

## De Doxastische Catch-22 van nuttige illusies

Maarten Boudry

ANTW 109 (3): 385–393

DOI: 10.5117/ANTW2017.3.BOUD

In zijn recensie van mijn boek *Illusies voor gevorderden* (Uitgeverij Polis, 2015) in dit tijdschrift (2016, nr.4) biedt Jan Willem Wieland een aantal interessante bedenkingen en tegenvoorbeelden bij mijn waarheidspleidooi, in het bijzonder bij de ‘Catch-22’ van nuttige illusies. Ik wil graag van de gelegenheid gebruik maken om die Catch-22 wat nauwkeuriger uit te werken, in het licht van Wielands opmerkingen.

In *Illusies voor gevorderden* onderzoek ik de hypothese dat sommige illusies heilzaam kunnen zijn voor lichaam en geest, of nuttig voor de samenleving. Foute overtuigingen kunnen gevaarlijk zijn, omdat onze overtuigingen hefbomen zijn om te handelen. Maar misschien kunnen we wel goed uitgekende illusies vinden, die niet alleen veilig zijn, maar bovendien concrete voordelen bieden tegenover de waarheid. Indien dergelijke illusies bestaan, wat staat ons dan te doen? Moeten we ze omarmen? Of is de waarheid toch steeds verkieselijker, zelfs indien kwetsend of verontrustend? Een centraal filosofisch argument in mijn betoog noem ik de Catch-22 van geloof in nuttige illusies, die ik hier de Doxastische Catch-22 zal noemen (in het boek vermijd ik filosofisch jargon).

Die Doxastische Catch-22, geïnspireerd door de beroemde roman van Joseph Heller (1994 [1962]), beschrijft een fundamentele paradox in de queeste naar heilzame illusies, die de vorm aanneemt van een *destructief dilemma*. De oorspronkelijke Catch-22 van Heller beschrijft een absurde regel voor militaire piloten tijdens de Tweede Wereldoorlog. Elke piloot was verplicht om aan missies deel nemen, tenzij hij aantoonbaar geestesziek was. In dat geval kreeg hij vrijstelling. Als een piloot echter goed bij zijn hoofd was, wou hij natuurlijk liever niet vliegen. De missies waren zinloos, levensgevaarlijk en leverden nooit iets op. Wie weigerde om aan een missie deel te nemen, gaf daarom blijk van uitstekende mentale gezondheid. Precies daarom moest hij vliegen. Iemand die echter vrijwillig

wilde vliegen, was beslist gek en hoefde daarom niet te vliegen. Ziedaar de Catch-22: 'Als hij vloog, was hij gek en hoefde hij niet te vliegen; maar als hij niet wilde, was hij gezond en moest hij.'

Nu loopt het onderzoek naar nuttige illusies ook vast in een pragmatische Catch-22. In het boek beschrijf ik die paradox informeel, maar ik wil ze hier zorgvuldiger uitwerken. Als we willen weten of een illusie al dan niet heilzaam is, moeten we eerst de voor- en nadelen onderzoeken. Sommige illusies zijn namelijk gevaarlijk, dus die willen we liever vermijden. Bovendien hebben illusies de neiging om zich te vertakken in ons web van overtuigingen, waardoor ze tot onverwachte neveneffecten leiden. Opletten geblazen dus. Stel nu dat we na onderzoek vastgesteld hebben dat een zekere foute overtuiging dat  $p$  niet alleen veilig is om te omarmen, maar bovendien netto voordelen biedt tegenover (het geloof in) de waarheid dat  $\sim p$ . Met andere woorden: we hebben een nuttige illusie gevonden. Probleem is dat we er niet doelbewust voor kunnen kiezen om  $p$  te omarmen: we kunnen niet zomaar een schakelaar in ons brein omdraaien teneinde  $p$  aan onze overtuigingen toe te voegen. Dat is in strijd met wat in de literatuur bekend staat als het principe van *doxastic involuntarism* (Adler 2002), en wat ik in het boek vertaal als 'geloofswillekeur' (cf. hoofdstuk 3).

Hiermee belanden we op de eerste tand van het destructieve dilemma: we weten dat  $p$  een nuttige illusie is, maar we kunnen er niet langer de vruchten van plukken, omdat we  $p$  niet langer *oprecht* kunnen geloven. Maar stel nu dat we  $p$  wél oprecht geloven, omdat we menen over goede redenen te beschikken dat  $p$  waar is. Dan kunnen we wel de vruchten plukken van het geloof in  $p$ . Het probleem is nu dat we niet langer kunnen weten of het geloof dat  $p$  heilzaam is. Misschien is het wel schadelijk, omdat het vroeg of laat tot een onzachte aanvaring met de werkelijkheid zal leiden. Noemen we de illusoire wereld waarin we vertoeven  $W^*$ , in tegenstelling tot de echte wereld  $W$ . Om te weten of geloof in  $W^*$  'nuttig is of eerder gevaarlijk', zoals Wieland het uitdrukt, moeten we de discrepantie tussen  $W$  en  $W^*$  meten en nagaan of geloof in  $W^*$  tot gevaarlijke situaties kan leiden. Daarbij moeten we ook de (utilitaire) kosten en baten van de overtuiging dat  $p$  vergelijken met die van de (ware) overtuiging dat  $\sim p$ . Maar daar zijn we per definitie niet toe in staat: zolang we oprecht zijn opgesloten in  $W^*$ , kunnen we het verschil tussen  $W$  en  $W^*$  niet meten. Dat is de tweede tand van het dilemma.

En als ik die informatie toch kan inwinnen, zodat ik alsnog een vergelijkende analyse kan uitvoeren? Dan hebben we onze illusie onvermijdelijk doorprikt, want dan namen we kennis van  $W$ . Dan belanden we weer bij de eerste tand van het destructieve dilemma. Samengevat: om te weten of een illusie heilzaam is, moet je haar effecten eerst onderzoeken; maar dan kan

je er niet langer in geloven. En als je er wel oprecht in gelooft, kan je er wel de vruchten van plukken, maar kan je niet langer weten of ze heilzaam is. Wieland (2016: 564) stelt de volgende formalisering voor van de Doxastische Catch-22:

- (P<sub>1</sub>) Voor alle illusies *x* geldt: we weten of *x* nuttig is, of leven oprecht in *x*.
- (P<sub>2</sub>) Als we weten of *x* nuttig is, dan hebben we *x* onderzocht en kunnen we *x* niet langer oprecht geloven, en de voordelen ervan proeven.
- (P<sub>3</sub>) Als we oprecht in *x* leven, dan moeten we erachter zien te komen of *x* nuttig is of eerder gevaarlijk, of moeten anderen ons hierover informeren, maar dan kunnen we *x* weer niet oprecht geloven.
- (C) Dus: we houden er geen illusies op na die nuttig zijn.

Nu merkt Wieland in zijn recensie op dat het ‘niet direct duidelijk is hoe hieruit zou moeten volgen dat illusies nooit nuttig kunnen zijn.’ Die conclusie (C) volgt inderdaad niet rechtstreeks uit de Doxastische Catch-22, en Wieland heeft een punt dat het boek niet helemaal duidelijk is over de structuur van het argument.

De Catch-22 sluit niet uit dat nuttige illusies *bestaan*. In zekere zin is het voorkomen van dergelijke illusies zelfs evident. Hier is een extreem voorbeeld, ontleend aan Stephen Stich (1990): ik leef in de waan dat mijn vlucht vertrekt om 10 uur in plaats van om 8 uur. Daardoor kom ik te laat aan op de luchthaven. Mijn vlucht is al vertrokken. Maar terwijl ik in de luchthaven wacht op een volgende verbinding, hoor ik op de radio dat mijn geplande vlucht is neergestort. Wat een geluk dat ik me vergiste over het vertrek! Mijn illusie was in dit geval niet enkel nuttig, ze was *levensreddend*. Maar uit dat punt volgt weinig: ik heb gewoon geluk gehad. Die nuttige illusie kan je niet consequent toepassen. Voor hetzelfde geld had ik een nieuw ticket geboekt voor de volgende vlucht, en was dié vlucht neergestort in plaats van mijn eerdere. In dat geval heeft dezelfde vergissing mijn leven gekost. In het eerste scenario ben ik gewoon een ‘cognitieve gelukzak’: iemand die om slechte redenen een illusie opdoet die, toevallig en zonder dat hij dat voorzien heeft, toch heilzaam blijkt.<sup>1</sup>

<sup>1</sup> Merk op dat zelfs in dit geval de illusie enkel nuttig is omdat ze de schadelijke effecten van een andere illusie weet te compenseren, i.c. de illusie dat je vlucht veilig zal landen. Met een extra portie waarheid kan je het probleem net zo goed en zelfs beter oplossen. Stel dat iemand je ware kennis kan verschaffen over het terroristische complot, of dat je dat met de gave der precognitie kan voorspellen. Dan blijkt dat je je illusie over het vertrek niet langer nodig hebt: je kan gewoon rationeel beslissen, op basis van accurate informatie over het vertrek, dat je voor geen geld op dat vliegtuig stapt.

De Catch-22 is een performatieve paradox: ze sluit niet uit dat nuttige illusies bestaan, maar ze sluit wel uit dat we ze op een (prudentieel) verantwoorde en (conceptueel) coherente wijze kunnen bezitten. In die zin functioneert de *Doxastische Catch-22* als een ultieme troefkaart in mijn betoog tegen nuttige illusies. In het boek geef ik eerst algemene redenen, los van de Catch-22, voor de these dat de meeste illusies gevaarlijk zijn, zelfs als ze aanvankelijk onschuldig lijken. Dat is bijvoorbeeld de moraal van de ‘dodelijke golfballenzoeker’ die Wieland aanhaalt (cf. hoofdstuk 5). De Catch-22 wordt eigenlijk pas in stelling gebracht wanneer we een uitzondering menen te hebben gevonden op die algemene stelregel dat illusies gevaarlijk zijn. Indien we toch een nuttige illusie gevonden hebben, belanden we in het destructieve dilemma dat ik hierboven beschreef. De illusie dat  $p$  kan dan wel nuttig zijn, maar we kunnen ze nooit tezelfdertijd op een conceptueel coherente en prudentieel verantwoorde wijze bezitten. Een weloverwogen keuze voor nuttige illusies is dus uitgesloten door de Doxastische Catch-22. In het licht van het bovenstaande, zou ik de premisse (3) van Wieland dan ook licht aanpassen, als volgt:

(P3) Als we oprecht in  $x$  leven, dan weten we niet of  $x$  nuttig is, en kunnen we de mogelijke schadelijke gevolgen niet inschatten. Mogelijk blijkt  $x$  toch heilzaam te zijn, maar dan hebben we gewoon *geluk* gehad.

De aangepaste conclusie luidt dan als volgt:

(C) Het is (conceptueel) onmogelijk om op een prudentieel verantwoorde wijze nuttige illusies te bezitten.

De redenering over de Doxastische Catch-22 en cognitief geluk kan je ook toepassen op meer realistische en subtiele voorbeelden. Stel dat een doctoranda er bij aanvang van haar onderzoek van overtuigd is dat ze haar doctoraat binnen een termijn van drie jaar kan indienen. In werkelijkheid zal het haar zes jaar kosten. Indien ze dat op voorhand had geweten, zou ze nooit de motivatie gehad hebben om eraan te beginnen. Niettemin is ze op het einde van de rit nog steeds blij dat ze haar doctoraat heeft behaald, en bedenkt ze zich dat haar zelfoverschattende illusie voordelig uitdraaide. Zonder die illusie had ze wellicht nooit de motivatie gevonden. Maar eigenlijk heeft ze gewoon *cognitief geluk* gehad. In een alternatieve wereld die slechts licht verschilt van de hare, sturen die extra drie jaar haar levensplannen helemaal in de war, en beklaat ze zich achteraf dat ze zichzelf zo verkeerd heeft inschat en zoveel tijd heeft verspild.

Om te weten of haar zelfoverschattende illusie nuttig dan wel schadelijk is, zou onze doctoranda bij aanvang de effecten van twee overtuigingen moeten vergelijken: (1) de ware overtuiging dat haar doctoraat zes jaar in beslag zou nemen (2) de foute overtuiging dat ze de klus in drie jaar kan afronden. Maar zelfs indien ze tot het besluit komt dat de foute overtuiging voordelig is, omdat die haar voldoende motivatie schenkt om het werk aan te vatten, zou ze die niet langer kunnen omarmen. De volgende interne monoloog is conceptueel incoherent: 'Als ik vandaag aan mijn doctoraat begin, zal ik pas binnen zes jaar kunnen afronden. Maar bij die gedachte zinkt de moed mij in de schoenen. Als ik er echter in slaag om mezelf wijs te maken dat ik de klus in drie jaar kan afronden, zal ik meer gemotiveerd zijn. Bovendien zal die illusie geen schadelijke effecten hebben. Als het zover is, dan zie ik wel in dat mijn overtuiging fout was'. Deze expliciete vorm van zelfbedrog botst op het principe van de *geloofsonwillekeur* en is daarom conceptueel onmogelijk.

### Uitwegen uit de Catch-22?

De enige mogelijke uitwegen uit de Doxastische Catch-22 zijn volgens mij de volgende: (1) paternalistisch bedrog en (2) gewilde onwetendheid. Die komen respectievelijk in hoofdstuk 4 en 8 van mijn boek aan bod. Maar ook die uitwegen zijn slechts gedeeltelijk succesvol, zo ben ik beginnen inzien. Laten we dat nog even van naderbij bekijken:

1. Bij paternalistisch bedrog speldt iemand anders ons nuttige illusies op. De bedrieger houdt een oogje in het zeil en ziet erop toe dat het met de bedrogene niet fout afloopt. Op het eerste zicht ontsnappen we daarmee aan de paradox, omdat bedrieger en bedrogene niet langer verenigd zijn in dezelfde persoon (daarmee ontsnappen we aan de traditionele filosofische paradox van zelfbedrog). Maar hoe moeten we die relatie tussen bedrieger en bedrogene zien? Geef ik toestemming om in zo'n paternalistische relatie te treden? Stel dat ik een paternalistische dokter heb die me geregeld met een kluitje in het riet stuurt: hij verzwijgt een diagnose, hij deelt alleen de labo-resultaten mee die me niet teveel zullen verontrusten, en hij schrijft af en toe wat homeopathische suikerbolletjes voor om mijn kwaaltjes te verlichten. Nu zijn er twee mogelijkheden. (a) Ofwel handelt de dokter op eigen houtje, zonder mijn voorafgaande toestemming. In dat geval kan je ernstige morele bezwaren uiten. Hoe weet de bedrieger dat ik liever in de waan leef dan de waarheid te kennen? Heb ik als patiënt geen recht op accurate

en waarheidsgetrouwe informatie ten allen tijde? (b) Ofwel – en deze situatie is conceptueel interessanter – heb ik mijn dokter voorafgaandelijk een soort expliciete toestemming gegeven voor een paternalistische houding. Stel dat ik mijn huisdokter tijdens mijn eerste consult ongeveer het volgende vertel: ‘Als patiënt wil ik alleen de waarheid kennen over mijn medische conditie als jij denkt dat dat goed voor mij is. Ik wil graag ook zoveel mogelijk van het placebo-effect profiteren. Ik zou liever hebben dat je me in de toekomst, als ik een of ander medisch probleem heb, enkel die zaken vertelt waar ik me als patiënt goed bij zal voelen en die mijn conditie zullen verbeteren. Zo nodig heb ik dus liever leugentjes om bestwil dan de waarheid.’ In dit scenario stuiten we echter op een indirecte variant van de Doxastische Catch-22: als ik mijn dokter ooit expliciet de toestemming gaf om me eventueel te bedriegen, dan weet ik bij elk bezoek dat de kans op bedrog bestaat. Werkt het bedrog dan nog wel? Als de dokter me bij een volgend consult een doosje pillen mee naar huis geeft, dan zal de volgende gedachte onwillekeurig bij me opkomen: ‘Zijn het deze keer echte pillen, of gaf hij me gewoon placebopillen?’ Die twijfel zal uiteraard niet bevorderlijk zijn voor het placebo-effect, omdat mijn overtuiging dat ik echte pillen slik een lagere subjectieve waarschijnlijkheid krijgt. Om ten volle van het placebo-effect te profiteren, zou ik op een of andere manier selectief moeten *vergeten* dat ik mijn dokter vooraf de toestemming heb gegeven om me te bedriegen. Maar dat botst weer op de geloofsonwillekeur.

2. Hoewel het conceptueel onmogelijk is om voor nuttige illusies te kiezen, kunnen we wel bewust kiezen voor onwetendheid. Ik kan er niet bewust kiezen voor kiezen om bepaalde kennis te loochenen, eens ik eraan werd blootgesteld, maar ik kan er wel bewust voor kiezen om die blootstelling te vermijden. Als ik een religieuze gelovige ben, kan ik bijvoorbeeld besluiten om geen boeken van Richard Dawkins of Bertrand Russell te lezen, uit angst dat die mijn geloof aan het wankelen zouden brengen. Of als ik niet wil weten of mijn vriendin er minnaars op nahoudt, kan ik besluiten om niet door haar briefwisseling te snuisteren. Waarom zou ik mezelf immer kwetsen? Zoals Shakespeares Othello zei: ‘He that is robb’d, not wanting what is stol’n / Let him not know’t, and he’s not robb’d at all.’ Maar in zulke gevallen verraadt de keuze voor onwetendheid al een soort *angstig vermoeden* van de waarheid. Waarom zou ik immers met een boog om atheïstische boeken heen lopen? Is mijn geloof dan zo wankel? En waarom zou ik zo beducht zijn om de briefwisseling van mijn vriendin te lezen? Ben ik bang

dat ik al ‘bestolen’ ben? Als ik echt overtuigd ben van mijn illusie, zou je verwachten dat ik er helemaal niet voor terugdeins om me bloot te stellen aan dat soort informatie. De strategie van ‘bewuste onwetendheid’ verraadt meestal dus een vermoeden van de ware toedracht. Voor zover gewilde onwetendheid conceptueel mogelijk is, resulteert ze niet zozeer in illusies (i.e. foute overtuigingen), maar in een *afwezigheid* van overtuigingen. Neem een filmrecensie van een *whodunnit* die de ontknoping van de film verklapt. Een dergelijke recensie bevat vaak de waarschuwing ‘spoiler alert’, om de argeloze lezer de keuze van gewilde onwetendheid te geven. Eens je een plotbederver hebt gelezen, kan je die kennis immers niet meer ongedaan maken. De meeste mensen willen liever niet weten hoe een spannende film afloopt voordat ze hem zelf gezien hebben (er zijn merkwaardig genoeg uitzonderingen), omdat de verwachtingvolle spanning en onzekerheid bijdragen aan het esthetische plezier van de filmervaring. Maar als je besluit om de recensie niet te lezen, wil dat niet zeggen dat je een of andere ‘illusie’ in stand houdt over de ontknoping van de film. Je hebt gewoon geen idee van die ontknoping, en je wil dat graag zo houden tot het zover is. Gaat de held echt dood op het einde van het verhaal (p), of blijft hij leven (~p)? Voor je naar de bioscoop gaat, geloof je noch het ene, noch het andere.

## Tegenvoorbeeld?

Wieland geeft op het einde van zijn recensie een mogelijke kandidaat voor een ‘nuttige illusie’, die mijn stelling kan ondergraven:

Echter, als je als individu de illusie hebt dat jouw leven interessant is en dat je projecten ertoe doen (in het licht van de zinloosheid van het bestaan, zie pp. 7-8), dan zal dat je helpen om dingen voor elkaar te krijgen waar je anders misschien niet eens aan zou beginnen. Een heilzame illusie?

Los van de Doxastische Catch-22, die een weloverwogen keuze voor nuttige illusies uitsluit, kunnen we ons wel de vraag stellen of deze overtuiging, indien je erdoor bevangen bent, inderdaad (1) een illusie is en (2) heilzame effecten heeft. Wat (1) betreft: in principe hoeft ons leven niet zinloos te zijn, zelfs indien je aanvaardt dat het universum uliem zinloos is. Maar veel hangt natuurlijk af van welke definitie je hanteert van begrippen als ‘ertoe doen’, ‘zinvol’ en ‘interessant’. Dat zijn rekbare begrippen, die geen duidelijke en objectieve maatstaf kennen. Inschattingen van wat ‘in-

teressant' is, lopen uiteen. Daardoor wordt het lastig om hard te maken dat iemand dienaangaande een echte 'illusie' (i.e. foute overtuiging) koestert. Is het een illusie om te denken dat je onderzoek interessant is? Elk persoon hanteert een ietwat eigengereide (*self-serving*) definitie van wat als 'interessant' geldt, waar zijn of haar eigen bezigheden gunstig uitkomen (Dunning, Meyerowitz, and Holzberg 1989). Een objectiever criterium in dat verband is 'interessant bevonden worden door anderen'. Vinden anderen dat je werk ertoe doet of niet?

Hier kunnen we ons makkelijk inbeelden dat iemand illusies koestert. Stel dat je je verbeeldt dat je filosofische collega's je werk interessant vinden, omdat ze schijnbaar interesse betonen in je bezigheden, en sociaal wenselijke antwoorden geven als je naar hun mening vraagt. In werkelijkheid echter zitten ze in de koffiekamer met hun ogen te rollen en vinden ze je hele onderzoek tijdverspilling. De overtuiging dat anderen je werk hoog inschatten, is vanzelfsprekend een aangename gewaarwording voor de persoon in kwestie (Taylor 1989). Maar is het ook een *nuttige* illusie? Dat betwijfel ik. Per slot van rekening zijn het je vakgenoten die de kwaliteit van je werk vroeg of laat moeten beoordelen, in wetenschappelijke commissies of als anonieme lezers bij academische tijdschriften. Daar zullen ze wél hun ongezoeten mening geven. Het systeem van blinde collegiale toetsing dient tenslotte om ons tegen positieve illusies te beschermen. Het is beter dat die tijdig doorgeprikt worden. Als je in de waan leeft dat je eigen werk baanbrekend en geweldig interessant is, ben je helemaal niet voorbereid op de onvermijdelijke ontzuivering. Wie tijdig kritische feedback krijgt, zal misschien eerst ontgoocheld zijn, maar kan wel iets van de kritiek opsteken. En als je werk gewoon zo hopeloos is dat kritische feedback niets zou uithalen? In dat geval kan je je ernstige vragen stellen bij de zinvolheid van een academische carrière. De werkelijkheid is weerbarstig, en vroeg of laat zullen je illusies er toch tegenaan botsen, hoe prettig ze ook aanvoelen. Weet je dan niet beter tijdig waaraan je toe bent?

## Bibliografie

- Adler, J. (2002) *Belief's own ethics*. London: MIT Press.
- Boudry, M. (2015) *Illusies voor gevorderden. Of waarom waarheid altijd beter is*. Antwerpen: Uitgeverij Polis.
- Dunning, D., Judith, A.M. en Holzberg, A.D. (1989) 'Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability', *Journal of Personality and Social Psychology* 35, pp.1082-1090.
- Heller, J. 1994 [1962] *Catch 22*. London: Vintage Books.



- Stich, S.P. (1990) *The fragmentation of reason: preface to a pragmatic theory of cognitive evaluation*. Cambridge (Mass.): MIT press.
- Taylor, S.E. (1989) *Positive illusions: Creative self-deception and the healthy mind*. Basic Books: New York.
- Wieland, J.W. (2016) 'Recensie "Illusies voor gevorderden"', *Algemeen Nederlands Tijdschrift voor Wijsbegeerte*, 108:4, pp. 563-65.

## Over de auteur

**Maarten Boudry** (1984) is wetenschapsfilosoof aan de UGent en auteur van *Illusies voor gevorderden. Of waarom waarheid altijd beter is* (Polis, 2015). Eerder publiceerde hij *De ongelovige Thomas heeft een punt* (Houtekiet, 2011) samen met Johan Braeckman. Binnenkort verschijnt bij The University of Chicago Press een bundel getiteld *Science Unlimited? On the challenges of scientism*, die hij samenstelde met de filosoof en bioloog Massimo Pigliucci.