make clear in the above is that the rise of mathematics in economics was not without criticism. Furthermore I have attempted to show that behind several of these criticisms voiced by important and influential economists lies the idea that economic reality cannot be captured in mathematical models or formulas, due to its complexity and openness.

Over time many more of such concerns have been voiced (Blaug, 2002). Amongst those voicing these concerns were, again, very influential economists such as Nobel laureates Milton Friedman and Paul Krugman who expressed their unhappiness with the fact that economics was becoming more and more a branch of mathematics than a social science dealing with real world economic problems. In spite of all these concerns however, economics has developed to become a thoroughly mathematical science (Lawson, 1997, 2012; Chick, 1998, 2001; Blaug, 2002; Marchionatti, 2004; Hodgson, 2013). Why is it that these concerns have not made their way into mainstream economics? Why is it that, in the words of Ronald Coase, "mathematics rides triumphant in economics?" (Coase, 1972, p. 415). I now turn towards a possible answer.

Complexity marginalized: the desire for prediction

The desire of economics to be like physics not only directed its approach to modeling but also elevated prediction as the supreme goal of the economic discipline. This predilection for prediction still exists today, as is evidenced by the continuous subscription of the economic science to Milton Friedman's famous argument that the test of a theory is its capacity for prediction (Hodgson, 2013, p. 14). It is the desire for the ability to predict that, I would argue, is the crucial force behind the marginalization of the calls for caution made by economists such as Marshall, Hayek and Keynes.

But first I want to take a closer look at where this desire for predictive power might come from. Why was it the ultimate goal in physics and what can this tell us about the goal of prediction in economics? Why did economics take over physics' goal, rather than only its methods? These are of course very big questions, and no doubt many different factors are in play. In this section, I want to expand on one factor in particular, which I claim to be of great importance. This factor is a development that Max Weber (1922, p. 117) has termed 'disenchantment' and that Charles Taylor (2007) has more broadly called 'secularization'.

In his 1922 *Science as a Vocation* lecture, Max Weber noted something similar to what Hayek and Keynes expressed about economics at their time:

> Nowadays in circles of youth there is a widespread notion that science has become a problem in calculation, fabricated in laboratories or statistical filing systems just as 'in a factory,' a calculation involving only the cool intellect and not one's 'heart and soul. (Weber, 1958, p. 113)

This idea that anything can be calculated is what Weber termed 'rationalization' and it is a crucial element of the process of 'disenchantment'. It is the idea that:

> (…) principally there are no mysterious incalculable forces that come into play, but rather that one can, in principle, master all things by calculation. (Weber, 1922, p. 117)

The phrasing "master all things" is crucial here. 'Disenchantment', or what Taylor calls 'secularization' is the process of men taking over from God the control of the world. It is the process of the marginalization of God in the natural world in the 19th and early 20th century (Taylor, 2007, pp. 94-95). Slowly, things that were normally assumed to be determined by God or other spirits came under men's control. Industrialization and the technological innovations that came with this process provided society with the idea that we can control nature and shape our own world. Furthermore it provided us with the tools to do so. Secularization moved the privilege to alter and direct the world from the domain of God to the domain of human beings (Taylor, 2007, pp. 121-125).

The crucial tool in bringing the world under the control of men was natural science (Taylor, 2007). For natural science provided not only the ability to explain what had happened but also to predict what would happen. And this ability to predict is the crucial element in men's ability to control and direct the world they inhabit (Taylor, 2007). We cannot control what 'way we will go', or what will happen, if we cannot predict the consequences of our actions. Without prediction, the world is out of our control.

And so the development of the natural sciences and its increased ability to predict natural events provided men with an increasing amount of control over the natural world and enabled them more and more to direct the natural world in the direction that they wanted it to go.

But the natural world is not the only world that mankind inhabits. For next to that natural world we also live in a world of social processes. I argue that the desire for control extends to that social world just as it did to the natural world and thus claim that the desire for predictive power in economics is the result of this broader development of mankind trying to gain control over the world they inhabit; the result of the desire to remove all mystical forces from that world.

In order to gain control over economic reality, we need to be able to predict. It is prediction that is required for economic policy, rather than explanation. If we cannot predict, it makes no sense to try and change the social world, for we do not know what the consequences of our actions will be. And it is in that respect also not very surprising that we turned to physics. For physics had proven itself to be extremely capable in predicting and at the time that economists turned to physics, men had already achieved a large degree of control over the natural world. So I would argue that physics-envy, or 'scientism' in Hayek's words, not only results from an admiration for its rigor and form, but also very much from an admiration for its ability to bring the world we live in under our control.

My claim is thus that the desire for predictive power in economics is an expression of man's attempt to remove all mysterious and mystical forces, such as invisible hands, from the social world, and to master all things by calculation. But it is precisely those mystical forces, those things beyond calculation that Marshall, Hayek and Keynes introduce, though in a different form, with their emphasis on the complexity of economic reality. Like Daniel Stein says:

> Complexity is almost a theological concept; many people talk about it, but nobody knows what 'it' really is. (Stein, 1989, p. xiii)

A complex reality is coherent in some recognizable way but it has a structure admitting surprise and novelty, which cannot be known beforehand. The problems of such real world complexity are thus, according to Hayek:

> (…) not, as one might at first suspect, difficulties about formulating theories for the explanation of the observed events - although they cause also special difficulties about testing proposed explanations and therefore about eliminating bad theories. They are due to the chief problem, which arises when we apply our theories to any particular situation in the real world. A theory of essentially complex phenomena must (unlike physics) refer to a *large* number of particular facts; and to derive a prediction from it, or to test it, we have to ascertain all these particular facts. (Hayek, 1989, p. 6).

And to ascertain all these particular facts is, for complex situations, virtually impossible (Kilpatrick, 2001). So precisely because the economic world is too complex to be captured by mathematics and to be predictable, it is also too complex to be fully brought under men's control.

There is thus a tension between real world complexity and prediction in the social sciences. In a complex world the possibilities for prediction are at best limited (Hodgson, 2013). This, I argue, is one of the main reasons for the marginalization of the pleas for caution by Marshall, Hayek and Keynes. Their emphasis on complexity was something that did not correspond to the desire to predict and control; it did not fit with the process of the disenchantment of the social world. In a complex world, where men has no control over what will happen, what basis is there for economic policy? It has been argued that this is also the idea behind Hayek's promotion of the adoption of laissez-faire economics (Kilpatrick, 2001). In any case it seems that subscribing to the idea of a complex economic reality highly limits the ability to say anything more about the future than that it is unpredictable. The line of thought that holds that mathematics is able to capture economic reality, and thus is able to determine the outcomes of economic events by calculation, has a much better fit with this general desire to take control of the world we live in. If we can calculate the outcomes of different actions and events in economic reality, we have a basis for policy-making, a basis for manipulating the world to make it fit the way we think it should be. This is something we had achieved in the natural world, and we desired the same thing for the social world we inhabit. Complexity, on the other hand, requires us to admit and, more importantly, to accept, that there are still 'mystical and mysterious' forces left in that social world: forces which are unpredictable and whose rationalization lies beyond our cognitive capacities.

That the victory of mathematical economics was the result of a predilection for prediction and control rather than one for mathematical rigor and beauty is something that is supported further by the fact that there have also been several alternative mathematical approaches which were able to deal with complexity, but that have been neglected by mainstream economic theory (Hodgson, 2013, p. 14). These alternatives (such as chaos theory or complexity theory) are able to capture complex systems in mathematics, but do so with a significant loss of predictive power. And so the marginalization of these approaches suggests that mainstream economics is not focused on mathematical models in general, but rather more specifically on mathematical models that yield predictions (Hodgson, 2013, pp. 14-15). A possible explanation for this desire for predictive power is thus, as I have argued above, the disenchantment of the social world.

Conclusion

In this paper I have argued that the development of economics towards the mathematical science that it is today has not been without critique. I have discussed concerns voiced by Alfred Marshall, Friedrich Hayek and John Maynard Keynes, which stress the importance of recognizing the fact that economic reality is highly complex and that this constitutes some serious limits to the use of mathematics in economics and to the predictability of economic reality in general.

Even though influential and important economists voiced these concerns, they have had virtually no impact on the development of the economic discipline as it exists today. The reason for this, I have argued, is that there is a tension between prediction and complexity. In a complex system the possibilities for prediction are highly limited at best. I have argued that there existed however a demand for predictive power that might be a result of the process that Weber has called 'disenchantment' involving men's attempt to gain control over the natural world they live in. I suggest that this desire to gain control exists for the social world just as it does for the natural world. In order to control, we need to be able to predict, and therefore the desire of men to have control over the world has been an important source of the demand for prediction. The conflict between this demand for prediction a complexity is the main reason behind the fact that the concerns voiced by economists such as Marshall, Hayek and Keynes, have remained in the margins and have not been able to affect mainstream economics. What I have tried to show in this paper is how it was possible that economics became a mathematical science *in spite of* the concerns that were raised against these developments by influential economists. This is so because the basis from which they criticized mathematics, real world complexity, was in conflict with the fundamental process of disenchantment.

Acknowledgements

I would like to thank Professor Geoffrey Hodgson for his nomination and encouragement to rewrite the paper into a publishable article, and I would like to thank the editors of the ESJP for the time and effort they invested in guiding me through that process. Furthermore, I would like to thank Philippe Verreault-Julien and Felix van den Ottolander for the helpful comments on the paper they provided in one of the EIPE PhD Seminars.

Sam van Dijck has a bachelor degree in History (cum laude) and in Philosophy (cum laude) from the Radboud University Nijmegen. He is currently enrolled in the research master in Philosophy and Economics at the Erasmus University Rotterdam and in the master Political Science at the Radboud University Nijmegen. His main research interests lie in the fields of global political economy and global justice, with a special focus on the political and ethical dimensions of tax avoidance.

'Economics, complexity and the disenchantment of the social world' was written for the course 'Living Philosophical Insights from Dead Economists' taught by prof. Geoffrey Hodgson.

Notes

1. In this paper, mathematics is supposed to refer to the kind of formal tools used by the majority of the economic profession, which includes calculus and optimization, set theory, linear algebra and game theory.

2. Hayek believed that any restrictions on freedom to conduct trade, on price levels or on quantities sold would have serious negative consequences for welfare. The most efficient tool to improve the general level of wealth, he claimed, is competition of private actors in a free market. Keynes on the other hand believed that such a system often leads to macroeconomic outcomes that are inefficient. He thus argued for the occasional active policy intervention by the public sector, especially by monetary and fiscal instruments.

References

Blaug, M. (2002). Ugly Currents in Modern Economics. In U. Mäki (ed.), *Fact and Fiction in Economics: Models, Realism and Social Construction*. Cambridge: Cambridge University Press.

Blaug, M. (2003). The Formalist Revolution of the 1950s. *Journal of the History of Economic Thought*, *25*(2), 145-156.

Chick, V. (1998). On Knowing One's Place: the Role of Formalism in Economics. *The Economic Journal*, *108*, 1859–1869.

Chick, V. and S.C. Dow. (2001). Formalism, Logic and Reality: A Keynesian Analysis. *Cambridge Journal of Economics*, *25*(6), 705-21.

Coase, R. (1972). Marshall on Method. *Journal of Law and Economics*, *18*(1), 25–31.

Friedman, M. (1966). *Essays in Positive Economics*. Chicago: University of Chicago Press.

Friedman, M. (1999). Conversation with Milton Friedman. In Snowdon, B and Vane, H. (eds), *Conversations with Leading Economists: Interpreting Modern Macroeconomists*. Cheltenham: Edward Elgar.

Hayek, F.A. (1989). The Pretence of Knowledge. *The American Economic Review*, *79*(6), 3-7.

Hayek, F.A. (1937). Economics and Knowledge. *Economica*, *4*(13), 33-54.

Hodgson, G.M. (1993). The Mecca of Alfred Marshall. *The Economic Journal*, *103*(417), 406-415.

Hodgson, G.M. (2013). On the Complexity of Economic Reality and the History of the use of Mathematics in Economics. *Filosofia de la Economia*, *1*(1), 125-128.

Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money*. London: Macmillan.

Keynes, J. M. (1939). Professor Tinbergen's Method. *The Economic Journal*, *49*(4), September, 558-68.

Keynes, J.M. (1924). Alfred Marshall. In: *The Collected Writings of J. M. Keynes, vol. X*, Cambridge: Macmillan.

Keynes, J.M. (1973). *The Collected Writings of John Maynard Keynes: The General Theory and After: Part II Defence and Development*. Royal Economic Society XIV.

Klamer, A. and D. Colander. (1990). *The Making of an Economist*. Boulder: Westview Press.

Kilpatrick, H. E. (2001). Complexity, spontaneous order, and Friedrich Hayek: Are spontaneous order and complexity essentially the same thing?. *Complexity*, 6, 16–20.

Krugman, P. (2009a). How Did Economists Get It So Wrong?. *New York Times*, September. Retrieved from http://www.nytimes.com/2009/09/06/magazine/06Economic-t.html.

Krugman, P. (2009b). Mathematics and Economics. *New York Times*, September. Retrieved from http://krugman.blogs.nytimes.com/2009/09/11/mathematics- and-economics/?_r=0.

Lawson, T. (1997). *Economics and Reality*. London: Routledge.

Lawson, T. (2012). Mathematical Modeling and Ideology in the Economic Academy: Competing explanations of the failings of the modern discipline? *Economic Thought*, *1*(1), 3-22.

Malthus, T. (1820). *Principles of Political Economy Considered with a View of their Practical Application*. London: John Murray.

Marchionatti, R. (2004). What don't economists know now that Marshall knew a century ago? A note on Marshall's "sophisticated informality." *Journal of Post Keynesian Economics*, 26(3), 441-460.

Marchionatti, R. (2009). J.M. Keynes, Thinker of Economic Complexity. *Working Paper No. 02/2009*. Retrieved from 'Centro di Studi sulla Storia e i Metodi dell'Economia Politica' website http://www.cesmep.unito.it/WP/2009/2_WP_Cesmep.pdf.

Marshall, A. (1961). *The Principles of Economics*. London: Macmillan.

Marshall, A. (1898). Distribution and Exchange. *The Economic Journal*, 8(29), 37–59.

McCloskey, D.N. (1994). *Knowledge and Persuasion in Economics*. Cambridge: Cambridge University Press.

Pigou, A.C. (1925). *Memorials of Alfred Marshall*. London: Augustus M Kelley Publishers.

Popper, K. (2002). *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Routledge.

Ricardo, D. (1817). *On the Principles of Political Economy and Taxation*. London: John Murray.

Samuelson, P.A. (1939). Interactions between the Multiplier Analysis and the Principle of Acceleration. *The Review of Economics and Statistics*, *21*(2), 75-78.

Schabas, M. (1989). Alfred Marshall, W. Stanley Jevons, and the Mathematization of Economics. *Isis*, *80*(1), 60-73.

Stein, D.L. (1989). *Lectures in the Sciences of Complexity*. Redwood City: Addison Wesley.

Taylor, C. (2007). *A Secular Age*. Cambrige: Harvard University Press.

Weber, M. (1958). Science as a Vocation. *Daedalus*, *87*(1), 111-134.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit http://creativecommons.org/licenses/by-nc/3.0/

ESJP
#9 | 2015

Dennett's Drafters

The Mysteries of the Multiple Drafts Model

Thijs Heijmeskamp

§1 Introduction

The mind-body problem is one of the great mysteries. How are my feelings and thoughts related to the nerve cells of my brain? This question not only concerns scientists and philosophers, but everyone… for your consciousness seems very much bound up with who you are. In his book *Consciousness Explained* Daniel C. Dennett defends his grand theory of consciousness. It is an extraordinary book, in virtue alone of being read by both philosophers and the general audience. In this book Dennett presents us his Multiple Drafts model of consciousness. The Multiple Drafts model is an explanation of how our consciousness works. However, Dennett not only wishes to sketch the mechanisms of our consciousness, he also wants to show new ways of thinking about resolutions to the traditional mysteries of consciousness. This entails a critique on, according to Dennett, the mainstream view of the nature of consciousness; a view he calls *Cartesian materialism*. Cartesian materialism holds the assumption that there is some sort of 'stage' to which experiences present themselves to a 'mind's eye', an internal viewer. Dennett calls this 'stage' the *Cartesian Theater*. He vehemently rejects this notion of consciousness as a Theater, for he thinks that this notion is illusionary and does not give us a correct picture of consciousness.

When I discuss consciousness in this paper, I mainly talk about visual consciousness, namely the awareness and appearance of an external world through vision, following Dennett who introduces his model through a discussion of the visual system and keeps his discussion mainly limited to perceptual consciousness.

In this paper, I argue that Dennett does not provide us with an explanation of consciousness. His model is based on a wrong characterization of our phenomenology and as a consequence he has to rely on the mysterious notion of *probe* in order to 'explain' phenomenal experience. Before I discuss Dennett's Multiple Drafts theory, I first identify the position Dennett argues against: Cartesian materialism. Afterwards I will give a short note on his methodology. Then the Multiple Drafts model will be introduced through the Phi Phenomena. Subsequently, I shall look further into phenomenal consciousness itself. Two questions arise concerning Dennett's Multiple Drafts theory: (1) How exactly is the content of our visual system created? And, (2) how do we form a unified, coherent conscious experience? Dennett has the problem that he cannot provide a satisfactory solution to these two questions. Both questions will be discussed in turn after which I conclude.

§2 Cartesian materialism

According to Descartes human beings are composed of a material body and an immaterial soul. Although the body and the mind can interact with each other, they are fundamentally different substances. The first being the *res extensa*, with its primary attribute extension, and the latter being the *res cogitans*, the thinking substance. The fact that we, human beings, are made up of these two substances separates us from the animals that only possess a body, not a mind. Animals only operate on mechanical notions whereas humans have free will. According to Descartes, the interaction between these two substances has to take place in the human body. Stimuli from the senses have to reach the mind and the (immaterial) mind must have control over the (physical) body, otherwise humans cannot move about in this world. Descartes pinpointed this place of interaction between mind and body in the pineal gland, an organ that sits in the midline of our brain and is attached to the rest of our nervous system. It is here, according to Descartes, that mechanical input from the body is translated to the mind and input from the mind translated into mechanical action. This is the theory of *Cartesian Dualism*.

Yet locating the place of interaction between mind and body did not help Descartes in solving one of the major problems of this theory: how do mind and body interact? Somehow the states of mind and body must be brought into relation; but if the mind is unextended and the body extended, how can this interaction take place? Placing the interaction in the pineal gland offers no clarification of this mysterious interaction.

Another consequence of Cartesian Dualism – besides the above mentioned troublesome relation between two substances – is that the theory assigns a center to the brain, a central place in the brain that integrates all conscious experience. All traffic from the senses has to pass through the pineal gland. Thus it is the pineal gland that houses consciousness. Although there are not many proponents of Cartesian Dualism nowadays, the idea of a centralized gateway or functional center is still present in many theories of consciousness. Dennett calls "the idea of such a centered locus in the brain *Cartesian materialism*, since it is the view you arrive at when you discard Descartes' dualism but fail to discard the imagery of a central (but material) Theater where 'it all comes together.'" (Dennett, 1991, p. 107) What enters this part of the brain is what you are conscious of. The *Cartesian Theater* is a metaphor for how consciousness sits in the brain. That is, the brain builds up a unified picture, a representation, which is 'viewed' by a central entity in order to become experience. Dennett admits that most likely no one today explicitly endorses Cartesian materialism, but he argues that the imagery of the Cartesian Theater is persuasive and "keeps coming back to haunt us." (Dennett, 1991, p. 107)

But positing such a place of central processing can in the end lead to faulty analysis. Those who subscribe to a Cartesian Theater (whether explicitly or implicitly) can give an explanation of some cognitive ability or process by providing a functionalist analysis, but rely on an internal agent to 'tie the knots together.'[1] So any loose end which cannot be explained by the functionalist[2] analysis is being attributed to this point of central processing. This point of central processing ends up having the cognitive abilities that needed explanation in the first place. According to the Cartesian materialist consciousness can only be understood by a Theater and an audience, because that which is being projected in the Theater must be viewed by an audience. This audience functions like an internal observer, the homunculus. Only what ends up being projected in the Theater can be conscious. If consciousness can only be understood via a Theater, we can also only understand the observer through the metaphor of a Theater. So the Theater and its audience need another Theater and audience in the head of the audience in order to be understood, leading to an infinite regress of Theaters.

One of Dennett's aims is to get rid of this notion of a centralized place of processing in the brain in order to escape Cartesian materialism. For him, there is no single brain area in which it all comes together. With this decentralized notion of consciousness, there is no need for a Theater and no need for a homunculus to live inside our brains. Dennett's Multiple Drafts model of consciousness must first be understood as an alternative for Cartesian materialism.

§3 Side-note on methodology

One of the reasons the false notion of the Cartesian Theater came into existence, was that people made the mistake of 'naively looking inward'. Dennett denies that people have immediate epistemic access to their conscious states. People can be mistaken about their own mental states. He rejects *introspectionism* – the idea that we have privileged access to our own thoughts and feelings and this access is somehow immune to errors (Dennett, 1991, p. 67). Descartes privileged his own thoughts with his 'Cogito ergo sum' and gave us the Cartesian Theater in the pineal gland. The phenomenologists adopted Descartes' first-person perspective, "in which I describe in a monologue (which I let you overhear) what I find in my conscious experience, counting on us to agree." (Dennett, 1991, p. 70)[3] On the basis of reflection on our own experiences, we can come to know what our conscious states are like.

According to Dennett, first-person methods of introspection have no privileged position over third-person methods. In fact, Dennett proposes a third-person perspective when it comes to studying our consciousness and inner world. He calls this perspective *heterophenomenology*, which stands in contrast to the earlier mentioned introspective phenomenology. Where Husserl bracketed the outer world, Dennett brackets the inner world. Husserl wanted to neutralize his metaphysical and empirical commitments and Dennett wants to neutralize his commitments to the ontological status of mental states. A third-person perspective cannot describe the inner world, so the heterophenomenologist assumes an agnostic attitude towards the ontological status of mental states. He studies not these mental states, but the behavior and beliefs subjects have about their inner world. Reports from subjects on their conscious experience are just further bits of evidence about the inner world. The heterophenomenologist, who wants to study consciousness, distrusts the first-person accounts that people give about their own qualitative experiences; he sees those accounts as idiosyncratic, unreliable and plagued with inconsistencies. The heterophenomenologist fictionalizes the reports of the first-person conscious experiences of subjects. The reports are seen as abstractions that describe the complex cognitive state of a subject. Dennett suggests that we should interpret the reports of first-person conscious experiences in the same way we interpret works of fiction. These reports should be read as novels. Subjects receive instructions from experimenters and give verbal feedbacks, which are all later converted to transcripts and studied by the heterophenomenologist. These heterophenomenologist texts are construed as a world of theorist's fiction (Wah, 2007). For the introspectionist the primary data are the experiences, for the heterophenomenologist the primary data are the utterances and behavior of the subjects they research. Heterophenomenology aims to be a scientific method to study consciousness. The heterophenomenologist does not suppose that the phenomenological accounts people give, have to be shared amongst all.

Although Dennett's skepticism about naïve introspectionism is justified, he may be too quick in dismissing first-person accounts. According to O'Regan & Noë, Dennett wrongly characterizes how things seem to perceivers (2001, p. 965). For instance, Dennett criticizes the notion that the visual field is in sharp detail and uniform focus from the center to the periphery. But according to O'Regan & Noë, normal perceivers are not aware of their visual fields in this way. They take the world to be solid, dense, detailed, and present. This wrong characterization is the result of Dennett reducing all first-person approaches towards consciousness to naïve introspectionism. O'Regan & Noë propose, instead of naïve introspectionism, an approach consisting in an "attentiveness to the complexity of the activity of perceptual exploration" (2001, p. 965). This approach would allow talk about the facts of our experience at a personal level. It would be possible to formulate substantive empirical questions on the first-person qualitative experiences.

§4 The Phi Phenomenon

Before Dennett discusses his model, he introduces the reader to the optical illusion called the *Phi Phenomenon*. Take two stationary dots separated by four degrees of visual angle. When these spots are lid in rapid succession, it seems there is a single spot moving between two points. When we give these spots different colors, say red and green, another interesting thing happens: the 'moving' spot appears to change color midway. This illusion is also persistent. Even armed with the knowledge of the Phi Phenomenon, you cannot help seeing the moving spot and the moving spot changing color. This is an odd thing, for how can the first spot seem to change color before the second spot is observed? The green spot cannot be attributed as content to any event until the light from the green spot has reached our eyes and triggered a neural response.

Dennett gives us two possible explanations, which he will both discard in favor of his own Multiple Drafts model of consciousness (1991, pp. 116-117). The first explanation is the Orwellian revision and contains a revision of memory. The second explanation is the Stalinesque revision and is a perceptual revision. In the first explanation, shortly after the second spot has entered consciousness, the mind makes up a narrative about the intervening events. It is this new event which enters memory. In the second explanation there is a delay in the brains editing room. The first spot is held in preconsciousness until the second spot arrives. In the editing room intermediate content is created and the finished (illusionary) product arrives at consciousness. According to Dennett, there is no possible reason to choose one explanation over the other. There is no way to demarcate the place and time in the brain where something enters consciousness (Dennett, 1991, p. 126).

According to Dennett, both the Stalinesque and the Orwellian revision fall away in the Multiple Drafts model, because there is no finish line anymore where everything has to be presented. There is no distinct moment of phenomenal awareness anymore. He claims that there is no real difference between both explanations. Both explanations still presuppose a Cartesian Theater, because both explanations assume a fixed point in the process where content becomes conscious. Dennett's Multiple Drafts model replaces the notion of a Cartesian Theater by parallel, multitrack processes of interpretation and elaboration. All information entering the nervous system is under continued revision (Dennett, 1991, p. 111). There is no center; instead the perspective of the observer is smeared out in time and space. The structure of the mind is not like the structure of a computer with a CPU, but more akin to the model of a Pandemonium.

The Pandemonium architecture was developed by Oliver Selfridge. The Pandemonium is a pattern recognition system that consists of four layers, each layer comprising units called *demons* who 'scream' for attention. The first layer records the sensory input. The second layer consists of feature detector demons, which detect certain features. For instance, one demon detects a horizontal straight line and another detects a curved line. The cognitive demons in the third layer are sensitive to these detected features. Each cognitive demon recognizes a certain pattern. The 'screaming' of a cognitive demon is determined by how much of their pattern is detected by the feature detector demons. Finally there is a decision demon that hears the shouting of the layers below and decides what pattern was presented in the layer below. The Pandemonium itself is far too simple to be a model of the mind, but Dennett is interested in the parallel nature of the Pandemonium and the absence of centralized processing (Schneider, 2007).

Neuroanatomical and neurophysiological research shows that the visual system in the human brain is only loosely hierarchical and functionally more interactive. Visual stimuli, such as the two dots in the Phi Phenomenon, may initially flow serially through the lowest level of the visual system, but then quickly reach modules on multiple levels (where more specific features are detected). A module is a specialized local section of the brain that detects certain features. These modules simultaneously process information about one and the same stimulus (Akins, 1996). This is what Dennett refers to as *parallel, multi-track processes*. Each module infers the presence of certain properties and the processing of visual stimuli gradually yield discriminations of greater and greater specificity. Perceptual modules carry out content discrimination, content fixation or feature detection, and cognitive sites make decisions or judgments or perform processes of interpretation and elaboration (Dennett, 1991, p. 134-135). Parts of the brain (modules) go into states that discriminate certain features. First there is the mere onset of stimulus, then location is discriminated, then shape, then color. Later (apparent) motion is discriminated and eventually object recognition takes place. Each module determines a certain feature of the world by ordinary computational means. At any moment there are multiple narrative fragments, or 'drafts', which are in various stages of editing. Some or all of these 'drafts' converge to gene-rate intentional behavior of the organism. The lesson Dennett takes from the Phi Phenomenon is that if one wants to settle on some moment of processing in the brain as the moment of consciousness, that moment is always arbitrary.

Modules that make decisions or judgments still seem to imply a homuncular theory of consciousness. Because persons make decisions and judgments, not parts of the brain. According to Dennett we must not understand decisions and judgment in the case of (perceptual) modules in the full sense that we understand decisions and judgments when it comes to persons. When it comes to the modules, 'decisions' and 'judgments' are metaphors for how a module determines a certain feature of the world. Dennett wants to replace the metaphors of the Cartesian Theater with new metaphors and he does not claim to replace the metaphor of the Cartesian Theater with a nonmetaphorical theory (Dennett, 1991, p. 455). The single-minded agent is to be broken down "into miniagents and microagents (with no single boss)" (Dennett, 1991, p. 458). In the following sections of this paper I address the success of Dennett's new army of metaphors in escaping the fallacies of the Cartesian Theater.

§5 When does consciousness arise?

According to the Multiple Drafts model, perception is accomplished in the brain by parallel, multi-track processes of interpretation and elaboration of sensory inputs. But when consciousness exactly arises, has not been discussed yet. The first problem for Dennett's theory is due to the decentralized nature of the Multiple Drafts theory: it presumes that the visual system processes different properties of stimuli at different sites and at different speeds. How then, is it possible that we have a coherent and unified experience of the world? This problem has both a spatial aspect and a temporal aspect. Spatial in the sense, as we have seen, that if properties of stimuli are being processed at different spatial sites, it is not clear how everything is brought together in a spatially unified whole. Temporally in the sense that since a modular conclusion[4] about a single event will be produced across a period of time and will be inter-mixed with other conclusions about earlier and subsequent events (Akins, 1996). This is known in the philosophy of mind as *the problem of binding*: how do various features of a visual scene come together in a unified experience? The process of discriminating content for our visual experience is not unified. So how does a single unified experience follow from a disunified process? There appears to be a gap between modular content discriminations and our personal experiences.

However, Dennett does not see the problem of binding as a real problem, for it presupposes Cartesian materialism. The problem assumes that the spatial unity of a perceptual experience must be mimicked by the spatial unity of the representations. Representations of temporally unified objects and events must occur in the same sequence as those objects and events. The assumption is that the phenomenological properties of our experience must match the physical properties of the neural vehicles (Akins, 1996). Dennett claims that representations of single, spatially unified objects need not themselves be spatially unified or singular. It is through symbolic representation that the brain differentiates the order in which 'conclusions' (Dennett's term) are produced. We perceive an ordered world of objects and events, because those temporal relations are symbolically represented by the brain. It is not the disunity of *content discriminations* that poses a problem for perceptual experience. It is only the disunity of *content* that provides a problem for understanding the form of our perceptual experience.

Thus Dennett puts aside the problem of binding, for there is no more need for physically unified representations. The problem that remains however is the problem of the unification of representational content. The pandemonium-like model of the brain leaves the brain with many distributed contentful states and modular conclusions. Some of these states die out and leave no further trace. Others leave trace on subsequent verbal reports of experience and memory. As soon as content discrimination has been accomplished, it becomes available for eliciting some behavior. Content arises, gets revised, contributes to the interpretation of other content or to the modulation of behavior, and yields over the course of time something like a narrative stream. At any moment in time there are multiple drafts of narrative fragments at various stages of editing in various places in the brain. It is through the notion of *probing* that Dennett wants to solve the problem of the unification of content. Probing is a process in which a consistent narrative thread is selected from among many of the states described (Dennett, 1991, pp. 134-135). Probing the narrative stream at different intervals will produce different effects and thus produce different versions of what occurred. Probes do not happen at regular fixed intervals, but are initiated sometimes by a need for action or sometimes by a self-imposed question. It is not necessary that probes give a full account of the events. The task of a probe is to unify information required to perform the task or solve the problem at hand. Probes arise irregularly in response to internal or external puzzles and initiate the integration of select subsets of information.

It is important to note that probing the narrative stream is a process of selection, not rewriting, re-ordering, translating or transporting the information to a central place so that the events can be played in a Cartesian Theater. Any narrative that gets precipitated provides a 'time line', a subjected sequence of events from the point of view of an observer. It is the answer to the probes that makes up our phenomenological consciousness. Probes are the bridge between the sub-personal system, where there are only content discriminations, and our psychological phenomenal experience, where we have an experience of an objective coherent world. If probes were absent, we would not be conscious of a narrative, since the answers to the probe form a part of a subject's experience. According to Kathleen Atkins (1988) the view that "probes are both necessary and sufficient for conscious experience" can be attributed to Dennett.[5] Without the probe there is no consciousness.

But, the question remains: when does consciousness exactly arise? Does consciousness arise during the binding of representational content or only after the probe is completed? If consciousness arises during the binding, at which stage of the process does it arise? Surely binding takes time. There seems to be neither a method for answering this question nor any evidence. The binding process is not available to the subject; we are not aware of it. We have no access to the temporal relations among the representational vehicles themselves and the individual conclusions that are reached, nor do we have access to the order in which independent representational contents are unified. Phenomenological accounts, first-person or third-person, are of no help because our experience is that of an orderly world. We are only aware of the results of the process, not the process itself. We see events in the external world, not the binding processes. Although the impossibility to know when consciousness exactly arises poses no problem for Dennett,[6] it does point to the mysterious nature of the probe and the essential role it plays in establishing conscious experience. More importantly, it remains unclear how the binding process works. How exactly does the probe unify the content of our representations? It is clear, however, that probing does not create a new representation. Our conscious experience is not the representation of the multitude of representations created by the perceptual modules. The content of our conscious experience is mirrored by their underlying neural representations. Once a feature detection or discrimination has been made by a specialized module in the brain, the information content is fixed and need not be rediscriminated by a central discriminator.

§6 Dennett's black boxes

When we return to the color Phi Phenomenon, both the Orwellian (revision of memory) and the Stalinesque (perceptual revision) explanation fall short, because on a small time scale there is no more distinction between a revision of perceptual input and a revision of memory. The brain does not build up a single final representation of the world; there is no final draft and canonical narrative that a researcher may or may not access. What happens is that the brain creates content, a perceptual module makes a judgment, and this content is available to govern activity and/or leave its mark on memory. Judgments made by perceptual modules only have to be made once. In the case of the Phi Phenomenon the content that is created is the motion between the dots. This means that the sensory input (the stimuli) does not equal perception. The world as we experience it is a grand illusion. Not in the traditional sense that we are given much less than we see, so what we think we see must arise through the workings of our brain. But in a new skeptic sense, that we do not have the experience we think we have; we are radically deceived by our brains about what our experience is (Noë, 2009, pp. 200-201). All of our sensory input goes through a process of interpretation and elaboration.

So perception is a parallel, distributed, erratic and non-linear process with no central locus. This means that it is possible (and very likely) there will be simultaneous information processing of signals carrying information from sequential stimuli. At each stop in the process there are individual perceptual modules, which carry out 'content discriminations', 'content fixations' or 'feature detections', or non-perceptual or cognitive sites which arrive at 'decisions' or 'judgments', or perform 'processes of interpretation and elaboration'. It is these modules together that form something like a narrative stream.

One of the problems is that we do not know what happens inside the modules. For instance, how a module carries out shape recognition remains a mystery. From a retinal image the brain must extract different kinds of information of the visual scene. Which criteria do the modules use to determine which details of the sensory information are salient or not? Dennett does not leave us with an explanation of how the content of our visual system is created. How is it determined what judgments must be formed by modules, and what to edit out or what content to fill in (for instance, the motion and color change in the Phi Phenomenon)? In other words, when do the modules ascribe meaningful content and on what basis? We can imagine that simple feature detection modules carry out some basic form of shape recognition, but how do more complicated cognitive modules or discrimination modules create content? Where do the rules come from upon which they 'act'? For now, these modules are like 'black boxes' for us; we do not know what they look like on the inside.

Dennett's model of the mind seems to be somewhat detached from the world, i.e. the world is only there to provide sensory stimuli and the relation between perception and visual stimuli is merely inferential. It is the visual system that creates content from the sensory stimuli. The criteria for determining the creation and selection of content can only be provided by the brain itself. Criteria are necessary, because there must be a correspondence between the specific stimulus and the created content, because if this relation were arbitrary, the visual system would not work. Positing that the modules themselves establish these criteria, is in itself not an answer for it still does not answer on what basis these criteria are established. The point is: instead of positing a single homunculus, Dennett posits multiple homunculi. Although these homunculi have simpler abilities and powers than the Cartesian homunculus, they carry little explanatory power. The (perceptual) modules still remain a mystery.

§7 Perceptual unity and the self

The modular make up of our visual system can leave us, especially in the short run, with multiple and contradictory conclusions. The only demand that can be laid upon such a system (or better said the system forces upon itself) is the demand for consistency. Because the world is consistent, our experience of the world must also be consistent. So multiple contradictory conclusions cannot coexist in the long run, and consistency checks can identify these inconsistencies. It is the role of probes to perform these consistency checks.

Our conscious experience is a unification of selected parallel and distributed processes. When these processes are probed they form different narratives, which in turn are continuously revised and updated. It is however unclear how these narratives are chosen. If consciousness is merely the result of edited visual data, there seems to be a need for an editor. Since there is no homunculus, no central place of processing, it is up to the visual modules and the probes to create a coherent unified conscious experience. This implies that the probes need more demands than the demand for coherence in order to select a coherent narrative or to adjust narratives in order to make them coherent. Some data have to be left out of conscious experience and this editing out has to be communicated between different modules in order to form narratives. These probes somehow interact with the content of modules from (in this case) the visual system, somehow combine some content of different modules and elect to ignore other content in cooperation with other probes, while the constructed narrative is under continuous revision. Dennett's probes provide us with an even bigger mystery than his modules. Not only do we not know how these probes work (how they interact with the perceptual modules), but we also do not know what elicits a probe.

When it comes to consciousness, as we have seen, it is the probes that do the heavy lifting. Consciousness arises when the stream of narratives is being probed. But why must the perceptual content be unified as a visual experience? The question why experience is needed in addition to all the computational processes is not addressed by the theory. In other words: what drives the process of the unification of our conscious experience? With the absence of a homunculus, the core of our identity, Dennett speaks of the 'self' as the center of narrative gravity (Dennett, 1991, p. 418). The self is not the source, but the result of the different narratives that are spun by 'us'; it is a metaphorical point where all aspects of our identity converge. This also means that our notion of 'me' can change and is different over the years. The self is formed by the unified narratives and cannot be the driving force behind the unification of our conscious experience. It seems that the probes are endowed with the cognitive powers Dennett want to explain.

§8 Conclusion

Dennett rejects all homuncular theories of consciousness and gives us a decentralized account of consciousness. O'Regan & Noë pointed out that Dennett is too quick in dismissing first-person accounts of our conscious experiences with his method of heterophenomenology. This leads him to a fundamentally wrong characterization of our consciousness and forces the problem on himself of explaining why our conscious experience is illusionary. Dennett starts out with the critique that it is a false assumption that our visual field is sharp in detail and uniform in focus from the center to the periphery. This leaves Dennett with a problem that follows him throughout his theory in that he has to account for the fact that normal perceivers have this false assumption. But normal perceivers do not take themselves as to experience all the environmental details at once. We see ourselves as situated in an environment.

According to the Multiple Drafts model, perception is accomplished in the brain by parallel, multi-track processes of interpretation and elaboration of sensory inputs. These content discriminations produce something like a narrative stream. Probing this stream at different places and times produces different effects and precipitates different narratives. There are many small agents screaming for attention. What we experience is a product of many processes of interpretation. Frustratingly, Dennett has very little to say about how these content discriminations work and it is unclear what governs the modules. Since the relation between these modules and the world is only inferential, there seems to be a need for a programmer.

Although our first-person account of consciousness says, according to Dennett, very little, about how our consciousness really arises Dennett still faces the problem of accounting for this unified experience. However, the probes seem to have the cognitive powers that Dennett tries to explain with his theory, thus falling into the same trap as the Cartesian materialists. The problem with the homunculus fallacy is that it attributes the whole mind to part of the system, thus offering no explanation at all. If Dennett's metaphorical model of the mind is to be successful in replacing the metaphor of the Theater, it needs to provide elucidation where homuncular theories do not. Dennett's metaphors must be more productive in explaining consciousness in unconscious terms in order for it to be a successful theory.[7] Even if we grant Dennett his point that his model of consciousness is merely intended as a metaphorical model - we are not forced to assign judgments in the strong sense to (perceptual) modules, like we do to persons - he needs to demonstrate that his metaphors are better tools for understanding consciousness. But these new tools seem to share certain faults with the metaphors used by Cartesian materialists. Although Dennett breaks up the homunculus into smaller homunculi, they maintain much of the mystery. Ultimately, Dennett does not provide us with an explanation why the changing of red to green in the Phi Phenomenon is conscious at all. In fact, Dennett does not appear to need consciousness at all for his model of the mind. In Dennett's functionalist explanation of visual perception it remains unclear why we have conscious experience or need conscious experience. The unification of the multiple drafts through probing is not a conscious experience (we don't have access to the probes) and we only experience its results consciously. Consciousness, understood this way, seems to be a mere side-effect of probes. Dennett's metaphors speak very little of consciousness.

Although the Multiple Drafts model aims to give an explanation of how our consciousness works, the model still relies on mysterious parts, in this case the probes, to account for conscious experience. Dennett fails to give an explanation on how a coherent conscious experience is established from the parallel multitrack processes. He successfully escapes the problem of binding, but cannot count for the unification of content in the brain. It seems that the Multiple Drafts theory requires a drafter.

Acknowledgements

I would like to thank prof. dr. F.A. Muller and Tjeerd Visser for comments on a earlier version, and the anonymous editors of the ESJP for helping me improve this paper even further.

Thijs Heijmeskamp (1988) obtained his BA in Philosophy at the Erasmus University Rotterdam. He is currently enrolled in the MA Philosophy program. His main research interests are in the areas of philosophy of mind and metaethics.

'Dennett's Drafters' was written for the master's course 'Consciousness From All Sides' taught by prof. dr. F.A. Muller.

Editorial note

Since this essay was written by a member of the editorial board of the Erasmus Student Journal of Philosophy, it was subject to a more extensive review procedure. For more information, see http://www.eur.nl/fw/english/esjp/submissions.

Notes

1. For example, according to Cartesian materialism, we build up an internal representation corresponding to what we experience. The data for vision are to be found on the retinas. Some scientists abide to the principle of Cartesian materialism and take the problem of the inverted retinal image seriously. The retina is the inner coat of the eye which is light sensitive. Because light passes through the lens of the eye, the image is inverted. According to some scientists, the brain has to adjust for this inversion (Noë, 2009, p. 144). But this problem is based on the misguided assumption that the 'image' projected on the retina, is an actual image. Seeing the retinal image as a picture doesn't explain vision. In order to explain vision, scientists still has to explain how the brain 'sees' the image on the retina.

2. Functionalism in philosophy of mind holds that a mental state is not defined by its internal constitution, but by its *function* in the system which it is part of.

3. I will not address whether Dennett's characterisation of the philosophical tradition of Phenomenology is correct. I will however address whether Dennett's own phenomenological account of our visual consciousness is correct.

4. Dennett denies that for experience we build up a single unified representation. The experience of for instance warm coffee comes about through a variety of sensory modules, each drawing their own perceptual conclusions. These conclusions from shape modules, color modules etc. are not rewritten into a single homogenous representational form.

5. It must be noted that Dennett never explicitly states this view in *Consciousness Explained*.

6. Dennett claims that no theory of consciousness (which, according to him must be a functionalist explanation) would be able to determine the exact moment of consciousness (1991, p. 401-404).

7. Assuming that consciousness can ultimately be explained in unconscious terms, which Dennett aims to do (Dennett, 1991, p. 454).

References

Akins, K. (1996). Lost the Plot? Reconstructing Dennett's Multiple Drafts Theory of Consciousness. *Mind & Language, 71*(1).

Dennett, D. C. (1991). *Consciousness Explained.* London: Penguin Books.

Noë, A. (2009). *Out of our Heads. Why you are not your brain, and other lessons from the biology of consciousness.* New York: Hill and Wang.

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24,* 939–1031.

Schneider, S. (2007). Daniel Dennett on the Nature of Consciousness. *Blackwell Companion to Consciousness.* Oxford: Blackwell Publishing, 313-324.

Wah, T. K. (2007). Heterophenomenology Debunked. *Journal of Humanities and Social Studies*, *1*(1). Retrieved from http://www.scientificjournals.org/journals2007/articles/1208.pdf.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit http://creativecommons.org/licenses/by-nc/3.0/