



## UvA-DARE (Digital Academic Repository)

### Conventies en moraal

Verbeek, B.

**Publication date**

2009

**Document Version**

Final published version

**Published in**

Als vuur: opstellen voor Govert den Hartogh ter gelegenheid van zijn emeritaat

**License**

Other

[Link to publication](#)

**Citation for published version (APA):**

Verbeek, B. (2009). Conventies en moraal. In P. Rijkema, G. van Donselaar, B. Verbeek, & H. Wijsbek (Eds.), *Als vuur: opstellen voor Govert den Hartogh ter gelegenheid van zijn emeritaat* (pp. 143-160). Boom Juridische Uitgevers.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Conventies en moraal

Bruno Verbeek\*

## Inleiding

In 1985 promoveerde Govert den Hartogh op een indrukwekkend proefschrift, getiteld *Wederkerige verwachtingen*. Daarin ontvouwde Den Hartogh zijn conventietheorie van sociale normen. Sociale normen zijn de formele en informele regels van het intermenselijk verkeer. Te denken valt hierbij aan bepaalde gewoonten en tradities, maar ook aan wetten en de regels van de (smalle) moraal.<sup>1</sup> Al deze regels werden door Den Hartogh geanalyseerd als patronen van wederkerige verwachtingen over elkaars gedrag, die als eigenschap hebben dat het voor eenieder individueel redelijk is om zich naar dit patroon te voegen. Dergelijke patronen duidde hij aan als conventies.

In deze bijdrage zal ik een aantal elementen van het conventionalisme van Den Hartogh kritisch tegen het licht houden. Dat zou als intellectuele vadermoord kunnen worden opgevat, maar ik verzeker u dat dit voortkomt uit mijn diepste respect en fascinatie voor het conventionalisme en vooral de versie van Den Hartogh.

## Het conventionalisme van Den Hartogh

Een goed voorbeeld van een conventie is de regel om rechts te houden in het verkeer. Stel dat op een smalle landweg twee voertuigen, A en B, elkaar naderen. Om elkaar te passeren moeten de bestuurders van beide voertuigen uitwijken. Elke bestuurder moet besluiten naar rechts of naar links uit te wijken. Het maakt beiden uiteraard niet uit aan welke kant ze elkaar passeren, zolang ze elkaar maar passeren, liefst zonder ongelukken en oponthoud. Heeft A een reden om naar rechts uit te wijken? Ja, maar alleen voor zover hij ervan uit mag gaan dat B ook naar rechts zal uitwijken. Is er een reden voor B om dat te doen? Ja, maar alleen voor zover B ervan uit mag gaan dat A naar rechts zal uitwijken. Kortom, de redenen van A en B om naar rechts uit te wijken verwijzen naar elkaar; ze zijn *interdependent*. A heeft reden om naar rechts uit te wijken alleen voor zover B dat heeft, en vice versa. Stel nu dat A en B reeds van elkaar kunnen verwachten

---

\* Hartelijke dank is verschuldigd aan Peter Rijkema voor diens commentaar op een eerdere versie van dit essay.

1. Met 'smalle moraal' bedoel ik het geheel van regels dat het samenleven überhaupt mogelijk maakt. Het moet worden onderscheiden van de regels van 'brede moraal', die onder meer aspecten van idealen van het goede leven betreffen. Zie ook Mackie, 1977, hoofdstuk 5.

dat de ander rechts zal houden, bijvoorbeeld omdat er precedents zijn uit het verleden, of omdat dit in deze gemeenschap een algemeen gedeelde verwachting is in dit soort situaties (misschien is deze verwachting wel publiekelijk uitgesproken en neergelegd in de Wegenverkeerswet), dan heeft zowel A als B voldoende reden om te besluiten rechts te houden. De algemeen gedeelde verwachting geeft doorslaggevende redenen, ja, *verplicht* in dat geval A en B rechts te houden.<sup>2</sup> Zo'n algemeen gedeelde verwachting is een *conventie*. De structuur van dergelijke algemeen gedeelde verwachtingen laat zich beschrijven als patronen van verwachtingen ten aanzien van gedrag en verwachtingen over dergelijke verwachtingen over en weer. Den Hartogh spreekt in dit verband van 'patronen van wederkerige verwachtingen'.

Conventies geven redenen, interdependente redenen, aan de leden van de gemeenschap om te doen zoals de conventie bepaalt. Het is rationeel om zo'n regel te gehoorzamen. Terugkerend naar het voorbeeld, kunnen we concluderen dat de regel om rechts te houden in het verkeer een conventie is: een patroon van wederkerige verwachtingen in die gemeenschap dat men rechts houdt in dit soort gevallen.

Dit soort patronen is niet alleen te vinden in situaties waarin er geen belangenconflict is tussen de betrokkenen, zoals in het verkeersvoorbeeld van hiervoor. Ook in situaties waar dat wel zo is, zoals het beruchte *prisoner's dilemma*, kunnen dit soort patronen van verwachtingen opkomen die de betrokkenen reden geven zich op een bepaalde manier te gedragen.<sup>3</sup> In al die situaties hebben de betrokkenen er belang bij dat er op de een of andere manier een onderlinge afstemming van gedrag wordt gerealiseerd. Conventies brengen dergelijke afstemming tot stand. Den Hartoghs stelling in *Wederkerige verwachtingen* is dat morele regels, net als conventies, als patronen van wederkerige verwachtingen opgevat moeten worden. Morele regels zijn patronen van wederkerige verwachtingen die rationele actoren interdependente redenen geven om te handelen zoals de regel bepaalt.

Voor velen is deze suggestie niet eens de moeite van het overdenken waard. Immers, conventies, zoals de verkeersregels of de regels van etiquette, zijn volslagen contingent en arbitrair: we moeten rechts houden in het verkeer – het had evengoed links kunnen zijn. Nederlanders kussen vrouwelijke bekenden drie keer – maar het had evengoed anders kunnen zijn. Morele regels daarentegen, zeker de fundamentele regels van de smalle moraal, zoals de regels die geweld en doodslag verbieden, zijn niet contingent in deze zin. Het had niet anders kunnen zijn: moraal verbiedt deze zaken noodzakelijkerwijze.

2. Hier dient zich de vraag aan wat precies met de term 'verplichting' wordt bedoeld in dit verband. Het project van Den Hartogh in *Wederkerige verwachtingen* is om deze notie te expliciteren. Verderop in dit essay bespreek ik Den Hartoghs definitie.

3. Den Hartogh (1985, hoofdstuk 9 en 13-14) maakt veel werk van het zogenoemde 'verzekeringsspel' en het *prisoner's dilemma*. Zie ook Schotter, 1981; Sugden, 1986; Skyrms, 1996, 2004. In Verbeek (2007) geef ik mijn versie van het conventionalisme aan de hand van het 'hawk-dove'-spel.

Den Hartogh heeft zelf zowel in *Wederkerige verwachtingen* als in later werk veel gedaan om deze kritiek te weerleggen. In deze bijdrage wil ik ingaan op een paar aspecten van deze weerlegging. Eerst zal ik kort enkele essentiële punten van het conventionalisme bespreken. Vervolgens zal ik het bezwaar van contingentie preciseren, om daarna twee tegenargumenten, die in een of andere vorm ook door Den Hartogh worden gehanteerd, kritisch tegen het licht te houden. Beide argumenten slagen er niet in om het verwijt van contingentie helemaal te weerleggen.

## Interdependentie

Laten we de conventie om rechts te houden in het verkeer nog eens nader beschouwen. We kunnen de situatie van beide bestuurders analyseren als een speltheoretisch coördinatieprobleem. Elk van beide bestuurders heeft de keuze uit twee mogelijke alternatieven: rechts houden of links houden. Ik neem aan dat beide bestuurders de gebruikelijke preferenties hebben van bestuurders in dit soort situaties: beiden willen veilig en snel thuiskomen, maar zullen indifferente zijn of ze dit doen door beiden rechts te houden of door beiden naar links uit te wijken. Dit levert de matrix van het pure coördinatie spel op:

		B	
		Rechts	Links
A	Rechts	(1, 1)	(0, 0)
	Links	(0, 0)	(1, 1)

Figuur 1

In deze matrix is de interdependentie van de redenen van A en B om rechts respectievelijk links te houden verdisconteerd. Stel dat A kan verwachten dat B rechts houdt. Dan heeft hij een reden om ook rechts te houden. Als A rechts zal houden, dan heeft B een reden om rechts te houden. A noch B heeft een onafhankelijke reden om rechts te houden – onafhankelijk van de redenen van de ander. Immers, als B links zal houden, heeft A ook reden links te houden, en dat geeft B weer reden links te houden. In de afwezigheid van enige onafhankelijke grond om te verwachten dat de ander rechts dan wel links zal houden, zullen A en B er niet uit kunnen komen. Zodra er voor een (of meerdere) van de betrokkenen een onafhankelijke reden is om rechts te houden, is de interdependentie van redenen verdwenen.

		B	
		Rechts	Links
A	Rechts	(1, 1)	( $\frac{1}{2}$ , 0)
	Links	(0, 0)	(0, 0)

Figuur 2

Stel dat de situatie van A en B die van figuur 2 zou zijn.<sup>4</sup> Nu heeft A reden om rechts te houden, ongeacht wat B doet. Immers, als hij verwacht dat B rechts houdt, realiseert hij zijn beste uitkomst, en als hij meent dat B links zal houden, realiseert hij nog altijd een betere uitkomst dan wanneer hij ook links houdt. A heeft dan een onafhankelijke reden om rechts te houden en als B veilig thuis wil komen, houdt hij ook rechts. B's reden om rechts te houden is afhankelijk van de reden van A, maar er is geen interdependentie van redenen.

Het is duidelijk dat interdependentie van redenen er alleen kan zijn wanneer er ten minste een andere combinatie van interdependente redenen is. Immers, stel dat het probleem van de bestuurders er als volgt uitziet:

		B	
		Rechts	Links
A	Rechts	(1, 1)	(0, 0)
	Links	(0, 0)	(0, 0)

Figuur 3

Nu prefereren beide bestuurders de situatie waar ze elkaar rechts passeren boven alle andere mogelijkheden en zijn ze volslagen indifferent tussen die andere mogelijkheden. Nu heeft A reden om rechts te houden. Als hij verwacht dat B links zal houden, heeft hij geen bijzondere reden om ook links te houden. Wat hij ook doet, de verwachte uitkomst is even goed. Hij zou dus evengoed rechts kunnen houden. Zelfs als hij twijfelt of B rechts zal houden, is het beter voor hem om rechts te houden. Er is geen reden voor A om links te houden, ongeacht wat B zal doen, terwijl er wel redenen zijn om rechts te houden, onafhankelijk van A's verwachtingen ten aanzien van B. Dat wil zeggen dat er in dit geval geen interde-

4. Uiteraard is een dergelijke matrix niet goed voorstelbaar waar het twee bestuurders op een smalle weg betreft. Dat is hier echter niet belangrijk, aangezien ik alleen dit punt over de afwezigheid van interdependentie wil maken.

pendentie van redenen meer is, en dat komt omdat er geen alternatieve uitkomst is die door interdependente redenen wordt ondersteund.<sup>5</sup>

Het is overigens niet zo dat die interdependentie alleen bij gelijkwaardige alternatieven tot stand komt. In figuur 4 blijkt dat ook bij suboptimale evenwichten er sprake is van interdependentie van redenen.<sup>6</sup>

		B	
		Rechts	Links
A	Rechts	(1, 1)	(0, 0)
	Links	(0, 0)	( $\frac{1}{2}$ , $\frac{1}{2}$ )

Figuur 4

Nu prefereert zowel A als B de uitkomst waarbij beiden rechts houden, maar, in tegenstelling tot figuur 3, hebben ze reden om links te houden voor zover er omdat de ander reden heeft om links te houden. (Als A links houdt, heeft B een reden om links te houden, en vice versa.)

### Opvallendheid en verwachtingen van een hogere orde

Wat moeten A en B nu doen om uit het probleem van figuur 1 te komen? Hoe kunnen we begrijpen dat echte bestuurders geen probleem hebben om in dergelijke gevallen elkaar rechts, dan wel links, te passeren? Er zijn diverse oplossingen geformuleerd in de loop der tijd. Niet al deze oplossingen voldoen in de ogen van Den Hartogh. In *Wederkerige verwachtingen* stelt hij een strikte voorwaarde: de oplossing moet recht doen aan de individuele rationaliteit en intentionaliteit van de deelnemers. Dat lijkt een heel plausible voorwaarde. Ze sluit bij voorbaat een aantal kandidaten uit. Zo verworpt hij alle vormen van functionalisme waarbij geslaagde coördinatie wordt verklaard vanuit de nuttige functie die dergelijke coördinatie heeft (zie bijvoorbeeld Den Hartogh, 1985, hoofdstuk 25). Dit lijkt relatief onproblematisch. Tal van filosofen zijn kritisch over dergelijke verklarin-

5. De matrix in figuur 3 is een grensgeval. Zodra de nutswaarden in de uitkomst 'links, links' kleiner worden dan (0, 0), hebben A en B beiden onafhankelijke redenen om rechts te houden: door rechts te houden vermijden ze in dat geval hun slechtste uitkomst. Zodra de nutswaarden van 'links, links' groter dan (0, 0) worden, is er weer sprake van interdependentie van redenen.
6. Kortom, is er sprake van interdependentie van redenen, dan en slechts dan is er meer dan een 'eigenlijk' evenwicht (*proper equilibrium*). Een evenwicht is 'eigenlijk' indien het zodanig is dat iedere deelnemer het strikt prefereert boven een andere uitkomst die bereikt had kunnen worden, gegeven de keuze van alle andere deelnemers (Lewis, 1969, p. 22).

gen.<sup>7</sup> Echter, deze voorwaarde maakt ook dat de hedendaagse standaardtheorie van conventies, de zogenoemde evolutionaire speltheorie, niet in aanmerking komt.<sup>8</sup> Deze theorie gaat namelijk niet uit van individueel rationele spelers.

Terug naar de oorspronkelijke vraag: hoe kunnen we begrijpen dat rationale actoren in de situatie van figuur 1 beiden rechts zullen houden? Op dit punt aangekomen, is het interessant om de theorie van Den Hartogh te vergelijken met die van Lewis (1969). Lewis betoogde, in navolging van Schelling (1960), dat rationale actoren in de situatie van figuur 1 een beroep kunnen doen op bepaalde kenmerken die een uitkomst op doen vallen en eruit laten springen. Zo zou het kunnen zijn dat rechts houden eruit springt omdat er een precedent is dat verkeersdeelnemers rechts houden, of omdat er een afspraak is gemaakt rechts te houden. Het zou ook kunnen zijn dat (sommige) betrokkenen gezamenlijk rechts houden prefereren boven gezamenlijk links houden. Maar, zoals Den Hartogh zelf later ook betoogde, het feit dat de wetgever heeft geboden om rechts te houden, is ook een signaal dat maakt dat rechts houden eruit springt.<sup>9</sup> Dergelijke zaken maken rechts houden opvallend.

Stel nu dat rechts houden opvallend is op een van de hiervoor genoemde manieren. Hoe kunnen we nu verklaren dat een dergelijke 'opvallendheid' maakt dat rationale bestuurders hun handelingen kunnen coördineren? Lewis reconstrueerde de redenering van de betrokkenen als volgt:<sup>10</sup>

- (1) A heeft reden naar rechts te gaan, als hij gelooft dat B reden heeft om naar rechts te gaan.
- (2) A gelooft dat B reden heeft om naar rechts te gaan.
- (3) Ergo: A heeft een reden om naar rechts te gaan.

Deze redenering is geldig. Premisse (1) is niet opzienbarend. Deze volgt uit de beschrijving van het probleem in figuur 1 en de aanname dat A rationeel is. (2) is daarentegen problematisch. Waarom mag A dat geloven? Welnu, zegt Lewis, A mag dat geloven omdat A een tweede-ordeverwachting heeft over B's opvattingen over de redenen van A:

- (4) A gelooft dat B reden heeft om naar rechts te gaan als B gelooft dat A reden heeft om naar rechts te gaan.
- (5) A gelooft dat B gelooft dat A reden heeft om naar rechts te gaan.
- (6) Ergo: A gelooft dat B reden heeft om naar rechts te gaan.

7. Bijvoorbeeld Hempel, 1959; Elster, 1989. Voor een riposte, zie Cohen, 1982.

8. Voor een karakterisering van deze versie van de speltheorie (zie Sugden, 1986; Skyrms, 1996, 2004). Verbeek & Morris (2008) geven een overzicht.

9. Den Hartogh, 2002. Vergelijk ook Vanderschraaf, 1995.

10. Lewis, 1969, p. 28-31. De formulering is de mijne om het bij het voorbeeld van figuur 1 te laten passen.

Dat is precies stap (2) uit de voorafgaande redenering. Opnieuw, de redenering is geldig. (4) is niet opzienbarend, maar (5) is de cruciale stap die tot de conclusie leidt. (5) is een tweede-ordegeloof: het is een geloof over een geloof. Is (5) aannemelijk? Alleen als er een gelijksoortig argument als (4)-(6) voor gegeven kan worden met een derde-ordegeloof in de premissen. En dat derde-ordegeloof moet ook weer worden gerechtvaardigd door een argument met een vierde-ordegeloof, enzovoort, enzovoort.

Nu is er een aantal zaken dat opvalt aan dit hele argument tot nu toe. Lewis neemt duidelijk aan dat hogere-ordeverwachtingen lagere-ordeverwachtingen kunnen rechtvaardigen: (5) is in feite de rechtvaardiging voor (2). Den Hartogh neemt dit element van Lewis' theorie over (bijvoorbeeld Den Hartogh, 1985, p. 5). Dat kan natuurlijk alleen indien de rechtvaardiging van die hogere-ordeverwachtingen onafhankelijk is van de rechtvaardiging van de lagere-ordeverwachtingen. Je kunt alleen concluderen dat Socrates sterfelijk is als je onafhankelijke redenen hebt om aan te nemen dat alle mensen sterfelijk zijn en dat Socrates een mens is. Op precies dezelfde manier is een argument voor een geloof van de  $n$ -de orde alleen aannemelijk indien voor de premisse met een geloof van de  $n+1$ -de orde een onafhankelijke grond is. Waar komt die onafhankelijke grond vandaan? Als die er niet is, dan is de hele constructie op drijfzand gebouwd.

Lewis meende dat er een onafhankelijke grond voor de hogere-ordeverwachtingen is. Dat maakt tevens dat de regressie die ik hiervoor signaleerde, niet problematisch is. Lewis betoogde dat de 'common knowledge', gemeenschappelijk gedeelde kennis, van de opvallendheid van een van de alternatieven de rechtvaardiging van de hogere-ordeverwachtingen oplevert. Om precies te zijn, de opvallendheid van rechts houden is gemeenschappelijk gedeelde kennis volgens Lewis dan en slechts dan als (Lewis, 1969, p. 52):

- (1) Zowel A als B reden heeft om te geloven dat rechts houden opvallend is.
- (2) Deze opvallendheid A en B aanleiding geeft dat ze beiden reden hebben om te geloven dat rechts houden opvallend is.
- (3) De opvallendheid van rechts houden A en B aanleiding geeft te geloven dat ze elk voor rechts houden zullen kiezen.

De combinatie van deze drie kenmerken van gemeenschappelijk gedeelde kennis, alsmede de gebruikelijke aannames van de individuele rationaliteit van A en B, is de basis voor ieder hogere-ordegeloof. We kunnen nu immers concluderen dat:

- (4) A reden heeft te geloven dat B rechts zal houden.
- (5) A reden heeft te geloven dat B gelooft dat A gelooft dat B rechts zal houden (en vice versa).
- (6) Enzovoort, enzovoort.

Met andere woorden, A heeft reden een *n*-de-ordegeloof over B's redenen om rechts te houden te accepteren, als rechts houden opvallend is. Op deze wijze laat Lewis zien dat de opvallendheid van rechts houden (deel van) de reden is voor A om rechts te houden.

Den Hartogh is niet overtuigd door dit argument, en terecht. Immers, als de opvallendheid van rechts houden een reden vormt om rechts te houden, dan is het hele bouwwerk van hogere-ordeverwachtingen overbodig (Den Hartogh, 1985, p. 80-81 en 94-95). Het feit dat rechts houden opvallend is, is voldoende reden om rechts te houden voor de betrokkenen. Bovendien is opvallendheid in dat geval een onafhankelijke reden om rechts te houden en geen interdependente reden. De reden voor A om rechts te houden is dat dit opvallend is – niet dat B reden heeft om rechts te houden!

Den Hartogh betoogt in *Wederkerige verwachtingen* dat we het beste dit idee van Lewis op kunnen geven en alleen het eerste inzicht, dat hogere-ordeverwachtingen lagere-ordeverwachtingen kunnen rechtvaardigen, overnemen. Hoe zit het dan met het bezwaar dat dit alleen opgaat indien die hogere-ordeverwachtingen een onafhankelijke basis hebben? Welnu, stelt Den Hartogh, dat hoeft helemaal niet. Het is genoeg dat die verwachtingen bestaan. Er is geen reden iedere keer individueel het hele patroon van hogere-ordeverwachtingen te reconstrueren:

'Wanneer (...) [coördinatie eenmaal geslaagd is] (...), zullen wij van elkaar in een volgende analoge keuzesituatie de keuzen verwachten die opnieuw tot die ontmoeting leiden. Maar daartoe hoeven we niet als vensterloze monade de praktische redenering van de andere die tot die verwachting leidt te repliceren.' (Den Hartogh, 1985, p. 91)

En:

'(...) als het patroon van verwachtingen eenmaal self-supporting is, dan zijn het de verwachtingen van de hogere orde die die van de lagere orde ondersteunen, en niet omgekeerd. Verwachtingen van [een lagere orde] kunnen gefundeerd worden op een onafhankelijke grond, óf gelegitimeerd worden door verwachtingen van [een hogere] orde, die niet zelf uit de lagere-orde verwachtingen zijn afgeleid. Het laatste is het geval bij konventies.' (Den Hartogh, 1985, p. 101)

Hiermee hebben we de drie centrale ideeën van het conventionalisme van Den Hartogh besproken:

1. Sociale normen, inclusief morele normen, zijn net als conventies patronen van interdependente verwachtingen over en weer, die individueel gedrag conform deze verwachtingen rechtvaardigen.
2. Dat doen ze doordat bij dergelijke verwachtingen hogere-ordeverwachtingen lagere-ordeverwachtingen rechtvaardigen: ik mag ervan uitgaan dat jij rechts

zult houden, omdat ik verwacht dat jij verwacht ... dat ik verwacht dat jij verwacht dat ik rechts zal houden.

3. Dat patroon van hogere-ordeverwachtingen, dat Den Hartogh consequent als 'konventie' aanduidt, hoeft niet iedere keer opnieuw door de betrokken individuen te worden afgeleid of te worden gerechtvaardigd.<sup>11</sup>

### Kritiek

Dit stelt ons in staat om de kritiek op het conventionalisme die ik in de inleiding aanstipte, naar waarde te schatten. Deze luidde dat in de conventionalistische theorie normen arbitrair en contingent lijken te zijn. Dat zijn ze uiteraard niet. Wellicht dat het tot op zekere hoogte contingent is of er links dan wel rechts in het verkeer gehouden zal worden, maar dergelijke contingentie is niet van toepassing op morele normen zoals het verbod op moord, geweld en andere fundamentele zaken.

De analyse laat zien dat het wel meevalt met de mate van contingentie en willekeur. Zelfs in het geval van rechts houden in het verkeer is het niet zo dat willekeurig welke regel zou kunnen bestaan. Een regel die eist dat verkeersdeelnemers bij iedere ontmoeting elk een muntje opgooien en rechts houden als kop boven komt, zal niet worden nageleefd. Zo'n regel is namelijk geen oplossing voor het onderliggende coördinatieprobleem. Evenzo zal een regel die eist dat verkeersdeelnemers elkaar rechts passeren op dagen waarvan de som van de getallen van de datum een priemgetal is, en links passeren op andere dagen, niet worden nageleefd. De regel is te ingewikkeld en zal te vaak tot vergissingen en daarmee gepaard gaande ongelukken leiden. Kortom, net als de basale morele normen zijn conventies niet arbitrair in deze radicale zin.

Bovendien hebben veel morele normen wel degelijk conventionele elementen – evengoed als in conventies, zoals de verkeersregels. Het verbod op moord verbiedt het opzettelijk doden van onschuldige slachtoffers. Merk op dat dit allerlei zaken ongespecificeerd laat. Wat telt als opzettelijk doden? Wat maakt een slachtoffer onschuldig? Is deze norm zonder uitzonderingen? Geldt ze voor iedereen? Een groot aantal van deze zaken is moreel gesproken alles behalve duidelijk, en wellicht is er variatie op deze of andere dimensies mogelijk. Het lijkt er eerder op dat het morele verbod op moord een formeel verbod is, waarbij de substantiële inhoud grote variatie kent. Dergelijke onderdeterminatie en variatie is nog veel groter wanneer we naar normen van rechtvaardigheid kijken, bijvoorbeeld op het

11. Den Hartogh spreekt vaak alsof de wederkerigheid van verwachtingen, de inbedding van lagere-ordeverwachtingen in hogere-ordeverwachtingen en de interdependentie van redenen dezelfde noties zijn. Dat is natuurlijk niet zo. Wederkerigheid van verwachtingen beschrijft het netwerk van verwachtingen over en weer in een groep; de inbedding zegt iets over de relatie tussen de verschillende opvattingen over het gedrag van anderen van een individu; interdependentie, ten slotte, is een notie die de grond van een reden voor handelen betreft.

gebied van eigendom, overeenkomsten en politieke verplichting. Kortom, als het verwijt aan het conventionalisme is dat morele normen arbitrair en toevallig zijn volgens deze theorie, is het antwoord dat dit maar ten dele het geval is, en voor zover de theorie een zekere mate van variatie toelaat, deze variatie aannemelijk is in het licht van de praktijk van onze morele normen.

Wellicht dat het eigenlijke verwijt een ander is. De centrale stelling van het conventionalisme is dat sociale regels – en morele normen zijn een soort sociale regels in de opvatting van Den Hartogh – interdependente redenen geven aan de betrokken individuen. We hebben hiervoor gezien dat dit alleen kan als er een alternatief is voor de regel in kwestie. Dat kan een volslagen gelijkwaardig alternatief zijn, zoals bij het coördinatieprobleem in figuur 1, maar het kan ook een inferieur alternatief zijn, zoals in figuur 4. Echter, als het zo is dat er eigenlijk geen alternatief is voor een bepaalde norm (en wellicht dat de norm die moord verbiedt, er zo een is), dan lijkt het erop dat de redenen die deze norm geeft, geen interdependente redenen zijn, maar onafhankelijke redenen. Dat betekent dan weer dat in voorkomende gevallen een individu zich niet af hoeft te vragen wat hij kan verwachten van zijn omgeving om vast te stellen dat er een reden is, wellicht zelfs een doorslaggevende reden, om af te zien van het plegen van een moord. Nog anders gezegd, het conventionalisme lijkt te impliceren dat het redengevende karakter van een norm afhangt van de sociale praktijken en omgeving waarin deze norm functioneert. Dat mag misschien zo zijn voor verschillende niet-morele sociale normen, zoals de regels van etiquette, maar morele normen steken niet zo in elkaar.<sup>12</sup>

We kunnen dit verwijt uiteenrafelen in twee onderdelen. Het eerste betreft de vraag waarom een individu een norm behoort te gehoorzamen. Het tweede deel van het verwijt betreft de vraag wat de norm die moord en doodslag verbiedt, rechtvaardigt. Voor de criticus hangen deze twee vragen samen. Als er geen goede rechtvaardiging voor de norm is, zo meent deze, dan is er ook geen reden voor een individu om deze te gehoorzamen. Het conventionalisme heeft een antwoord op de eerste vraag – al zal dat de tegenstanders niet aanstaan. Volgens het conventionalisme behoort men een norm te gehoorzamen (mede) omdat (voldoende) anderen dat doen. Hoe zit het dan met de tweede vraag? Hoe rechtvaardigt het conventionalisme een morele norm?

### Rechtvaardiging van de norm

Een conventionalistische rechtvaardiging van een morele norm zal aan een aantal criteria moeten voldoen. Op de eerste plaats moet het zo zijn dat de rechtvaardiging recht doet aan de individuele rationaliteit en intentionaliteit van degenen

12. Den Hartogh ontkent dit expliciet. Morele normen zijn voor hem sociale normen. Etiquette is niet essentieel anders dan moraal: beide zijn verplichtend. Bijvoorbeeld Den Hartogh, 1985, p. 7 en 393-394.

die de norm volgen. Op de tweede plaats zal het antwoord de criticus gerust moeten stellen en laten zien dat de rechtvaardiging van de norm zelf niet berust op toevallige sociale praktijken en omstandigheden. Ten slotte moet een conventionalistisch antwoord geen afbreuk doen aan het antwoord op de eerste vraag. De rechtvaardiging van een norm moet zo zijn dat deze nog steeds gekarakteriseerd kan worden als een netwerk van wederkerige verwachtingen dat interdependente redenen voor handelen geeft aan de betrokkenen, en geen onafhankelijke redenen.

Hoe kan een conventionalist nu een norm rechtvaardigen? Ik zal dit bespreken aan de hand van de fundamentele morele norm die moord verbiedt. Een eerste suggestie is dat er goede gronden zijn die deze norm rechtvaardigen. Moord is slecht; het berokkent slachtoffers en nabestaanden grote schade. Niemand, behalve de ergste psychopaat, zal dit ontkennen. Den Hartogh lijkt deze rechtvaardiging te onderschrijven.

‘Ik heb in dit boek verplichtingen geïdentificeerd met legitieme verwachtingen. Men zou daartegen de volgende objectie naar voren kunnen brengen. Misleiding, marteling, moord zijn niet daarom laakbaar omdat wij zulke daden niet van elkaar verwachten, maar omdat zij anderen schaden. Het zal nu duidelijk zijn wat ik op die tegenwerping te antwoorden heb. Ik heb niet beweerd dat iets pas een kwaad is als er een gerechtvaardigde verwachting bestaat het niet teweeg te brengen, en evenmin dat zo'n gerechtvaardigde verwachting de enige geldige reden zou zijn om zich van het teweeg brengen van dat kwaad te houden. Dat iets reeds in zichzelf een kwaad is, verklaart integendeel mede waarom wij van elkaar gaan verwachten het niet te veroorzaken. Maar dat moet wel gebeuren wil er sprake zijn van een verplichting. De plichtenleer is beperkter dan de goederenleer, er bovendien van afhankelijk, maar er niet mee identiek.’ (Den Hartogh, 1985, p. 393)

Een conventionalist, geïnspireerd door dit citaat, zou deze slechtheid kunnen opvatten als individuele waardering van de betrokkenen voor een uitkomst van de matrix. In plaats van figuur 1 levert dit de volgende matrix op:<sup>13</sup>

		B	
		Onthoudt zich van moord	Moord & doodslag
A	Onthoudt zich van moord	(1, 1)	(0, 0)
	Moord & doodslag	(0, 0)	(-1, -1)

Figuur 5

13. Ik heb de labels die de handeling van rij- en kolomkiezer identificeren, aangepast.

De nutsgetallen in deze figuur staan nu voor de relatieve morele waarde van de resulterende situaties. Het moreel gesproken beste is als A en B zich beiden onthouden van moord en doodslag. Het een na beste is indien slechts een van de twee zich hieraan overgeeft, en de situatie waarin beiden moordend door het leven gaan, is, moreel gesproken, een ramp. In dit geval zou A zich moeten onthouden van moord. Immers, stel dat B zich ook onthoudt van moord, dan wordt (1, 1) gerealiseerd. Als B niet afziet van moord, dan wordt (0, 0) gerealiseerd, hetgeen nog altijd moreel gesproken beter is dan (-1, -1). Echter, in deze situatie heeft A geen interdependente reden om af te zien van moord. Hij heeft reden daarvan af te zien, ongeacht wat B doet. Nu zal de criticus deze conclusie toejuichen, maar een conventionalist kan dit niet accepteren en zal vasthouden aan de gedachte dat de verplichting die een sociale of morele norm oplegt, op interdependente redenen berust. Dit kan dus niet de rechtvaardiging zijn in een conventionalistische theorie voor de norm die moord verbiedt.

Een tweede suggestie hoe een conventionalist het morele gebod tegen moord zou kunnen rechtvaardigen, is in feite te vinden bij Lewis. We hebben gezien hoe voor Lewis een conventie individuele redenen voor handelen geeft aan de betrokkenen. Het feit dat ik verwacht dat jij verwacht dat ik verwacht dat ... jij rechts zult houden, samen met mijn belang bij het veilig passeren op de weg, geeft mij reden om rechts te houden. Conventies zijn niets anders dan stabiele patronen van zulke hogere-ordeverwachtingen die gedeelde kennis zijn in een gemeenschap. Wat rechtvaardigt die wederkerige verwachtingen dan? Met andere woorden, wat rechtvaardigt de conventie? Lewis' antwoord was dat gemeenschappelijk gedeelde kennis van de opvallendheid van rechts houden de conventie rechtvaardigt. We kunnen Lewis' antwoord toepassen op het verbod op moord. De morele waarde van een situatie waarin moord en doodslag zijn verbannen, is moreel gesproken opvallend.

Met andere woorden, we zouden kunnen aannemen dat de situatie meer op die van figuur 6 lijkt, waarbij de combinatie waar beiden afzien van moord en doodslag opvallend is om redenen die buiten de matrix liggen.

		B	
		Onthoudt zich van moord	Moord & doodslag
A	Onthoudt zich van moord	(1, 1)	(0, 0)
	Moord & doodslag	(0, 0)	(1, 1)

Figuur 6

Het is belangrijk om te zien dat deze suggestie de criticus een heel eind tevreden zou moeten stellen. Immers, nu is, net zoals bij de conventie rechts te houden, de verplichting af te zien van moord afhankelijk van de rechtvaardiging van de

conventie. Op het moment dat de norm tegen moord niet gerechtvaardigd kan worden met een beroep op de gemeenschappelijk gedeelde kennis van de morele opvallendheid van de uitkomst, is de individuele verplichting af te zien van moord ook niet langer gerechtvaardigd. Kortom, rechtvaardiging en verplichting hangen samen – precies zoals de criticus verlangde.

De redenen om deze rechtvaardiging van de norm die moord verbiedt af te wijzen, zijn dezelfde als de redenen van Den Hartogh om de opvallendheid van rechts houden als rechtvaardiging voor de hogere-ordeverwachtingen in figuur 1 af te wijzen. Immers, deze rechtvaardiging stelt het conventionalisme voor een dilemma. Ofwel de opvallendheid van het wederzijds afzien van moord is een reden om af te zien van moord, maar dan is het hele bouwwerk van wederkerige verwachtingen overbodig. Er is dan geen interdependente reden om van moord af te zien – alleen een onafhankelijke reden. Dat past niet in de conventionalistische theorie van moraal. Ofwel de opvallendheid van het wederzijds afzien van moord is geen reden om af te zien van moord – waarmee de rechtvaardiging van de norm tegen moord opnieuw op losse schroeven komt te staan.

Misschien hebben we de geciteerde suggestie van Den Hartogh in het laatste hoofdstuk van *Wederkerige verwachtingen* nog steeds niet goed uitgewerkt. Moord en doodslag zijn slecht, maar pas als we van elkaar verwachten, op de juiste wederkerige manier, dat we ons van dat soort handelingen onthouden, ontstaat er een plicht om dit te doen. De (morele) waarde van de afwezigheid van dergelijk geweld geeft op zich nog geen plicht om daarvan af te zien. We moeten dus een manier vinden waarop we zowel de morele waarde van het wederzijds afzien van geweld als de interdependentie van redenen in stand houden. We kunnen dat bereiken door de matrices van figuur 5 en 6 te combineren:

		B	
		Onthoudt zich van moord	Moord & doodslag
A	Onthoudt zich van moord	(2, 2)	(0, 0)
	Moord & doodslag	(0, 0)	(1, 1)

Figuur 7

Nu is de situatie waarin beiden van geweld afzien beter dan wederzijdse moord en doodslag vanuit het oogpunt van de goederenleer. Verder is het zo dat op het eerste gezicht de reden voor A om van moord af te zien afhangt van de vraag of B reden heeft om af te zien van dergelijk gedrag en vice versa. Wellicht is dit het soort rechtvaardiging dat past binnen een conventionalistische theorie. We hebben én de morele veroordeling van het kwaad van moord en doodslag én het interdependente karakter van de verplichting van dergelijk gedrag af te zien bewaard, zonder afbreuk te doen aan de individuele rationaliteit en intentionaliteit van de



betrokken actoren. We hebben, kortom, aan alle randvoorwaarden van een rechtvaardiging binnen een conventionalistische theorie van verplichting voldaan.

Zou dit antwoord de criticus tevreden stellen? Ik denk het niet. Het lijkt alsof er nog steeds een plicht tot moord en doodslag gerechtvaardigd zou kunnen worden. A en B hebben interdependente redenen om af te zien van geweld. Maar zoals we reeds eerder hebben gezien, kunnen interdependente redenen alleen bestaan als er ten minste een alternatieve situatie is die ook door dergelijke redenen in stand gehouden kan worden. Stel nu dat het een feit is dat B een moordeenaar en geweldenaar is. Moet A nu afzien van moord? Nee. Sterker, het lijkt erop alsof A in deze situatie de plicht heeft om ook te moorden! Net zoals de plicht om af te zien van moord en doodslag berust op het bestaan van wederkerige verwachtingen hierover, zo berust de plicht om te moorden op het bestaan van wederkerige verwachtingen hierover.

Den Hartogh zou dit moeten beamen. Bij de bespreking van een identieke matrix als figuur 7 zegt hij:

‘Voor het eerst met deze keuze gekonfronteerd zouden beide spelers zonder aarzeling alternatief a [in figuur 7 “onthoudt zich van moord”; BV] kiezen. Toch kan er een “konventie” bestaan om [moord & doodslag; BV] te kiezen. Er wordt voorrang gegeven aan het verkeer dat de rotonde oprijdt, hoewel iedereen kan inzien dat het verstandiger zou zijn het verkeer dat zich op de rotonde bevindt voor te laten. Hoe kan zo’n “konventie” in stand blijven? (...) Niets anders dan de wederkerige verwachtingen houdt de spelers bij de b-keuze [“moord & doodslag” in figuur 7; BV]. De enige legitimatie voor mijn verwachtingen is dat jouw verwachting de reden is voor jouw keuze: ik kan die verwachting alleen maar vaststellen, niet op grond van verdere redenen reconstrueren.’ (Den Hartogh, 1985, p. 92-92)

Dat gaat te ver. Men zou kunnen betogen dat het niet van een individu verlangd kan worden af te zien van moord en doodslag in Hobbes’ natuurtoestand waarin alles geoorloofd is en men aan mag nemen dat iedereen een moordzuchtige vijand is.<sup>14</sup> Echter, het enige dat daarmee gerechtvaardigd kan worden, is dat in zo’n geval niet verlangd kan worden dat een individu zich als enige onthoudt van moord en doodslag – niet dat het *verplicht* is te moorden.

Den Hartogh meent echter van wel. Dat dit volgt uit Den Hartoghs opvattingen over de aard van sociale verplichting, blijkt ook als we naar zijn definitie van verplichting kijken:

14. Ik denk hierbij aan de discussies over de vraag of moraal meer mag eisen van mensen dan redelijkerwijs verwacht mag worden. Centrale auteurs zijn Williams (1971, 1985) en Kagan (1989).

A is verplicht tegenover B een handeling a te doen dan en slechts dan als

- (1) A verwacht dat B b zal doen.
- (2) Als B b doet geeft B er de voorkeur aan dat A a doet.
- (3) Als B b doet heeft B, gegeven A's voorkeuren, zekere gronden om te menen dat A a zal doen, maar deze gronden zijn ontoereikend om B's keuze voor b volledig te rechtvaardigen.
- (4) B verwacht dat A a zal doen.
- (5) A heeft de morele preferentie onder de in (1) t/m (3) aangegeven omstandigheden aan de in (4) genoemde verwachting van B te voldoen.
- (6) Betreffende voorwaarde (1) t/m (5) bestaat bij A en B [gemeenschappelijk] gedeelde kennis. (Den Hartogh, 1985, p. 131)

Stel dat A en B nu in de situatie zitten die door figuur 7 wordt beschreven. Beiden prefereren een situatie waarin beiden zich onthouden van dergelijk destructief gedrag. Echter, die optimale situatie is alleen sociaal, gezamenlijk toegankelijk. Stel nu dat A verwacht dat B geweld zal gebruiken (de ‘moord & doodslag’-keuze). Aan voorwaarde (1) is dan voldaan. Gegeven de waarderingen van A en B voor de mogelijke situaties, moeten we dus aannemen dat B er de voorkeur aan geeft dat A dat ook doet (voorwaarde (2)). Omdat de redenen voor de ‘moord & doodslag’-keuze interdependent zijn, is aan voorwaarde (3) voldaan. Als B nu verwacht dat A de ‘moord & doodslag’-keuze zal maken – en dat zou hij zomaar kunnen doen, daar is geen verdere grond voor nodig in Den Hartoghs theorie! –, is ook aan voorwaarde (4) voldaan. Als we aannemen dat dit alles gemeenschappelijk gedeelde kennis is, dan lijkt het erop dat A inderdaad verplicht is de ‘moord & doodslag’-keuze te maken, tenzij we aannemelijk kunnen maken dat A niet de ‘morele preferentie’ zal hebben dat te doen.

Echter, uit lezing van *Wederkerige verwachtingen* blijkt dat ‘morele preferenties’ hier disposities zijn te handelen zoals verwacht wordt in een context van de voorkeuren en verwachtingen van anderen. Dit soort ‘preferenties’ kenmerken faire, betrouwbare personen. Het zijn die eigenschappen van een persoon die maken onder sommige omstandigheden (bijvoorbeeld wanneer men vertrouwd wordt) de ander ‘niet te laten zitten’.<sup>15</sup> Het zijn dus geen voorkeuren vóór een bepaald soort uitkomsten, maar juist voorkeuren voor een bepaald soort gedrag onder omstandigheden. Met andere woorden, A en B zouden in deze situatie heel goed het slachtoffer kunnen zijn van hun coöperatieve deugden en pessimistische attitudes ten aanzien van de voorkeuren en verwachtingen van de ander: ‘(...) het resultaat van voldoende plichtbetrachting [kan] zijn dat beiden er per saldo op achteruit gaan’ (Den Hartogh, 1985, p. 132). De criticus kraait victorie! Het interdependente karakter van de verplichting die een sociale norm oplegt, maakt het onvermijdelijk dat de vigerende sociale norm contingent is: er kan geen rechtvaardiging voor worden gegeven, zonder afbreuk te doen aan het uitgangspunt van het conventi-

15. Over deze eigenschappen heb ik elders uitgebreid geschreven (zie Verbeek, 2002, 2007).

onalisme. Het lijkt erop dat we hier de kern van waarheid van de kritiek hebben gevonden.

### De scheiding van verplichting en rechtvaardiging

Ik denk dat een andere reactie op de kritiek nodig is. De kritiek kwam uiteindelijk voort uit de klacht dat het conventionalisme van Den Hartogh niet uit kan leggen wanneer een sociale of morele norm gerechtvaardigd is. Pogingen dat wel te doen op een manier die verenigbaar is met het conventionalisme van Den Hartogh, falen of lopen stuk op het probleem dat de theorie niet uit kan leggen dat als een norm niet gerechtvaardigd is, dit het individu ontslaat van de verplichting. Het was juist dit laatste probleem dat leidde tot de weliswaar theoretische, maar daarom niet minder storende, mogelijkheid van het bestaan van een verplichting tot moord en doodslag.

Dit probleem kan worden ontlopen (maar niet opgelost) door rechtvaardiging en verplichting volledig uit elkaar te trekken, verder dan Den Hartogh wellicht geneigd is te doen. We moeten erkennen dat het conventionalisme geen rechtvaardiging *kan* geven voor een norm. We moeten de criticus gelijk geven in zijn klacht dat de theorie openlaat dat onrechtvaardige regels toch verplichten. Het conventionalisme is echter helemaal geen theorie waarmee normen gerechtvaardigd kunnen worden. Rechtvaardiging valt buiten het bereik van de theorie. Al wat van de theorie verwacht kan worden, is een specificatie van de voorwaarden voor het bestaan van een verplichting.<sup>16</sup>

De conventionalist die Den Hartogh volgt, zal dus tegen de criticus moeten volhouden dat een handeling verplicht kan zijn zonder dat dit impliceert dat de handeling goed is of dat de norm die de verplichting oplegt, een goede norm is. De conventionalist zal daarnaast moeten volhouden tegen de criticus dat de rechtvaardiging van een norm niet impliceert dat die norm verplicht. Evenmin impliceert het ontbreken van zo'n rechtvaardiging dat de norm niet verplicht. Rechtvaardiging van de norm en de verplichting deze te gehoorzamen zijn dan volledig onafhankelijk.

Het conventionalisme laat daarmee de mogelijkheid open dat voor veel geldende normen en regels geen (volledige) rechtvaardiging kan worden gegeven. Op zijn hoogst kan er een verklaring worden gevonden die berust op feiten over onze biologische, sociale en psychologische geschiedenis waarom *deze* norm het is geworden.<sup>17</sup> In dergelijke verklaringen zal toeval een grote rol spelen. Op de vraag waarom we de regels en normen hebben die we hebben, waarom onze sen-

16. Deze voorwaarden zijn samen voldoende indien men van mening is dat ook onrechtvaardige normen verplichtingen scheppen. De voorwaarden zijn noodzakelijk als men van mening is dat alleen rechtvaardige normen verplichtingen kunnen scheppen.

17. Dergelijke geschiedenis kan (maar hoeft niet) een evolutionaire geschiedenis zijn. Een goed voorbeeld van een dergelijke verklaring is Sugden (1986).

timenten zijn zoals ze zijn, zal een conventionalist de schouders moeten ophalen en het antwoord overlaten aan de vakwetenschappen. We kunnen Den Hartoghs favoriete citaat van David Lewis gebruiken in dit verband:

'Conventions are like fires: under favorable conditions a sufficient concentration of heat spreads and perpetuates itself. The nature of the heat does not depend on the original source of heat. Matches may be our best fire starters, but that is no reason to think of fires started otherwise as any the less fires.'  
(Lewis, 1969, p. 92)

Soms wordt een brand aangestoken, soms is een willekeurige blikseminslag voldoende. De vraag of het terecht was dat het vuur zo woedt, en waarom de bliksem nu juist hier in moest slaan, is irrelevant voor de brand en de noodzaak hem te blussen. Hetzelfde geldt voor de verplichting die sociale regels ons opleggen: de vraag of de norm gerechtvaardigd is of verklaard kan worden, doet niets af aan de verplichtingen die zij oplegt.

Dit is geen grote afwijking van Den Hartoghs theorie in *Wederkerige verwachtingen*. Het betekent echter wel dat we de aspiraties en de focus van het conventionalisme moeten beperken. Het conventionalisme is geen morele theorie die ons de aard van alle morele rechten en plichten kan verklaren. De theorie stelt ons in staat te begrijpen hoe het kan dat algemeen aanvaarde regels 'normativiteit' hebben – om een modern anglicisme te gebruiken. Het conventionalisme kan uitleggen hoe sociale regels, morele en andersoortige, redenen geven aan individuele actoren. Het doet dit op een manier die verenigbaar is met een naturalistisch wereldbeeld enerzijds en de plaats van een rationeel delibererend individu in die wereld anderzijds. Het is beslist geen theorie die ons voorschrijft welke normen we zouden moeten hebben. Evenmin is het een theorie die alle morele vereisten kan omvatten. Moraal is zo divers dat het onwaarschijnlijk is dat er een correcte allesomvattende theorie van bestaat, of dit nu een conventionalistische is of niet. De vergissing van de criticus is te menen dat het conventionalisme wel een dergelijke theorie is.

### Literatuur

- Cohen, G.A. (1982) 'Functional explanation: reply to Elster', *Political Studies*, 28(1), p. 129-135.
- Elster, J. (1989) *Nuts and bolts for the social sciences*, Cambridge: Cambridge University Press.
- Hartogh, Govert den (1985) *Wederkerige verwachtingen* (diss. UvA).
- Hartogh, Govert den (2002) *Mutual expectations: a conventionalist theory of law*, Den Haag: Kluwer Law International.
- Hempel, C.G. (1959) 'The logic of functional analysis', in: L. Gross (red.), *Symposium on Sociological Theory*, Evanston, IL: Row, Peterson, p. 271-307.

- Kagan, S. (1989) *The limits of morality*, Oxford: Clarendon.
- Lewis, D. (1969) *Convention: a philosophical study*, Cambridge, MA: Harvard University Press.
- Mackie, J.L. (1977) *Ethics: inventing right and wrong*, London: Penguin Books.
- Schelling, T. (1960) *The strategy of conflict*, Cambridge, MA: Harvard University Press.
- Schotter, A. (1981) *The economic theory of social institutions*, Cambridge: Cambridge University Press.
- Skyrms, B. (1996) *Evolution of the social contract*, Cambridge: Cambridge University Press.
- Skyrms, B. (2004) *The stag hunt and the evolution of social structure*, Cambridge: Cambridge University Press.
- Sugden, R. (1986) *The economics of rights, co-operation and welfare*, Oxford: Blackwell.
- Vanderschraaf, P. (1995) 'Convention as correlated equilibrium', *Erkenntnis*, 42(1), p. 65-87.
- Verbeek, B. (2002) *Instrumental rationality and moral philosophy: an essay on the virtues of cooperation*, Dordrecht: Kluwer.
- Verbeek, B. (2007) 'The authority of norms', *American Philosophical Quarterly*, 44(3), p. 245-258.
- Verbeek, Bruno & Christopher Morris (2008), 'Game theory and ethics', *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition; Edward N. Zalta, red.) (<[plato.stanford.edu/archives/fall2008/entries/game-ethics/](http://plato.stanford.edu/archives/fall2008/entries/game-ethics/)>).
- Williams, B. (1971) 'A critique of utilitarianism', in: J.J.C. Smart & B. Williams (red.), *Utilitarianism for and against*, Cambridge: Cambridge University Press, p. 77-150.
- Williams, B. (1985) *Ethics and the limits of philosophy*, Cambridge, MA: Harvard University Press.