



Universiteit
Leiden
The Netherlands

Boekbespreking van: Anarchie, staat en het Prisoner's Dilemma

Hees, M.V.B.P.M. van; Lieshout, R.H.

Citation

Hees, M. V. B. P. M. van, & Lieshout, R. H. (1995). Boekbespreking van: Anarchie, staat en het Prisoner's Dilemma. *Acta Politica*, 30: 1995(2), 244-248. Retrieved from <https://hdl.handle.net/1887/3450343>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3450343>

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

ging himself observed, the aristocracy argued about the legitimacy of younger titles (p. 14) and took pride in its lineage – its noble *blood*.

Secondly, political theory before the Revolution was more than a discourse on aristocracy, even in aristocratic circles. In his zest to reconstruct the classicist ethos, Kinneging totally blacks out more familiar subjects like religion, religious war, church and state, tolerance and every possible influence the Enlightenment may have had on aristocratic political thought. To a large degree, Kinneging's central thesis is based on a gross reduction of *wie es eigentlich gewesen*: what really happened happened among the ruling class, among the mainstream of the ruling class (defined as the unmodern part of the aristocracy), on the greater part of their bookshelves particularly, in those parts of texts that do not smell of modernity, religion or deviation from three centuries of 'normal science'. The odd result is that one begins to wonder whether legendary figures like d'Alembert, Voltaire, de Lamettrie, de Sade, Diderot and their aristocratic friends existed at all.

Marcel Wissenburg

H. Pellikaan, **Anarchie, staat en het Prisoner's Dilemma**, Eburon, Delft 1994.

Het zal beoefenaren van de politicologie wel niet zijn ontgaan dat binnen de politieke economie en aanverwante formele benaderingen van politieke verschijnselen de oude volkswijsheid 'het leven is een pijpkanneel, elk zuigt eraan en krijgt zijn deel' het heeft moeten afleggen tegen de wijsheid dat het leven een Prisoner's Dilemma is. De sociaal-politieke werkelijkheid is één groot probleem van collectieve actie, aangezien de preferenties van individuen ten aanzien van de verwezenlijking van een gemeenschappelijk belang zodanig zijn dat ieder van hen besluit niet bij te dragen aan de verwezenlijking van dit gemeenschappelijke belang, ongeacht wat de anderen doen, ook al verkiest ieder van hen de uitkomst dat iedereen bijdraagt aan het gemeenschappelijk belang, boven de uitkomst dat niemand een bijdrage levert. Met de menselijke soort is het helaas zo gesteld dat het slechts door middel van dwang mogelijk zal zijn economische crises, ecologische rampen en totale sociale ontwijking te voorkomen. Het is deze modieuze wijsheid die Huib Pellikaan in zijn proefschrift *Anarchie, staat en het Prisoner's Dilemma* wenst te ontcrachten. De wapens die hij daartoe in de strijd werpt zijn scherpzinnige analyses, een ontzagwekkende belezenheid, ontvouwingsschalen, tientallen spelmatrices én het actorperspectief.

Pellikaan koestert dezelfde ambitie als Michael Taylor en Robert Axelrod. Hij wil laten zien dat onder bepaalde omstandigheden individuen wel degelijk spontaan met elkaar zullen samenwerken teneinde een gemeenschappelijk doel te verwezenlijken. Aan de hand van deze analyse wil hij de grenzen verkennen van de individuele vrijheid en van het overheidsingrijpen. Pellikaan verwerpt echter de claim van Taylor en Axelrod dat het (herhaalde) Prisoner's Dilemma een adequaat kader biedt om die omstandigheden op het spoor te komen. Die claim is volgens Pellikaan tweeledig

van aard. Enerzijds is er de politiek-filosofische claim, die inhoudt dat actoren in de natuurlijke toestand niet tot de oprichting van een Leviathan zullen overgaan. Anderzijds is er de politiek-economische claim, die luidt dat in een geordende samenleving overheidsingrijpen ten behoeve van de verwezenlijking van gemeenschappelijke belangen achterwege kan blijven omdat actoren zullen besluiten op vrijwillige basis met elkaar samen te werken.

De dissertatie bestaat uit drie delen. Het eerste deel is gewijd aan de politiek-filosofische claim. Pellikaan onderzoekt of de situatie van de individuen in de natuurlijke toestand zoals beschreven door Hobbes in zijn *Leviathan*, op vruchtbare wijze kan worden geanalyseerd met behulp van de speltheorie, en zo ja, welk speltheoretisch model het keuzeprobleem van de betreffende individuen het best beschrijft. Verschillende modellen passeren daarbij de revue. Pellikaan argumenteert dat de natuurlijke toestand, in tegenstelling tot wat velen menen, niet als een Prisoner's Dilemma kan worden opgevat, en ook niet, zoals onder meer door Hampton wordt gesteld, als een coördinatiespel, maar moet worden geïnterpreteerd als een onderhandelingspel. Volgens Pellikaan betekent dit dat individuen in de natuurlijke toestand wel degelijk zullen kiezen voor het oprichten van een gemenebest en dat Hobbes' rechtvaardiging van de staat derhalve volledig overleefd blijft: de politiek-filosofische claim moet worden afgewezen. In het tweede deel staat de politiek-economische claim centraal. Deze wordt onderzocht aan de hand van een analyse van het Prisoner's Dilemma-superspel. Volgens Pellikaan vertonen zowel het PD-superspel van Michael Taylor als dat van Robert Axelrod onoverkomelijke tekortkomingen. Op grond hiervan verwerpt Pellikaan de stelling van Taylor en Axelrod dat iteratie van het PD het mogelijk maakt dat individuen zich voorwaardelijk coöperatief tegenover elkaar zullen opstellen en op basis van vrijwilligheid zullen samenwerken teneinde een collectief goed te realiseren. Het derde deel van het proefschrift draait om het onderscheid tussen het waarnemersperspectief en actorperspectief. Het gaat hierbij om de veronderstelde preferenties van individuen. Binnen het waarnemersperspectief wordt volgens Pellikaan extern bepaald – dat wil zeggen, door de waarnemer – wat die preferenties zijn, terwijl binnen het door Pellikaan geprefereerde actorperspectief wordt uitgegaan van de motieven zoals de individuen zelf zeggen die te hebben. Deze motieven worden vervolgens met behulp van een specifiek rationaliteitsbegrip – namelijk dat rationaliteit neerkomt op consistentie tussen motieven en preferenties – gekoppeld aan de individuele preferenties. In dit verband formuleert Pellikaan ook zijn, wat wij zouden willen noemen, 'politiek-empirische claim': de vraag of overheidsingrijpen noodzakelijk is om een zeker collectief goed te produceren, kan slechts worden beantwoord na empirisch onderzoek naar de motieven van degenen die aan de verwezenlijking van dit collectieve goed zouden moeten bijdragen. Doordat actoren zich niet alleen laten leiden door zelfzuchtige motieven, maar ook heel andere motieven kunnen hebben, is het goed mogelijk dat zij vrijwillig een bijdrage zullen leveren indien dat van hen wordt gevraagd. Dat individuen inderdaad motieven kunnen hebben die leiden tot samenwerking op vrijwillige basis, toont Pellikaan vervolgens aan met behulp van een empirisch onderzoek naar twee hypothetische besluitvormingssituaties, het 'Student's Dilemma', als voorbeeld van een tweepersons Prisoner's Di-

lemma, en het 'Milieuprobleem', dat model staat voor een N-persoons probleem van collectieve actie.

Het boek van Pellikaan is intellectueel zeer stimulerend. Alles wordt nog eens op een rijtje gezet en Pellikaans aanpak zorgt ervoor dat de lezer vanzelfsprekendheden en automatismen die mogelijk in zijn gebruik van de speltheorie zijn ingeslopen op het spoor komt en vervolgens nog eens kritisch beziet. Dit geldt heel sterk voor Pellikaans verhandeling over de correcte speltheoretische interpretatie van de Hobbesiaanse natuurlijke staat. Pellikaans bespreking van de mogelijke speltheoretische modelleringen leidt ontegenzeggelijk tot een beter begrip van de theorie van Hobbes. Het gehanteerde onderscheid tussen 'covenants to keep the peace' en 'covenants to keep his promise' is hier een fraai voorbeeld van. Hetzelfde geldt voor zijn behandeling van het argument tegen de dwaas, één van de cruciale passages uit de *Leviathan*. Al moet gezegd dat Pellikaan zich lijkt te vergalopperen wanneer hij stelt dat het argument van Hobbes erop neerkomt dat deze: 'in feite alle andersdenkenden als asociale elementen de toegang tot zijn gemenebest ontzegt' (p. 86). Ons inziens valt de stelling heel goed te verdedigen dat Hobbes de dwaas slechts hetzelfde probeert duidelijk te maken als dat wat de reus Fasolt de god Wotan voorhoudt wanneer de laatste, net als de dwaas, probeert onder een 'half-spot, half-forward contract' uit te komen: 'Was du bist, bist du nur durch Verträge' (Das Rheingold, tweede scene).

Teleurstellend is dat Pellikaan zeer veel aandacht besteedt aan de vraag welke speltheoretische modellen de natuurlijke staat *niet* adequaat beschrijven, maar dat hij slechts zeer summier de contouren aangeeft van het model dat naar zijn mening de natuurlijke toestand wel goed beschrijft: het onderhandelingsmodel. Omdat het onderhandelingsmodel slechts in algemene termen wordt beschreven, is niet duidelijk waarom dit model de natuurlijke staat beter zou beschrijven dan de in de literatuur voorgestelde spelen. Het is daardoor bovendien onduidelijk waarop Pellikaans afwijzing van de politiek-filosofische claim is gebaseerd. Gesteld wordt dat rationale actoren zullen samenwerken om het gemenebest op te richten, omdat het: 'plausibel (is) dat de individuen in de natuurstaat het gemenebest zullen prefereren boven de natuurstaat' (p. 147). Een van de kerninzichten uit de speltheorie is nu echter juist dat de relatie tussen preferenties en gedrag niet altijd zo ondubbelzinnig te formuleren valt: de verdachten uit het Prisoner's Dilemma worden immers veroordeeld tot een langere gevangenisstraf, omdat ze zo graag een kortere gevangenisstraf willen.

Ook de twee hoofdstukken waarin de politiek-economische claim van Taylor en Axelrod wordt geanalyseerd, blinken uit door de wijze waarop de lezer wordt gedwongen de fundamentele van de speltheorie en van collectieve-actiemodellen nog eens te doordenken. Fraai is de wijze waarop Pellikaan laat zien dat in het licht van David Humes sombere visie op de menselijke natuur (we zijn onverbeterlijke opportunisten) Taylor ten onrechte meent dat Hume in zijn *Treatise of human nature* het menselijk gedrag plaatst in de context van een herhaald Prisoner's Dilemma (pp. 153-154). Zoals bekend proberen Taylor en Axelrod aan te tonen dat een herhaling van het Prisoner's Dilemma kan leiden tot het ontstaan van samenwerking. Taylor veronderstelt daarbij dat het aantal rondes van het PD-superspel oneindig is. Het Prisoner's Dilemma wordt oneindig vaak gespeeld. Pellikaans centrale argument tegen

Taylor richt zich op deze oneindigheidsassumptie en op het beperkte aantal superstrategieën waarover de individuen beschikken in Taylors model. Pellikaan betoogt dat Taylor in plaats van de oneindigheidsassumptie had moeten veronderstellen dat de spelers over onvolledige informatie beschikken. Zij weten weliswaar dat het spel een keer ophoudt, maar weten niet precies wanneer dat het geval zal zijn: het aantal rondes is onbepaald. Vervolgens introduceert Pellikaan de superstrategie $E\Phi$, die zegt dat een speler alleen in de laatste ronde voor noncoöperatie kiest. Aangevoerd wordt dat de introductie van deze superstrategie leidt tot de verdwijning van het door Taylor gevonden coöperatieve Nash-evenwicht. Pellikaan ziet daarbij echter over het hoofd dat de strategie $E\Phi$ juist datgene veronderstelt wat in een superspel met een onbepaald aantal rondes is uitgesloten: kennis van het aantal rondes. Immers, hoe zou een individu kunnen weten wanneer de laatste ronde is als het aantal rondes onbepaald is? Hoewel de kritiek van Pellikaan op de oneindigheidsveronderstelling overtuigend is, realiseert hij zich dus niet dat de introductie van onvolledige informatie het spel drastisch verandert. Omdat Pellikaan geen PD-superspel met onvolledige informatie analyseert, is zijn conclusie dat samenwerking moeilijk of zelfs helemaal niet tot stand zal komen indien het aantal rondes eindig maar onbepaald is, op zijn minst voorbarig. De politiek-economische claim met betrekking tot spontane samenwerking, voor zover die gebaseerd is op het oneindige karakter van een herhaald Prisoner's Dilemma, blijft recht overeind staan.

Twijfels bestaan er ook met betrekking tot de waarde van Pellikaans empirisch onderzoek. Pellikaan onderzoekt namelijk geen gedrag (wat hij ook ruiterlijk toegeeft), maar vraagt individuen naar hun motieven en bekijkt vervolgens of motieven en preferenties consistent zijn. De reden waarom hij dit doet, zo verzekert hij verschillende keren, is omdat: 'de assumptie van rationaliteit een empirische basis dient te hebben' (p. 242), om de speltheorie 'een empirische fundering te geven' (p. 328). Wanneer we verder voorbijgaan aan het punt dat het zoeken naar de fundering van empirische theorieën slechts leidt tot een oneindige regressie, dan blijft de vraag waarom de speltheorie een dergelijke fundering nodig zou hebben. De speltheorie doet geen uitspraken over de motieven die aan preferenties ten grondslag liggen, maar onderzoekt de strategische dimensie van besluitvormingssituaties bij *gegeven* preferenties. Als onderzoekers in hun pogingen uitspraken over de werkelijkheid te doen, speltheoretische modellen hanteren die onjuiste veronderstellingen bevatten over de preferenties van individuen, dan is dat geen manco van de speltheorie maar van de gedragsassumpties die door de onderzoekers (bewust of onbewust) zijn ingevoerd. Als de speltheorie in dezen al iets kan worden verweten, dan is het niet dat zij de verkeerde assumpties over de aard van de individuele preferenties hanteert, maar dat zij geen assumpties hieromtrent hanteert. De speltheorie is geen theorie die uitspraken doet over de werkelijkheid, maar een instrument met behulp waarvan uit een empirische theorie uitspraken over de werkelijkheid kunnen worden afgeleid. Wat bij het gebruik van de speltheorie in de sociale wetenschappen ter discussie staat is niet haar mogelijke empirische fundering, maar of de met behulp van de speltheorie afgeleide evenwichtsuitkomsten zich in de empirie ook daadwerkelijk voordoen. Een nauw aan het voorgaande gerelateerd probleem is dat Pellikaans betoog wordt

ontsied door het inmiddels welhaast klassiek te noemen misverstand over de betekenis en de relaties tussen de begrippen preferenties, nut en eigenbelang – nog afgezien van het feit dat waar volgens Hobbes de individuen in de natuurlijke staat naar zelfbehoud streven, Pellikaan hen gedreven ziet door eigenbelang (p. 74). De opmerking dat de speltheorie uitgaat van 'de veronderstelling dat het individu streeft naar maximalisatie van zijn eigenbelang. Het eigenbelang wordt vorm gegeven in zijn preferentieordening' (p. 10), kan exemplarisch worden genoemd. De formulering is verwarrend en feitelijk onjuist. De speltheorie spreekt niet over belangen, maar over preferenties. Zoals gezegd, zijn de preferenties primitieven van het model: zij zijn niet verder gespecificeerd en de speltheorie doet dan ook geen uitspraken over wat er verder in wordt 'vorm gegeven'.

Anarchie, staat en het Prisoner's Dilemma is een veelzijdig boek. Vele onderwerpen passeren de revue: van vangstbeperking bij vissers, de onverenigbaarheid van Gods geboden met de rationele-keuzetheorie (p. 58), tot en met het strategisch belang van een goede paraplu. Niettemin had het vertoon van belezenheid wel wat mogen worden ingeperkt. Ontegenzeggelijk weet Pellikaan de rode draad in zijn betoog goed vast te houden. Tegelijkertijd haalt hij zoveel overhoop wat hij na bespreking als onjuist of niet relevant terzijde schuift, dat het voor de lezer toch moeilijk wordt die rode draad vast te houden. Voor de leesbaarheid zou het een goede zaak zijn geweest indien Pellikaan zich had weten te beheersen en zijn regelmatige uitstapjes naar het werk van auteurs die ook wel eens iets over een bepaald onderwerp hebben gezegd, maar dat verder geen enkele functie heeft voor zijn betoog, had geschrapt.

Bovenstaande kritiek kan echter niet verhelen dat *Anarchie, staat en het Prisoner's Dilemma* voor alles een indrukwekkende prestatie is. In de slotalinea van zijn voorwoord spreekt Pellikaan 'de (irrationele?) hoop uit dat in de toekomst de politieke economie in Nederland dezelfde aandacht zal krijgen als in de Angelsaksische landen' (geen paginanummer). Een proefschrift als het zijne vormt een goede reden aan te nemen dat de toekomstige bloei van de politieke economie in Nederland spoedig een rationele verwachting zal zijn.

M.V.B.P.M. van Hees en R.H. Lieshout

J.C. Dagevos, **Naar een sociaal interactiemodel: over de dynamiek van wetenschap, macht en afhankelijkheid**, Bilthoven, Prime Press 1994.

Dit proefschrift heeft reeds de nodige publieke aandacht op zich gevestigd en is niet oncontroversieel bevonden (zie NRC, 29-9-1994). Een nadere analyse van de wetenschappelijke inhoud van deze studie is dan ook op zijn plaats. De studie presenteert een nadere aanzet tot theorievorming op het gebied van sociale interactie. Evenals het onlangs gepubliceerde boek *Anarchie en hiërarchie* van R.H. Lieshout (Coutinho, 1993) leunt deze studie sterk op het werk van promotor De Vree (Prime Press, 1982/1990). Het uiteindelijke doel van alle inspanningen is het formuleren van

een theorie van sociaal gedrag tussen (al dan niet formeel georganiseerde groepen van) mensen teneinde dit gedrag in de empirie te kunnen verklaren en voorspellen.

Daartoe wordt eerst in hoofdstuk 1 uiteengezet welke wetenschappelijke houding tegenover theorievorming aangenomen dient te worden (§1) en hoe deze houding aansluit bij de opvattingen van met name Kuhn, Lakatos en Popper betreffende de totstandkoming van wetenschappelijke kennis over observeerbare verschijnselen (§2-5). In §1 wordt een gezonde wetenschappelijke twijfel en volharding aanbevolen tijdens het proces van theorievorming om uiteindelijk succes te bereiken. Daarna wordt in hoofdstuk 1 betoogd dat theorie-ontwikkeling veel tijd vergt. De hypothesen gevat in een theorie moeten dan ook niet te snel aan falsificatieprocedures worden onderworpen teneinde de verdere uitwerking en nuancering van de theorie mogelijk te maken. En ingeval van falsificatie van hypothesen dient een kritische houding te worden aangenomen betreffende de waarde van dergelijke falsificatie voor (delen van) de theorie.

Merkwaardig genoeg wordt echter nergens uiteengezet aan welke wetenschapstheoretische en methodologische criteria het produkt van het proces van theorievorming, namelijk de theorie zelf, moet voldoen. In het licht van deze studie is dat vreemd. Het aangehaalde werk van Popper biedt voldoende aanknopingspunten hiervoor. Popper (Hutchinson, 1975³, p.111-123) stelt, dat een theorie beter is naarmate de empirische inhoud ervan groter is. Dit betekent dat een theoretisch model beter is naarmate de empirische inhoud van de theoretische begrippen éénduidiger is gedefinieerd. Binnen de door Popper aangehangen deductief-nomologische verklaringswijze wordt dit gerealiseerd door de (afgeleide) theoretische begrippen te operationaliseren via niet tot de theorie behorende 'initial conditions' van indicatoren van observeerbare verschijnselen. De kritiek van Quine op het logisch-empirisme heeft nadien geleid tot de hedendaagse opvatting binnen de wetenschapstheorie dat deductief-nomologisch verklaren een grensgeval is van inductief-statistisch verklaren (zie Stegmüller, Springer Verlag, 1983², Band I, p.1-10/940-978). Quine's kritiek bestaat eruit dat elke verklaring van observeerbare verschijnselen extra-logische elementen bevat (in empirische termen), die uitspraken vanuit de gededuceerde theorie van toepassing maken op de desbetreffende empirische verschijnselen. De extra-logische elementen worden dus toegevoegd aan een gededuceerde theorie en vormen samen een (nieuwe) specifieke theorie betreffende een bepaalde categorie van observeerbare verschijnselen. En de specificatie van deze extra-logische elementen ontbreekt nu juist binnen de deductieve theorievorming betreffende menselijk gedrag en sociale interactie in de hoofdstukken 4 tot en met 9 van de onderhavige studie. Hierdoor blijft het ontwikkelde theoretische raamwerk empirisch leeg, waardoor het zeer lastig wordt om dit raamwerk op zijn validiteit en betrouwbaarheid te toetsen. Een eindeloze discussie over correcte operationalisaties van de theoretische begrippen en correcte specificaties van de theoretische (causale) relaties tussen empirische indicatoren ligt in het verschiet.

Over het proces van theorievorming zelf wordt weinig uiteengezet, behalve op p.17 en 18 waar een aantal uitspraken hierover wordt gedaan met verrekende implicaties. Op p.17 wordt gemeld: 'Om de doorzichtigheid, kracht en de subtiliteit van