

PERIODIEK

— P A R O D O S —

VERENIGING

TOT INSTANDHOUDING

VAN DE

INTERNATIONALE SCHOOL

VOOR WIJSBEGEERTE

IN DIT NUMMER

- Kants categorische imperatief
- Wetenschap tussen prachtig en machtig
- Paradoxaal denken
- Filosofische puzzel
- Nieuwe boeken

HOLLE FRASE, LEGE WOORDEN

Drie onmogelijke versies van de Categorische Imperatief

door M.L.J. Wissenburg

Inleiding

In dit korte artikel¹ wil ik, populair gezegd, een drietal zeepbellen, in de vorm van alternatieve formuleringen van Kants categorische imperatief, blazen en vervolgens doorprikken. Zoals Aristoteles in het begin van boek V van zijn *Ethica Nicomachea* al zei: soms kunnen we iets leren kennen door te kijken naar dat wat het niet is. Ik wil zo toelichten waarom de categorische imperatief, een van de belangrijkste bouwstenen van het liberale politieke en politiek-filosofische denken, niet uit lege woorden bestaat (in tegenstelling tot hetgeen Kant meent), en in welke zin hij geen holle frase is (in tegenstelling tot hetgeen Kants critici beweren). De conclusies die ik trek zijn toegespitst op de categorische imperatief, maar - werk dat ik aan de lezer overlaat - ze kunnen ook met andere politieke en ethische beginselen worden toegepast: grondrechten, mensenrechten, natuurlijke rechten, principes van sociale rechtvaardigheid, etc.

Laat ik niemand vervelen met een lange biografie van Kant. Het volstaat hier te zeggen dat hij, behalve een groot, ook een onleesbaar filosoof was, vrij wereldvreemd en zeer precies. Op enkele korte reizen na heeft hij zijn hele leven doorgebracht in één kleine universiteitsstad: Königsberg in Oost-Pruissen (nu als Kaliningrad in de USSR gelegen), waar hij in 1724 werd geboren en in 1804 overleed. Volgens de anekdotes over zijn dagindeling hanteerde hij blijkbaar een ongeschreven interpretatie van zijn categorische imperatief: "Handel zo, dat jij en anderen de klok op je gelijk kunnen zetten". Kants precisie en wereldvreemdheid komen ook tot uitdrukking in zijn belangrijkste ethische werken, de *Grundlegung zur Metaphysik der Sitten* (1785) en de *Kritik der praktischen Vernunft* (1788). Na de eerste druk van beide werken ging hij het zetduiveltje met eindeloos geduld en tot op de komma's en dubbele punten te lijf, waarna tweede drukken verschenen die bewezen dat ook het zetduiveltje niet voor rede vatbaar was. Een (gelukkig klein) deel van de Kant-interpretatie draait daarom nog steeds om interpretatie en vervoegingen.

Verreweg de bekendste bijdrage van Kant aan de filosofie is de categorische imperatief. Het zetduiveltje en Kant hebben samen gezorgd voor (ten minste) dertien verschillende formuleringen van de categorische imperatief. Er zijn evenwel drie grondvormen aan te wijzen, waar de overige (minimaal) tien formuleringen niet fundamenteel van afwijken.

Vrijheid en regels

De oorsprong van de categorische imperatief ligt in de metafysica. Volgens Kant houdt ethisch handelen het midden tussen determinisme en willekeur. Ethisch handelen is bewust en gewild regelmatig handelen. Een van rede verstoken wezen (bij voorbeeld een plant of dier) handelt wel regelmatig, maar is zich noch bewust van de regels die het volgt, noch van de onmogelijkheid die regels te breken: het valt onder het rijk der natuurwetten. Een redelijk wezen daarentegen kan zich, dankzij de rede, bewust zijn van een zekere wilsvrijheid: hij moet eerst zijn wil bepalen voor hij kan handelen. Dat die vrijheid bestaat valt niet te bewijzen; volgens Kant is het een intuïtief gegeven ervaring.

Redelijke wezens (God, de engelen en de mens - Kant hanteerde een traditioneel-christelijk wereldbeeld) zijn echter ook beperkt in hun vrijheid, juist door (het bewustzijn van) die vrije wil. Wie iets wil omwille van iets anders, door een drijfveer die buiten zijn redelijke aard bestaat, handelt niet vrij, niet naar een zelfgekozen wet, maar uit willekeur. Kant spreekt in dat verband van de plicht die ethisch handelen kenmerkt: het gaat niet om de gevolgen van het handelen, maar om de intentie, de motieven van het verantwoordelijke wezen. Als die motieven van het toeval, van empirische omstandigheden, afhankelijk zijn, kan er intuïtief geen sprake zijn van vrij bepaald handelen. Wat blijft er over van de idee van een goede, deugdzame menselijke wil, vraagt Kant, wanneer we daaruit alle inhoud en objecten wegdenken? De plicht "an sich", de vorm van de wil. Hij onderscheidde daarmee de autonomie (letterlijk: eigen wet) van de heteronomie (wet van buiten) van de wil. Samengevat in een jargon waar Kant zich niet voor zou schamen: de wil moet zichzelf tot principe verkiezen; alleen dan is hij redelijke en autonoom.

De mens, een regels ontwerpend en volgend wezen, kent meerdere soorten regels. In de eerste plaats moeten maximes van imperatieven onderscheiden worden. Maximes zijn subjectief, wat wil zeggen dat ze voor één bepaald persoon gelden, bij voorbeeld: "wanneer ik morgen nog gas, water en licht

wil hebben, moet ik vanavond mijn rijke buurman beroven”. Imperatieve daarentegen zijn algemeen geldig. Binnen de klasse der imperatieve kunnen weer hypothetische van categorische regels onderscheiden worden, naarmate ze niet, respectievelijk wel, onvoorwaardelijk geldig zijn. Elke imperatief van het “als-doel, dan-middel” type is enkel onder voorwaarden geldig. “Wie wil dat zijn tekst begrepen wordt, moet begrijpelijk schrijven” is alleen geldig onder de voorwaarde dat we begrepen willen worden.

De Categorische Imperatief

Kant nu stelt dat de redelijke ethische regel, waaraan alle andere regels afgemeten kunnen worden, algemeen en onvoorwaardelijk geldig moet zijn - een categorische imperatief dus. Elke andere regel veronderstelt heteronomie, doelen buiten de voor alle redelijke wezens gelijke wil.

Zo komt Kant tot de eerste formulering van zijn categorische imperatief: “Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde” (BA 52)². Het is een volstrekt formele imperatief, waarin niets gezegd wordt over wat gedaan moet worden. Kant geeft uitsluitend een randvoorwaarde aan: van de regel waarnaar je handelt moet je kunnen willen dat ze universaliseerbaar (“algemeen”) is. De formulering “kunnen willen” is opvallend. Voor de vrije wil, de mens in het “rijk van de vrijheid”, mag het niet uitmaken of een regel in de fysische wereld, het “rijk van de noodzaak”, ook werkelijk universeel geldig kan zijn. (Kant neemt gewoon aan dat kunnen en willen naadloos op elkaar aansluiten (BA 80)). Als de wil afhankelijk was van wat mogelijk is, dan was hij niet autonoom maar heteronoom.

De tweede formulering van de categorische imperatief volgt bijna rechtstreeks uit de eerste: “Handle so, als ob die Maxime deiner Handlung durch deinen Willen zum allgemeinen Naturgesetze werden sollte” (BA 52; een alternatieve formulering: BA 82). Het belangrijkste verschil met de eerste versie is dat er nu sprake is van een natuurwet in plaats van een rechtsregel. Kant verlangt blijkbaar dat we ons een moment indenken de god te zijn die de natuurwet gaat opstellen - een wet die ook de wetgever zal binden. Voor de eis van universaliseerbaarheid maakt dat geen verschil³, voor de vrijheid van de wil echter wel. Het is niet mogelijk dat de vrije wil aan een natuurwet onderworpen is, immers dan was hij niet vrij en geen wil (daarom ook zegt Kant “als ob”) - maar toch moeten we dat als wetgever kunnen willen. Dat is paradoxaal: de wil moet zichzelf in feite durven ontkennen, hij moet bereid

zijn figuurlijk zelfmoord te plegen, terwijl zulke verlangens wat Kant betreft immoreel zijn. Vanuit Kants theologie is zo'n ontkenning van de wil bovendien ondankbaar ten opzichte van de schenker van vrijheid - en het is helemaal moeilijk voor te stellen dat God zelf dood zou kunnen willen zijn. Kant draaft dus door, en wel omdat hij - vergeef me de beeldspraak - op zijn stokpaardje zit: de plicht. Wat hem betreft is er maar een kleine gedachtesprong nodig van de autonome mens die zijn handelen op de idee van de plicht “an sich” afstemt, naar de mens die plichtmatig handelt, die zo zeer van goede wil is, dat het automatisch zijn intentie is zijn plicht te doen (cf. BA 81).

Een opmerking terzijde: Kants al te gemakkelijke identificatie van plicht en plichtmatigheid heeft de deontologische⁴ ethiek een in dat opzicht onverdiend slechte naam bezorgd. Hij heeft daarmee namelijk aanleiding gegeven tot een weinig subtiele kritiek, die zou illustreren hoe zeer deontologisch handelen contra-intuïtief is. Het argument is gelukkig nooit in de leerboeken terecht gekomen, maar in discussies wordt het nog wel eens als doodoener gehanteerd: “Wat is u liever, dat u goed gedaan wordt omdat iemand u goed wil doen, of dat het u wordt gedaan omdat het nu eenmaal moet?” Er bestaan talloze varianten van dit argument, alle even zeer “wide off the mark”. In de eerste plaats is er geen principiële tegenspraak tussen formeel handelen naar de plicht en inhoudelijk handelen om een goed doel te bereiken - de categorische imperatief beperkt alleen de mogelijke goede doelen door te eisen dat ze universeel geldig zijn. En in de tweede plaats kan er bij iemand die deontologisch handelt geen sprake zijn van tegenzin (“het moet nu eenmaal”): hij of zij handelt per definitie naar “de idee van de wil op zich” en wil zo handelen.

De derde en laatste grondvorm van de categorische imperatief wijkt sterk af van de beide andere: “Handle so, daß du die Menschheit, sowohl in deiner Person, als in der Person eines jeden andern, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchest” (BA 66). In de alternatieve formulering op BA 82 heet het kortweg dat iedereen een redelijk wezen, zowel zichzelf als anderen, “steeds ook als doel” moet beschouwen. (Merk overigens op dat Kant niet zegt dat een mens nooit als middel gebruikt mag worden; weer een wijd verbreid misverstand.) Kant verlangt ook hier universaliseerbaarheid, spreekt echter niet langer over de wil en de formele intenties, maar over de inhoud van de wil. In het algemeen doet het er, ethisch gesproken, niet toe welke middelen je tot welke doelen aanwendt, zolang de categorische imperatief (in de eerste of tweede vorm) maar gerespecteerd wordt.

Wanneer het echter gaat om het gebruik van redelijke wezens wordt het wat anders. Een redelijk wezen, met een zelfbewustzijn en wil (een "Person"), ervaart zich namelijk eerst als doel op zich, een doel ten opzichte waarvan de waarde van alle andere doelen, wensen en middelen relatief is (BA 64-65); een absoluut doel ook dat alle verdere doelen, wensen en middelen bepaalt. Op dat postulaat past Kant vervolgens de eis van universaliseerbaarheid toe: als voor mij geldt dat ik niet enkel middel ben maar vooral doel van mijn wil, zuiver en alleen omdat ik bepaalde universeel-menselijke eigenschappen heb, dan geldt dat ook voor anderen. De maxime die daaruit afgeleid wordt lijkt sterk op de beroemde christelijke en confuciaanse Gulden Regel: "wat gij niet wilt dat u geschiedt..." - met dien verstande dat de derde vorm van de categorische imperatief uiteindelijk op universaliseerbaarheid en niet op reciprociteit berust.

Er lijkt echter achter deze formulering nog een ander postulaat schuil te gaan: de normatieve betekenis van het zelfbewustzijn. Waarom is het feit dat de mens zich als absolute waarde en doel op zich ervaart, ook een norm? Kant probeert die sprong onder andere te rechtvaardigen door te stellen dat mensen noden hebben (met name aan erkenning door anderen) en doelen, waarbij ze vaak de hulp van anderen nodig hebben. Wie het doel wil, moet ook de middelen willen, en zonder de anderen te respecteren en te behandelen als gelijken zal niemand zijn noden bevredigd zien of zijn doelen bereiken. Zo wordt, meent Kant, het feit een norm.

Kants pogingen de derde variant van de categorische imperatief met de andere twee in overeenstemming te brengen hebben aanleiding gegeven tot veel commentaar. De consensus onder zijn critici is dat hij een paar vooronderstellingen binnensmokkelt, misschien wel heteronome principes, namelijk de medemens en zijn geluk, doelen buiten de wil op zich. Welke vooronderstellingen dat precies zijn, en hoe Kant ze binnensmokkelde, is voor dit verhaal niet zo interessant.⁵ Belangrijk is de constatering dat hij het deed, en de vraag waarom. Probeert Kant met deze vorm van de categorische imperatief misschien tegelijk een universeel egoïsme te rechtvaardigen, en het te "beschaven" door er de scherpe kantjes af te vijlen? Ik kom daar in de laatste paragraaf op terug.

Alternatieve categorische imperatieven

De drie grondvormen van de categorische imperatief zijn alle formeel. Ze geven uitsluitend aanwijzingen over de wijze waarop gehandeld moet

worden, en zeggen niets over de waarde van bepaalde doeleinden. Zelfs de derde versie gaat niet verder dan de stelling dat andere personen het evenbeeld van de handelende persoon zijn, in wezen identiek met hem zijn, en als gelijken moeten worden behandeld. Alles lijkt te mogen, zolang er maar een passende formulering van de categorische imperatief op kan worden geplakt: armoede, politieke ongelijkheid, oorlog, paternalisme, leugens om bestwil. De tekst van de categorische imperatief ligt namelijk niet vast. Wat Kant betreft is de categorische imperatief het idee van de onvoorwaardelijk en algemeen geldige regel op zich. Hoe die regel voor ons precies geformuleerd wordt is niet zo belangrijk (getuige ook de vele varianten die Kant zelf heeft uitgedacht). Woorden op zich zijn leeg; het gaat om het idee waarnaar ze verwijzen. Ik noem dit de these "Lege Woorden 1".

De these "Lege Woorden 2" is afkomstig van Kants tegenstanders. Zijn critici hanteren regelmatig twee technieken. Ten eerste pogen ze aan te tonen dat de categorische imperatief niet in staat is onderscheid te maken tussen goed en kwaad: het besluit van de ene mens dief te worden, en dat van de ander arts te worden zouden beiden het produkt zijn van een vrije wilsbeschikking.⁶ Anderen gebruiken, ten tweede, graag dilemma's waarin de categorische imperatief tot contra-intuïtieve oordelen komt over toestanden als armoede, politieke ongelijkheid, en dergelijke. Deze beide lijnen van kritiek, waarlangs gepoogd wordt Kants principe als "lege woorden" te diskwalificeren, als regels die alles toelaten, zal ik verder buiten beschouwing laten. Mij gaat het erom te laten zien dat de categorische imperatief in zekere opzichten geen holle frase is. Dat is een stuk gemakkelijker dan Kant met intuïties weerleggen. Kant zou daarop immers kunnen antwoorden dat dergelijke intuïties grotendeels cultuurprodukten zijn, daarom heteronoom en niet gerechtigd over goed en kwaad te oordelen, laat staan die begrippen te definiëren. De discussie die daaruit volgt is lang, langdradig en ingewikkeld.

Beperken we ons dus tot de idee van de categorische imperatief zelf. Hoe abstract deze ook mag zijn, er zijn imperatieven denkbaar die op het eerste gezicht bij de geest van Kants principe passen, op het tweede gezicht echter onverenigbaar blijken te zijn. Als criterium voor imperatieven en maximes is de categorische imperatief daarom geen holle frase. Al geeft hij niet regelrecht en positief aan wat we moeten doen (gokken, doden, beminnen), hij geeft wel aan waar de grens tussen goed en kwaad ligt. De argumenten die uiteindelijk tot die conclusie voeren weerleggen echter niet alleen de these "Lege Woorden 2", maar ook "Lege Woorden 1".

De vrijwillekeurig gekozen drie pseudo-categorische imperatieven die ik wil voorstellen (en aanvankelijk verdedigen) zal ik C4, C5, en C6 noemen, de grondvormen van Kant C1, C2 en C3. Laten we zeggen dat C4 Rambo's motto is ("A man's got to do what a man's got to do"), C5 me in een droom is geopenbaard, terwijl C6 niet mijn vondst is, maar een in de speltheorie veelbesproken regel:

C4: Handel zo, als je handelen moet (alternatief: doe, wat je doen moet);
 C5: Handel zo, als je niet kunt laten (alternatief: doe wat je niet laten kunt);
 C6: Handel zo dat, als anderen ook zo handelen, dat het meest in je voordeel uitwerkt.

Elk van deze drie alternatieven kan worden opgevat als een categorische imperatief. Allereerst zijn het geen subjectief geldende regels, met andere woorden: het zijn geen maximes maar imperatieven. Bovendien zijn ze niet hypothetisch. Zelfs C6, waarin een conditie ("als", en niet "wanneer", anderen ook zo handelen" - vergelijk "wanneer" en "als" je uit het ziekenhuis komt) wordt gesteld, is een categorische regel, omdat het gedrag van de handelende persoon niet afhankelijk is van hun daadwerkelijk gelijk handelen. Formeel gesproken hebben we blijkbaar te maken met categorische imperatieven.

Ook in inhoudelijk opzicht valt te pleiten voor C4-6. Zo is C4 een korte vorm van C1 en C2, waarin de nadruk wordt gelegd op het vervullen van de plicht in abstracto, dus onafhankelijk van de inhoud van die plicht, van doelen en middelen. Precies zoals Kant eist, stelt C4 dat uitsluitend de wil achter het handelen ethisch relevant is, en dat die wil overeen dient te komen met het idee van de plicht op zich. Bovendien bespaart C4 een fors aantal woorden door het kantiaanse "kunnen willen" uit de formule te schrappen. Die termen zijn overbodig, omdat het, gegeven enerzijds de menselijke vrijheid de wil te bepalen, en anderzijds de afhankelijkheid van alle handelen van een wil, onmogelijk is tegelijk volgens de categorische imperatief te handelen, en dat niet te willen.

C5 is weer afgeleid van C4, en kan op dezelfde gronden verdedigd worden. De idee van de plicht is hier nog sterker aanwezig: "moeten" is "niet kunnen laten" geworden. Dat past uitstekend bij Kants visie op het menselijk handelen. We kunnen immers van alles "laten", elke bepaalde inhoud van de wil afwijzen, alle heteronomie afwerpen, maar als autonoom persoon zijn er twee dingen waar we niet omheen kunnen, die we niet kunnen "laten" -

de noodzaak überhaupt onze wil te bepalen, en het feit dat die wilsbepaling altijd de vorm van een plicht moet hebben.

C6 ten slotte stamt, zoals ik al zei, uit de speltheorie. De maxime is afgeleid van de regel die het gezamenlijk best haalbare resultaat oplevert in dilemma's naar het model van het zogeheten Prisoner's Dilemma. (Let wel: afgeleid van; er zit een addertje onder het gras, maar daarop kom ik later terug.) Stel dat twee dieven, A en B, opgepakt zijn en nu, zonder onderling overleg te hebben kunnen voeren, voor de rechter komen. Als ze niet bekennen, krijgen ze elk 1 jaar gevangenisstraf; bekennen ze beiden dan wordt dat 4 jaar; en als slechts een van beiden bekent, krijgt deze 0 jaar, en de zwijger 6 jaar. In een schema zien de resultaten er zo uit (het eerste cijfer in elke cel geeft de straf van A aan, het tweede die van B):

	B bekent niet	B bekent
A bekent niet	1,1	6,0
A bekent	0,6	4,4

Voor A zou de ideale oplossing natuurlijk zijn dat A bekent en B zes jaar moet zitten; voor B geldt precies het omgekeerde. Voor elk van beiden zou het daarom het verstandigst zijn precies het tegendeel van C6 te doen: handel zo dat, wanneer de ander niet zo handelt, dat in je voordeel uitwerkt. Wanneer ze vervolgens op eigen houtje hun grootste persoonlijke voordeel zouden nastreven (door te bekennen), eindigen ze met een zeer slechte uitkomst: vier jaar elk. Konden ze daarentegen met elkaar overleggen en bindende afspraken maken, dan zouden ze C6 volgen, zwijgen, en er beiden afkomen met een jaar. Voor geen van beiden is dat persoonlijk de beste uitkomst - maar onder de omstandigheden is het wel de best haalbare.

In een recente discussie is weer⁷ de suggestie gewekt dat C6 een interpretatie van de kantiaanse categorische imperatief is - althans: Hans Achterhuis stelt in zijn recensie van Abraham de Swaans *Zorg en de Staat* dat de Swaan dat doet (zie *Krisis* 1990/40, p.10). De Swaans repliek is dubbelzinnig: hij geeft de schuld aan een ander (Rapoport) en laat in het midden of er wel van een fout sprake is (*Krisis* 1990/40, p. 26; zie ook *Zorg en de Staat*, Amsterdam 1989, p. 48 en p. 266 noot 78). Hoe dan ook, er valt iets voor C6 te zeggen. Zoals ik al heb opgemerkt gaat het om een wilsbepaling onafhankelijk van anderen: het doet er niet toe wat zij werkelijk doen of besluiten. (In een

Prisoner's Dilemma is het zelfs voordelig als de ander niet op dezelfde wijze handelt!) Maar een sterker argument ten gunste van C6 is dat het, juist door die onafhankelijkheid van dat wat anderen doen, ook onafhankelijk is van consequenties op zich, ondanks de term "voordeel". C6 is dus geen heteronome maar een autonome regel. Waarom? Omdat de uitkomst er op zich niet toe doet wanneer deze afhankelijk is van het handelen van anderen. Stel dat anderen niet zo handelen dat het in uw voordeel uitwerkt, en u wel - dan nog hebt u C6 gevolgd, dus goed gehandeld, of de consequenties nu voordelig zijn of niet.

Pseudo-categorische imperatieven

Een halve waarheid impliceert een halve leugen. In de vorige paragraaf heb ik liefst drie van dergelijke halve waarheden verteld. Elk van de drie daar voorgestelde imperatieven is weliswaar te verenigen met Kants idee van de onvoorwaardelijke en algemeen geldige ethische regel - maar alleen wanneer ze op een bepaalde manier gelezen worden.⁸ Nu is dubbelzinnigheid niets nieuws, en de dubbelzinnigheid van C4-6 is daarom op zich ook niet interessant. Het is dat wel als illustratie van de stelling dat de categorische imperatief geen holle frase is, zoals critici betogen. Wanneer we overwegen waardoor, om welke redenen, het mogelijk is "echte" van "valse" kantiaanse principes te onderscheiden, doen we een merkwaardige ontdekking: ook Kants grondvormen zijn dubbelzinnig. Er is dan ook een tweede zin waarin we kunnen zeggen dat de categorische imperatief geen lege woorden bevat, en wel in tegenstelling tot Kants overtuiging (these "Lege Woorden 2") dat de tekst van zijn imperatief er niet toe doet, omdat alleen de idee erachter eeuwig, onveranderlijk en daarmee ethisch maatgevend is.

Het meest opvallende verschil tussen C4-5 enerzijds en C1-3 anderzijds is natuurlijk de factor van "de anderen". In de formuleringen C1-3 is steeds een ander aanwezig, als object van wet, natuurwet, of handelen, en daarmee als garantie van de universele geldigheid van de maxime - terwijl C4-5 evengoed subjectieve regels, maximes, konden zijn. C4 en C5 hebben, met andere woorden, een premisse nodig zonder welke we ze niet als categorische imperatieven kunnen herkennen: "de regel geldt voor mij en elk ander persoon". Maar is die ander in de categorische imperatief werkelijk nodig? Hij is immers niet een Ander in de stijl van Levinas (de beroemde ethicus en Kant-criticus), een fundamenteel andere Ander, het Andere überhaupt - maar hij is mijn gelijke en evenbeeld. Deze andere mens, die toch niet werkelijk anders is, is een standaard-vooronderstelling van de westerse

filosofie. Al in de stoïsche metafysica, bijvoorbeeld, werd verondersteld dat het menselijk subject de wereld als object(en) van zich leert onderscheiden, en in sommige objecten zijn gelijke vindt en ervaart, waarmee de basis wordt gelegd voor sociale en politieke verbanden. Deze ander is (zo suggereren in elk geval de Levinas-exegeten) te reduceren tot een ik, een tweede ik dat zich elders in tijd of ruimte mag bevinden, maar in essentie identiek met mij is. In feite zou de westerse filosofie, inclusief die van Kant, daardoor geen Ander erkennen en respecteren, alleen een Ik-subject en een Het-object.

Hoe het ook zij, het begrip "de ander" is er een dat in meer betekenissen blijkt te kunnen worden begrepen. Om te mogen volhouden dat C4-5 met C1-3 verenigbaar zijn volstaat het niet de premisse van gelijkheid tussen mij en de ander te introduceren. We moeten bovendien aannemen dat deze ander de mij-gelijke-ander van Kant is. En dat betekent dat niet alleen de woorden van C4-5 moeten worden geïnterpreteerd om in de geest van Kants idee van een categorische imperatief te worden begrepen: hetzelfde geldt voor die van C1-3.

Voor C4 en C5 geldt bovendien dat hier de termen "moeten" respectievelijk "niet kunnen laten" zo moeten worden uitgelegd dat ze "volg je lusten" uitsluiten, evenals elke vorm van determinisme, daar Kant beide afwijst (cf. BA 43-45). C4 en C5 hebben dus nog een premisse nodig: ze moeten dezelfde opvatting van plicht in zich dragen als C1-3. En dat betekent, nogmaals, dat voor de woorden van C1-3 hetzelfde geldt: ze moeten correct worden geïnterpreteerd.

C6 levert gelijksoortige moeilijkheden op. De regel moet zo worden uitgelegd dat de consequentie van het handelen voor de ethische waarde van het handelen irrelevant is. Maar helemaal juist is dat toch niet: het handelen volgens C6 blijft gericht op wat, persoonlijk en tijdelijk, voordelig wordt genoemd. Op zich is dat geen probleem. Ook Kant weet heel goed dat mensen niet anders kunnen dan handelen met zekere middelen tot zekere doelen, die op hun beurt worden gewaardeerd als al of niet subjectief voordelig. Waar Kant zich mee bemoeit is de interpretatie van het begrip voordeel. C6 is een regel voor "morele egoïsten", zoals Amartya Sen dergelijke lange termijn-denkers noemde. Hun zelfzuchtige motief kan goed-kantiaans leven en handelen zijn, of rijkdom, of lustbevrediging, of vrijheid van gevangenisstraf - dat laat C6 in het midden. Kant wil echter onderscheid maken tussen het efficiënt handelen van de arts en dat van de gifmenger (BA 41) door universaliseerbaarheid en onvoorwaardelijkheid

van hun maximes te verlangen. C6 voldoet daar niet zonder meer aan: je kunt niet willen dat anderen die regel ook toepassen, tenzij je weet dat hun voordeel jouw voordeel is.

Om C6 met C1-3 onder een noemer te brengen is dus weer een premisse nodig, en ditmaal een forse en ingewikkelde. Het “voordeel” in C6 moet een ethisch onvoorwaardelijk en universeel voordeel in kantiaanse zin zijn. De enige voorstelbare invulling daarvan is het voordeel voor (om C3 aan te halen) de essentie van de persoon als doel op zich, ofwel het voordeel voor de autonomie van de wil. Voor de derde keer impliceert dit dat ook de letter van Kants principes geïnterpreteerd moet worden naar de geest.

Voor wat betreft de onenigheid tussen Achterhuis en de Swaan c.s. kunnen we daarom concluderen dat het gelijk aan Achterhuis' kant is. C6 is geen interpretatie of verschijningsvorm van de categorische imperatief, het is eerder omgekeerd: de categorische imperatief kan worden begrepen als een verbijzondering van C6. C6 laat in het midden wat A en B als voordeel beschouwen. Als dat voor beiden bewegingsvrijheid is, levert C6 in het Prisoner's Dilemma 1 jaar gevangenisstraf op bij samenwerking, 4 jaar bij individueel optreden, of (voor een van beiden de meest voordelige uitkomst) 0 respectievelijk zes jaar bij het breken van een afspraak. Als A en B daarentegen kantianen zijn, en C1-3 volgen, is er maar een uitkomst mogelijk: bekennen en vier jaar uitzitten. In de originele speltheoretische versie van C6 moet bovendien “wanneer” in plaats van “als” staan. Daarmee wordt de “goedheid” van het eigen handelen afhankelijk van wat anderen doen: als zij, achteraf, anders blijken te hebben gehandeld dan ik verwachtte, heb ik “slecht” gehandeld. Of, in het geval van het Prisoner's Dilemma: als A en B samen besluiten te zwijgen, en B breekt de overeenkomst door toch te spreken (B krijgt 0, A 6 jaar) - dan heeft A slecht gehandeld. Alleen al daarom is de speltheoretische variant van C6 niet met de categorische imperatief verenigbaar.

Het is duidelijk dat de these “Lege Woorden 1” weerlegd kan worden. De categorische imperatief laat zich niet eindeloos oprekken tot alle denkbare maximes erin onder te brengen zijn - niet zonder premissen die een imperatief als C5 iets geheel anders laten betekenen dan hij in het dagelijks leven betekent: een oproep tot instant lustbevrediging, of de permissie daartoe. Met dergelijke premissen komen we tot geheel andere oordelen over “goed of slecht volgens deze imperatief” dan zonder. Met andere woorden: het universum van het door de categorische imperatief gerecht-

vaardigde is eindig.

Maar ook deze “Lege Woorden 2” moet verworpen worden. Het onderscheidend vermogen, dat we zoëven aan de categorische imperatief moesten toeschrijven, brengt met zich mee dat de exacte bewoording van Kants principe wel degelijk relevant is. Een onder woorden gebrachte categorische imperatief is niet meer strikt formeel, maar iets dat zelf geïnterpreteerd moet worden.

Zolang we aan de tekst van een categorische imperatief dezelfde interpretatie geven als Kant, bestaat er geen aanleiding vraagtekens te zetten bij de hypothese dat woorden probleemloos verwijzen naar begrippen. Wanneer woorden echter iets anders gaan betekenen voor verschillende mensen (zoals we mogen aannemen dat het geval is bij Kant en Levinas met betrekking tot “de Ander”, of bij Kant en speltheoretici over autonomie) dan is er een probleem. Dan blijken woorden namelijk niet naar buitenmenselijke, eeuwige begrippen te verwijzen, niet naar entiteiten die “anders” zijn dan wij, maar juist naar door de mens gemaakte en voor adequaat versleten begrippen, zelf- of sociaal geconstrueerde concepties van de “ware” werkelijkheid - of misschien zelfs wel naar concepties die onbedoeld, “onveroorzaakt”, in en door menselijke interactie zijn ontstaan en evolueren. Kant heeft met zijn categorische imperatief precies gedaan wat Marx als de competentie van de filosoof beschouwde: hij heeft de wereld geïnterpreteerd, en niet, zoals hij zelf geloofde, haar verklaard.

Noten

1. Met dank aan Martin V.B.P.M. van Hees, Erik Heijerman, en Grahame E. Lock.
2. Verwijzingen geschieden naar pagina's van de *Grundlegung zur Metaphysik der Sitten*, resp. de eerste (A) en tweede (B) editie van 1785 en 1786. Varianten van deze eerste grondvorm van de categorische imperatief komen voor op BA 17, BA 70, BA 76, BA 81 (tweemaal), BA 82, BA 124-125, en ten slotte in de *Kritik der praktischen Vernunft*, eerste oplage (1788), p. 54.
3. Vergelijk de formulering op BA 82, waar Kant stelt dat de wil een regel moet volgen die “algemeen geldig” zou moeten zijn, waarmee hij het midden houdt tussen wet en natuurwet.
4. Deontologie (van het Griekse deon = plicht, “het nodige”) beoordeelt handelingen alleen op hun intrinsieke waarde; het tegenovergestelde con-

sequentialisme kijkt alleen naar de gevolgen van het handelen.

5. Voor een bondige weergave van de belangrijkste posities in het debat over de discrepantie tussen de eerste twee grondvormen en de derde zie bv. William A. Galston, *Justice and the human good*, Chicago 1980, p. 38.

6. Een van de bekendste Kant-critici van deze eerste lijn was Henry Sidgwick, met "The kantian conception of freedom", pp. 511-516 in zijn *Methods of Ethics* (1901⁶).

7. Oudere voorbeelden zijn bij voorbeeld Richard Brandt in zijn *A Theory of the Good and the Right*, Oxford 1979, p. 278, en William Galston impliciet in *Justice and the human good*, p. 40.

8. Ik wil niet beweren dat er maar één interpretatie van Kant mogelijk is - waar het om gaat is dat sommige interpretaties uitgesloten zijn.