

De ethiek van Christine Korsgaard

Bachelorscriptie / leeronderzoek Wijsbegeerte

Gerard Versteegen

Begeleiding: Katrien Schaubroeck

3^e versie, 11 oktober 2012

Document: Paper Korsgaard v20

10.799 woorden

1 Inleiding

Twee onlosmakelijk met elkaar verbonden vragen spelen in de ethiek een belangrijke rol. De ene is: wat is goed en kwaad, hoe kan ik beide onderscheiden. Een antwoord op deze vraag geeft helderheid over de inhoudelijkheid, de *substantialiteit* van morele claims die we onszelf en elkaar opleggen. De andere vraag is: waarom zouden we ons iets gelegen laten liggen aan deze morele claims. Beantwoording van deze vraag moet ons iets vertellen over de *fundering*, de rechtvaardiging van die claims. Waaraan ontleen ze hun autoriteit?

Eén van de hedendaagse moraalfilosofen die een aan Kant ontleend en op belangrijke punten nieuw, eigen antwoord heeft ontwikkeld op deze vragen is Christine Korsgaard. De kernvraag die leidend is geweest in dit paper, is: hoe valide, hoe 'waar' is haar theorie? 'Waar' hier niet in de betekenis van de werkelijkheid representerend maar in de zin zoals Rorty over waarheid sprak (Rorty, 1995): hoe *overtuigend* is de onderbouwing, de rechtvaardiging van haar model, van de claims die zij daarin doet?

Korsgaard heeft een nieuwe interpretatie van Kant's moraalfilosofie ontwikkeld. Ze noemt de vraag naar de rechtvaardiging van morele claims de 'normatieve vraag'. Deze vraag speelt een belangrijke rol in haar gehele oeuvre, en staat centraal in *Sources of Normativity* (Korsgaard, 1996a, in verwijzingen verder afgekort als SoN).

Ze heeft een model gebouwd waarin ze niet alleen een antwoord geeft op deze normatieve vraag maar tegelijk een theorie ontwikkelt over de rationaliteit van ons handelen en de bron van onze waarden. Haar model is sterk gebaseerd op het ethisch denken van Kant. Maar op een aantal punten wijkt ze van Kant af, punten die nog aan de orde komen. Dat leidt tot een andere, nieuwe fundering van moraliteit, een ander antwoord op de vraag: waaraan ontleen morele claims hun autoriteit? Het leidt ook tot een andere invulling van die morele claims, waarbij naast plichten en respect ook persoonlijke relaties en genegenheid als moreel relevant worden gezien.

De centrale claim in Korsgaard's *Sources of Normativity* is dat de rechtvaardiging van onze morele verplichtingen is af te leiden uit onze reflexieve structuur. Wij hebben, als mens, de vrijheid om al dan niet verlangens die in ons op komen te onderschrijven. Uit dit vermogen tot zelfreflectie leidt zij, in een aantal stappen, onze morele plichten af. Ik noem dit verder de *funderingsclaim*.

De vraag die ik in dit paper probeer te beantwoorden is: hoe succesvol, hoe overtuigend is Korsgaard in het verdedigen van haar funderingsclaim? Mijn doelstelling is om aannemelijk te maken dat Korsgaard in de onderbouwing van haar theorie op enkele wezenlijke punten niet overtuigend is, en dat zij tegelijk ons een nieuwe interpretatie biedt van Kant, een interpretatie die aan zijn denken een nieuwe, meer op menselijke relaties gerichte inkleuring geeft en die ons enkele belangrijke nieuwe ethische inzichten verschaft.

Dit paper is verder als volgt opgebouwd.

De paragrafen 2 en 3 omvatten een korte reconstructie van de belangrijkste componenten van Korsgaard's theorie. In deze reconstructie heeft *Sources of Normativity* een centrale rol

gespeeld, en heb ik me daarnaast gebaseerd op *Creating the Kingdom of Ends* (Korsgaard 1996b) en *The Constitution of Agency* (Korsgaard 2008). In paragraaf 2 heb ik geprobeerd de kern te schetsen van Korsgaard's theorie rond handelen en moraliteit. In paragraaf 3 besteed ik speciale aandacht aan drie nauw onderling samenhangende thema's van haar werk, namelijk aan waarden (waar komen die volgens Korsgaard vandaan?), aan haar meta-ethische positie, constructivisme, en aan substantialiteit (welke inhoud geeft Korsgaard aan moraliteit?).

In paragraaf 4 richt ik me op drie kwetsbaarheden in haar theorie.

De eerste betreft de vraag in hoeverre binnen haar model rationeel handelen en immoreel gedrag kunnen samengaan. In mijn interpretatie laat zij die mogelijkheid open. Dit leidt tot de mogelijkheid van immorele redenen, en tast de normatieve kracht van haar model aan.

De tweede kwetsbaarheid betreft de door Korsgaard veronderstelde vrijheid die ieder van ons heeft om zelf onze eigen waarden, doeleinden, en daarmee onze eigen identiteit te kiezen. Ik vecht die vrijheid aan, mede gebaseerd op argumenten van Charles Taylor en Harry Frankfurt. Een derde kritiek gaat over haar fundering van de morele wet. Zij stoelt die op de claim van deelbaarheid van redenen. Ik stel daar tegenover dat immorele redenen moeilijk als deelbaar denkbaar zijn.

De nieuwe inzichten die Korsgaard ons biedt komen impliciet naar voren in de loop van mijn argumentatie in de paragrafen 2, 3 en 4. Deze betreffen haar toepassing van het concept praktische identiteit, haar claim over de principiële onoplosbaarheid van bepaalde morele vraagstukken, en haar focus op menselijke relaties als bron van moraliteit.

In de laatste paragraaf, conclusies, vat ik naast mijn kritiekpunten deze nieuwe inzichten samen.

2. De kern van Korsgaard's model

De startsituatie

In lijn met Kant ligt het startpunt van Korsgaard's onderzoek naar de bronnen van normativiteit bij de reflexieve aard van de mens. Dat illustreert ze aan de hand van het onderscheid tussen ons denken en dat van een dier.

Voor een dier vormen waarnemingen zijn overtuigingen, en verlangens vormen zijn wil. Een dier is wel betrokken in bewuste activiteiten, maar is zich daar niet, op een hoger niveau, van bewust (SoN 93). Wij, mensen, kunnen wel onze aandacht richten op onze overtuigingen en onze verlangens. We kunnen hierover nadenken. Dat creëert een probleem, het probleem van normativiteit. Stel, ik heb een sterk verlangen, en krijg daarmee een sterke impuls om te handelen. Maar ik doe een stap terug, en kijk naar de impuls. Nu domineert de impuls me niet meer, en ik heb een probleem: zal ik handelen? Kan ik de impuls ondersteunen? Zo ja, dan heb ik een *reden* om te handelen. Zo nee, dan heb ik de *plicht* om af te zien van de handeling¹.

Korsgaard gaat ervan uit dat wij, mensen, vanwege onze aard noodzakelijkerwijs handelen uit redenen, dus uit verlangens die we na reflexieve toetsing hebben geaccepteerd.

Dit proces van reflexieve acceptatie geeft ons *vrijheid*. We worden niet meer geleefd door onze verlangens maar maken zelf keuzes over welke impulsen onze drijfveren zijn. In dit model spreekt Korsgaard over twee zelden: het denkende zelf, dat reflecteert en kiest, en het handelende zelf, dat aan de ene kant impulsen ervaart en aan de andere kant wordt aangesproken door het denkende zelf.

¹ Korsgaard hanteert 'verplichting' (obligation) in een betekenis die afwijkt van de meer gangbare betekenis, in die zin dat verplichtingen voor haar alleen verwijzen naar situaties waarin reflectie leidt tot het afwijzen van een impuls, dus tot verboden (bijvoorbeeld SoN 102, 113). Aangezien dit verder geen relevante rol speelt in dit paper ga ik hier niet verder op in.

Korsgaard's Categorisch Imperatief

De vraag die nu open staat is: hoe bepaal ik welke impulsen ik wel en niet zal ondersteunen? Op grond van welke criteria doe ik dat?

Kant's antwoord daarop is de Categorisch Imperatief (verder afgekort CI). Korsgaard volgt Kant daarin, maar wijkt op een wezenlijk punt af van het Kant's originele versie.

Kant gaat in zijn afleiding uit van de vrije wil van de mens. Daaronder verstaat hij een rationele causaliteit die effectief is zonder te zijn bepaald door een externe oorzaak. De twee punten waar het hier om draait zijn *geen externe oorzaak* en *causaliteit*.

Als externe oorzaak wordt alles beschouwd buiten de wil van de mens – inclusief zijn neigingen en verlangens. Dus is de vrije wil volledig zelf bepalend.

Daarnaast is de vrije wil een causaliteit, en dus, zegt Kant, gebonden aan wetten. Vrijheid is derhalve iets anders dan wetteloosheid. De mens handelt uit redenen. Die redenen moeten zijn afgeleid van principes (die Kant maximes noemt, de wetten die de mens zichzelf oplegt). Waar komen die principes vandaan? Niet van buiten, dan zou de mens niet meer vrij zijn. Dus de mens moet zichzelf zijn eigen principes voorschrijven. En daarmee is de mens autonoom: hij schrijft zichzelf de wet.

Uit deze autonomie leidt Kant rechtstreeks zijn CI af, die luidt: handel alleen volgens die maximes waarvan je zou kunnen willen dat het een universele wet is - dus een wet die voor iedereen zou kunnen gelden. Die eis, potentieel geldig voor iedereen, maakt zijn CI tot een moreel principe. Voor Korsgaard is die afleiding problematisch. In haar ogen is uit onze autonomie slechts af te leiden dat wij volgens wetten handelen. Haar CI heeft daardoor een beperktere claim: handel alleen volgens maximes waarvan je kunt willen dat het een wet is (SoN 98).

De maximes of principes die ons handelen sturen moeten dus tot wetmatigheid leiden en ze moeten een zekere mate van constantheid bezitten. Korsgaard noemt dit het constitutieve principe van rationele activiteit (Korsgaard 2008, 12).

Korsgaard's CI eist dus nog niet dat onze maximes geldig zouden kunnen zijn voor alle mensen, en behelst dus nog geen morele component. Daarvoor introduceert zij als tweede stap de 'morele wet'. Deze komt dicht in de buurt van Kant's CI en luidt: 'handel alleen volgens maximes waarover alle rationele wezens het eens zouden kunnen worden om alleen op grond daarvan te handelen, in een werkbaar coöperatief systeem' (SoN 99).

Ze biedt een eigen onderbouwing voor deze morele wet. Die verloopt in twee stappen: eerst introduceert ze het begrip 'praktische identiteit', daarna geeft ze drie argumenten als onderbouwing voor de morele wet.

Het begrip praktische identiteit

Onze reflexieve structuur dwingt ons om een conceptie te hebben van onszelf, van onze eigen identiteit. Mijn denkende zelf besluit in een proces van reflectie welke verlangens het wel of niet zal volgen op grond van principes, wetten die ik mezelf heb opgelegd. Ik identificeer mezelf met die principes, ze vormen een uitdrukking van wie ik ben (SoN 100).

Deze conceptie van identiteit is niet theoretisch maar praktisch. Het is niet kennis, wetenschap over feitelijkheden die mijzelf betreffen. Nee, het verwijst naar een beschrijving waaronder ik mijn leven waard vind om geleefd te worden, en mijn handelingen waard om uit te voeren. Praktische identiteit is een complex begrip, de gemiddelde mens heeft vele van zulke concepties. Korsgaard noemt verschillende voorbeelden. Je ziet jezelf als mens, burger van het Koninkrijk van Doeleinden, de door Kant als ideaal geschetste morele gemeenschap van alle mensen, die allen de morele wet als principe hanteren. Je denkt over jezelf als vader of moeder. Je ziet jezelf als pure behartiger van alleen je eigen belangen (en bent een egoïst). Ieder van die identiteiten reflecteert een set van principes of waarden, en leidt tot redenen en verplichtingen.

Redenen en verplichtingen zijn dus verbonden aan onze praktische identiteit, en ieder mens heeft normaal gesproken een veelheid aan identiteiten, vanuit de verschillende rollen die hij heeft – die hij 'is' -, en die bepaald worden door de gemeenschappen waarvan hij deel uitmaakt. Korsgaard is het ook tot op zekere hoogte eens met het communitarisme, met het idee dat een mens wordt gevormd en gedefinieerd door de gemeenschappen waartoe hij behoort (SoN 120)².

Deze constructie van het begrip praktische identiteit gebruikt Korsgaard niet alleen als opstap voor haar fundering van de morele wet (die komt hierna aan de orde). Het wordt daarmee voor haar ook mogelijk om enkele wezenlijke praktische problemen van ethisch handelen te benoemen en te verklaren, die veel andere moraalfilosofen niet aan de orde stellen.

Onze veelheid aan rollen leidt tot verschillende redenen en verplichtingen, die onderling tegengesteld kunnen zijn. Het bekende voorbeeld van Sartre: moet ik als soldaat mijn land dienen of thuis blijven om mijn zieke moeder te verzorgen.

Iedere verplichting, zegt Korsgaard, is ongeconditioneerd. Iedere verzaking van een verplichting leidt tot het verzaken van je eigen principe, en ontkracht wie je bent. Hoe kunnen we dan, met tegengestelde verplichtingen die uit verschillende rollen voortkomen, onze identiteit bewaren? Korsgaard biedt een uitweg in twee stappen. In de eerste plaats betekent het eenmalig schenden van een principe niet meteen volledig verwerpen. Je kunt soms van een principe afwijken zonder het principe los te laten. Daarbij zijn er natuurlijk wel grenzen aan hoe vaak je een uitzondering maakt. Dus het verzaken van een principe leidt niet meteen tot identiteitsverlies. Daarnaast moet je onderscheid maken tussen verschillende verplichtingen. De ene is veel 'dieper' dan de ander, omdat het principes betreft die voor jou van veel groter belang zijn. Dat betekent dat het verlies van een beperkt deel van je identiteit acceptabel kan zijn, als consequentie van een commitment aan je meest belangrijke principes.

Onderbouwing van de morele wet

Hoe komen we vanuit Korsgaard's CI tot een morele wet, tot een invulling van wat juist is? Door te kijken naar die praktische identiteit die we allemaal gemeen hebben, namelijk ons mens-zijn, in Kant's termen ons burgerschap van het Koninkrijk der Doeleinden. Welke wetten, welke principes zou zo'n burger moeten adopteren? Het antwoord, de morele wet, heb ik hierboven al aangeduid. Hoe leidt Korsgaard deze af, hoe fundeert ze deze?

Daarvoor geeft ze, in mijn interpretatie van *Sources of Normativity*, drie deels op elkaar aansluitende argumenten, die ik zal aanduiden met het transcendentale argument, het naturalistische argument en het taal-argument.

Het transcendentale argument dat Korsgaard schetst is een direct afgeleide van Kant's strategie om tot een a priori bewijsvoering te komen van een propositie P door (1) aan te tonen dat omstandigheid O een logische oorzaak of conditie van P is, in de zin dat $O \rightarrow P$, en (2) te laten zien dat niet-O niet denkbaar is. Korsgaard volgt in deze argumentatie de volgende stappen (SoN 123-124). (1) De mens handelt noodzakelijkerwijs uit redenen, die voortvloeien uit specifieke concepties van zijn identiteit. (2) Die noodzaak vloeit voort uit de aard van de mens: een zelfreflexief wezen dat redenen nodig heeft om te handelen. (3) Dus je moet je menselijkheid waarderen om überhaupt iets te kunnen waarderen. (4) Dus we moeten mensen waarderen, het zijn doeleinden in zichzelf.

Korsgaard acht dit argument op zichzelf ontoereikend. Voorafgaand aan de beschrijving (SoN 122 – 125) geeft ze een conditie aan die ze aan het argument verbindt: 'Voorlopig zal ik aannemen dat *onzelf* als mens waarderen met zich meebrengt *andere mensen* op dezelfde wijze te waarderen' (SoN 121, cursivering van mij). Dit is de cruciale laatste stap van haar argumentatie, die juist nodig is om het argument tot een morele conclusie laten te komen.

² Tot op zekere hoogte. Ze wijkt van het communitarisme af door ook de mensheid in zijn geheel als gemeenschap te zien, en daar morele waarden vanaf te leiden, wat hierna verder wordt toegelicht.

Later (SoN 134) geeft ze aan: het argument toont wel aan dat mijn menselijkheid normatief is voor mij, en dat jouw menselijkheid normatief voor jou. *Het toont echter niet aan dat jouw menselijkheid normatief is voor mij.*

Omdat ze het transcendentaal argument niet accepteert moet ze een ander onderbouwing ontwikkelen voor deze laatste, cruciale, stap. Daarvoor kiest ze een onverwachte strategie. Ze streeft ernaar om aan te tonen dat redenen naar hun aard niet privé zijn maar publiek. Als dat immers lukt dan worden jouw redenen en jouw menselijkheid normatief voor mij.

Privéredenen definieert ze als redenen die iemand heeft en die alleen voor hem of haarzelf normatief zijn. Publieke redenen zijn redenen die voor iedereen normatieve kracht hebben. De Kantiaanse argumentatie probeert het publieke karakter van redenen aan te tonen, vanuit de aanname dat redenen privé zijn. Die aanname bestrijdt Korsgaard. Redenen, zegt ze, zijn in hun essentie publiek (SoN 135). Dat publieke karakter vloeit voort uit het wederzijds uitwisselen, het *delen* van redenen. Publieke redenen zijn deelbare redenen, en redenen zijn op z'n hoogst tijdelijk privé, ze zijn inherent deelbaar.

Korsgaard onderbouwt deze claim op twee manieren.

Ze start met wat ik heb genoemd het *naturalistische argument*. Het delen van redenen, zegt ze, vloeit rechtstreeks voort uit onze *sociale aard*. Wij mensen zijn sociale dieren, als gevolg daarvan zijn onze redenen publiek en deelbaar (SoN 135/136).

Het is goed mogelijk dat Korsgaard dit argument niet heeft bedoeld ter onderbouwing van een rechtvaardiging van moraliteit. Wat daarvoor pleit is haar afsluiting van dit argument: 'Dus het soort argument dat we nodig hebben is [...] er één die aantoonde dat onze redenen nooit op voorhand privéredenen zijn geweest', daarmee implicerend dat er blijkbaar nog geen argument is gegeven in het voorafgaande (SoN 136). Maar een andere lezing lijkt ook goed verdedigbaar, kijkend naar haar verdere toelichting: 'Hoewel het algemeen erkend is dat mensen in feite sociale dieren zijn, achten moderne moraalfilosofen het normaal gesproken niet toegelaten om dit feit te gebruiken in argumenten die zich richten op het rechtvaardigen van moraliteit. [...] Maar als onze sociale natuur diep is, dan kan en hoort een justificatie van moraliteit er een beroep op te doen' (SoN 135/136).

Daarna brengt ze haar tweede (of, als het naturalistische argument nooit als argument bedoeld is geweest haar eerste) argument in om de deelbaarheid van redenen te onderbouwen, het *taalargument*. Daarin trekt ze een parallel met Wittgenstein's punt over het publieke karakter van een taal. Wittgenstein zegt dat er niet iets als een privétaal kan bestaan. Betekenis is relationeel want het is een normatief concept. Als je zegt 'een kat is een zoogdier' dan zeg je dat men een kat *hoort* te rekenen tot de groep van zoogdieren. En dat vooronderstelt minstens twee entiteiten: een 'wetgever' die de claim uit, en een 'burger' die zou moeten gehoorzamen. Op precies dezelfde wijze zijn ook redenen relationeel en normatief. Zeggen dat R een reden is om A te doen is zeggen dat je A *hoort* te doen om reden R. En ook nu heb je een wetgever nodig, die de reden uitvaardigt, en een burger, die A hoort uit te voeren.

Betekeningen en waarden zijn geen mentale activiteiten, ze hebben betrekking op *relaties* die we hebben met onszelf en met anderen.

Alle redenen zijn dus deelbaar.

Hoe werkt dat in de ogen van Korsgaard in de praktijk?

Als we een verlangen ervaren, kijken we of we dit al dan niet willen ondersteunen, en als een reden tot handelen willen accepteren. Normaal gesproken zullen we dat doen, tenzij er redenen zijn om dat niet te doen. Onze gewone impulsen hebben in die zin een bepaalde status, namelijk een soort automatisch recht om minstens gehoord te worden. *Redenen van anderen* hebben een zelfde soort status. We hebben, zegt Korsgaard, geen redenen nodig om rekening te houden met de redenen van anderen. We hebben een reden nodig om daar *niet* rekening mee te houden. Als meest eenvoudige voorbeeld schetst ze de situatie dat als ik jou roep, terwijl je wegloopt, ik jou daarmee een reden geef om te stoppen (SoN 140).

Deze onderbouwing van moraliteit, op basis van de deelbaarheid van redenen onderstreept haar visie op de kern van moraliteit, die ligt in de *relaties tussen* mensen. Elders zegt ze (Korsgaard 1996c, 275): ‘Bij moraliteit gaat het in de kern niet om situaties waarin ik jou iets doe of jij mij iets doet, maar om situaties waarin we samen iets doen. Het gaat er niet om wat we te weeg brengen, maar hoe we ons tot elkaar verhouden.’

Persoonlijke relaties

In lijn met Korsgaard's tot nu toe geschetste benadering heeft ze ook een eigen, van vele andere neo-Kantianen afwijkende, visie op het morele belang van persoonlijke, affectieve relaties. In veel interpretaties van Kant worden redenen voortvloeiend uit genegenheid of sympathie niet beschouwd als morele redenen, omdat ze niet worden gezien als plichten.

Een voorbeeld van een hedendaagse moraalfilosoof die genegenheid afwijst als morele drijfveer is Darwall. Dat blijkt onder meer wanneer hij redenen gebaseerd op respect vergelijkt met redenen voortvloeiend uit zorg (Darwall 2006, pag. 126 – 130). Deze laatste wijst hij af als morele redenen, omdat hij ze beschouwt als agent-neutraal, en dus niet ‘second-personal’³. Hij licht dit als volgt toe: wanneer iemand handelt om redenen van zorg (‘in feeling sympathetic concern’), ‘wordt hij niet gedreven door wat de ander waardeert of prefereert (zoals bij respect), maar door wat je zelf denkt dat de ander echt zou baten’. Hij kwalificeert redenen van zorg als naar hun aard consequentialistisch.

Korsgaard schets een beeld van een persoonlijke relatie dat sterk afwijkt van dat van Darwall. Bij Darwall staat de ander, ook in een persoonlijke relatie, op dusdanige afstand dat ik niet kan weten, maar hoogstens vermoeden wat de ander wil, drijft, prefereert. Daarmee wordt handelen uit sympathie een kwestie van inschatten van wat voor de ander het beste is, het wordt consequentialistisch en valt daarmee buiten het bereik van de moraal.

Voor Korsgaard is juist het *kenmerk* van een persoonlijke relatie dat je de drijfveren van de ander kent. Ze beschrijft een persoonlijke relatie als een wederzijds commitment om rekening te houden met elkaars beelden, belangen en wensen. Het delen van redenen, voor Korsgaard het hart van moraliteit, is juist waar het in een persoonlijke relatie om gaat. Ze vertaalt dit in Kantiaanse termen door te spreken over een ‘Koninkrijk van Twee’, als een miniversie van Kant's ideaal, een Koninkrijk van Doeleinden dat alle mensen omvat. Persoonlijke relaties vormen een kernonderdeel van de verschillende concepties van iemands praktische identiteit, in de rol van bijvoorbeeld ouder, kind of vriend. Ze zijn *constitutief* voor onze praktische identiteit (SoN 127). En daarmee vormen ze een bron van verplichtingen. Die kunnen conflicteren met andere, morele, verplichtingen. Korsgaard zegt daarover (SoN 128): ‘*er is geen reden waarom je relatie met de menselijkheid in de wereld altijd van groter belang voor je zou moeten zijn dan je relatie met een individueel persoon. Waarom zouden de wetten van het Koninkrijk van Doeleinden meer kracht hebben dan de wetten van een Koninkrijk van Twee?*’. Daarmee, zo interpreteer ik haar retorische vraag, lijkt ze aan de verplichtingen voortvloeiend uit persoonlijke relaties dezelfde normatieve kracht toe te kennen als aan algemene morele verplichtingen, voortvloeiend uit de menselijkheid van iedere ander.

3. Waarden, constructivisme en substantialiteit

De probleemstelling van dit paper gaat over de vraag hoe succesvol Korsgaard is in het verdedigen van de funderingsclaim van haar ethisch model. Ik heb in het kort hierboven een reconstructie geschetst van dit model. Voordat ik de vraag van de probleemstelling beantwoord wil ik nog nader stilstaan bij drie nauw onderling samenhangende thema's in haar werk : (1) wat

³ ‘Second personal’ verwijst naar Darwall's claim dat handelingen hun morele waarde ontleen aan claims van de ander, in zijn denken zijn morele redenen naar hun aard altijd second-personal

is waarde, oftewel wat maakt iets waardevol? (2) wat houdt haar meta-ethische positie, die ze zelf aanduidt met constructivisme, in? en (3) wat is het juiste?, de vraag naar de substantialiteit van moraliteit.

Waarden

Wij mensen hechten waarde aan bepaalde dingen. Waar komt die waarde vandaan? Korsgaard zegt: die komt van onszelf, wij projecteren waarde op de wereld. Wat van waarde is, het goede, construeren wij zelf. Het is gegrond in onze zelfreflexieve aard. Onze reflectie is daarbij dus niet theoretisch, in de zin dat deze kennis ontleend wordt aan de wereld. Hij is praktisch en richt zich op ons eigen handelen.

Onze waarden zijn inherent verweven met de concepties van onze identiteit. Ik identificeer mezelf in verschillende rollen (als vader, student, etc.) met verschillende sets van waarden. Die identiteiten zijn voor een belangrijk deel contingent. Je bent geboren in een land, met een bepaalde cultuur. Je bent het kind van jouw ouders. Je bent misschien van kinds af aan opgevoed in een bepaald geloof. Ieder van die rollen brengt bepaalde waarden met zich mee, waarden waarvoor je niet vooraf bewust gekozen hebt maar die onlosmakelijk verbonden zijn met die identiteiten.

Toch, zegt Korsgaard, kun je je van die waarden losmaken, juist omdat ze contingent zijn. 'Je kunt stoppen om jezelf als een moeder te zien of als een Quaker. Of, als de feiten dat onmogelijk maken, kan de betreffende identiteit zijn normatieve kracht verliezen' (SoN 120).

Je waarden, je doeleinden zijn voor Korsgaard dus altijd een kwestie waarover je zelf een bewuste keuze kunt maken, waarover je zelf vrijheid hebt.

Hierop is één uitzondering. Die betreft de enige conceptie van je identiteit die niet contingent is, die van mens. Je identiteit als zelfreflexief wezen kun je niet van je afzetten – en dus ook niet de daaraan verbonden morele waarden.

Constructivisme

Korsgaard noemt haar visie op de bron van waarden en van normativiteit constructivisme. Verschillende auteurs geven verschillende interpretaties over wat constructivisme inhoudt, en ieder voor zich onderscheiden ze ook verschillende varianten.

Ik kies Korsgaard's eigen benadering als uitgangspunt, en verduidelijk deze nader door gebruik te maken van een onderscheid dat Sharon Street maakt (Street 2009).

Korsgaard maakt haar positie als constructivist duidelijk in haar paper *Constructivism and Realism* (Korsgaard 2008, 302 e.v.) door deze te vergelijken met moreel realisme.

Daarbij hanteert ze een interpretatie van realisme die ze elders substantieel realisme noemt (SoN 112). Ze doelt niet op de visie dat morele proposities waar kunnen zijn – in die betekenis valt haar constructivisme ook onder realisme (en noemt deze procedureel realisme). Het gaat haar om de vraag *waarom* morele claims waar kunnen zijn. Substantieel realisme verwijst dan naar de opvatting dat proposities die morele concepten gebruiken waar of onwaar kunnen zijn, omdat die concepten verwijzen naar normatieve feiten of entiteiten die bestaan onafhankelijk van die concepten zelf.

Korsgaard wijst die visie af, en is in die zin een anti-realist. Bij ethiek gaat het niet om het vinden van kennis, en om het vervolgens toepassen van die kennis in de praktijk. Morele concepten zijn geen namen van feiten of situaties die we tegenkomen in de wereld. Ze verwijzen naar de oplossingen van problemen. Het concept 'het juiste' verwijst naar het probleem hoe we ons moeten gedragen. En 'het goede' is de naam voor een oplossing voor de vraag hoe we moeten kiezen tussen opties, of tussen doelstellingen.

Zo'n concept is op zichzelf slechts een verwijzing, en heeft nog geen inhoud. Daarvoor hebben we *concepties* nodig. Een conceptie is een principe dat een oplossing kan bieden voor het probleem. Rawls heeft het begrip constructivisme gemunt. Korsgaard gebruikt zijn conceptie van rechtvaardigheid ter illustratie. Het concept rechtvaardigheid verwijst naar de oplossing van

het probleem hoe mensen die samenwerken hun lusten en lasten moeten verdelen. Rawls' antwoord daarop, zijn conceptie, is zijn Original position – procedure.

Korsgaard zegt: onze morele concepten *verwijzen* niet naar feiten, ze *construeren* een menselijke realiteit, gebruik makend van de juiste conceptie, waarmee ze het probleem oplossen waaruit het concept voortvloeit (Korsgaard 2008, 324).

Een conceptie ontleent zijn normatieve kracht aan de mate waarin deze in staat is om voor het probleem dat voor handen is de beste (of enige) oplossing te geven.

Tot dusver Korsgaard zelf. Anderen geven een andere invulling van het begrip, zonder daarbij overigens in conflict te komen met Korsgaard's benadering. Sharon Street beschrijft constructivisme als de visie, binnen de ethiek, dat de waarheid van een normatieve claim wordt afgeleid door een bepaald praktisch gezichtspunt in te nemen. Een onderscheid dat Street maakt en hier van belang is betreft twee varianten, die zij aanduidt als beperkt ('restricted') constructivisme en meta-ethisch constructivisme (wat ze ook wel aanduidt als diepgaand, 'thoroughgoing') (Street, 2009). Dit onderscheid betreft de vraag in hoeverre, binnen het praktische gezichtspunt dat als uitgangspunt fungeert, er substantieve normatieve elementen zijn meegenomen. Die normatieve elementen leiden tot een beperking van de set van normatieve claims die het gezichtspunt oplevert.

Rawls' rechtvaardigheidstheorie is een voorbeeld van beperkt constructivisme. Ingebed in het door hem geconstrueerde praktische gezichtspunt, de originele positie, vind je normatieve elementen als oordelen over de aard van eerlijke onderhandelingscondities, de vrijheid en gelijkheid van mensen, en de irrelevantie, vanuit moreel oogpunt, van individuele eigenschappen als ras en geslacht.

Meta-ethisch cognitivisme daarentegen komt tot normatieve waarheden vanuit een puur formeel praktisch gezichtspunt. Met andere woorden uitgaande van een positie waarin geen enkele substantieve waarde wordt voorondersteld maar louter het proces van waarderen expliciet wordt gemaakt worden normatieve waarheden afgeleid.

Korsgaard maakt zelf, in een wat andere terminologie, precies hetzelfde onderscheid. Zij onderkent verschillende graden van diepgang van constructivisme (Korsgaard 2008, 324). Verwijzend naar Rawls' rechtvaardigheidstheorie stelt ze vast: in de originele positie laat hij mensen kiezen wat het beste voor hen is. Die notie van 'het beste' is een morele, of normatieve notie. Rawls construeert dus een conceptie van rechtvaardigheid, maar niet een conceptie van het goede.

Hoe diep kun je gaan met constructivisme, vraagt Korsgaard zich af. Kunnen ook onze meest basale redenen zelf worden geconstrueerd? En haar antwoord is: ja, mijn interpretatie van Kant is daar een voorbeeld van.

Haar constructivistische positie wijkt dus wezenlijk af van die van andere neo-Kantianen als Rawls of Scanlon. Op basis van niet meer dan een constructie van het rationele proces van zelfreflectie, gezien als noodzakelijk startpunt van handelen, komt zij tot een conceptie van het goede en het juiste.

Substantialiteit

Het derde thema waar ik nader bij stil sta gaat over de vraag: wat is voor Korsgaard moreel juist? Welke maatstaven, welke normen biedt haar model om in een concrete situatie te kunnen beoordelen of sprake is van moreel juist handelen?

Op het eerste gezicht komt haar visie sterk overeen met die van Kant. Op verschillende plaatsen refereert ze naar moreel handelen in termen die geheel aansluiten op Kant's CI, bijvoorbeeld: 'Wanneer een actor bepaalt of ze een maxime als een universele wet kan willen ...' (Korsgaard 2008, 324).

Haar eigen definitie van de morele wet luidt (met cursivering van mij): 'handel alleen volgens maximes waarover alle rationele wezens het eens zouden kunnen worden om alleen op grond daarvan te handelen, in een werkbaar *coöperatief* systeem' (SoN 99). Deze wijkt dus af van

Kant's CI. Korsgaard geeft bij die afwijking geen verdere verklaring of uitleg. In mijn ogen verwijst dit verschil naar een belangrijk onderscheid tussen haar en Kant.

Door in haar omschrijving niet te spreken over een *universele* wet maar over wetten (maximes) *waarover je het eens kunt worden* maakt ze helder waar in haar ogen de bron van moraliteit ligt: namelijk in de *relatie tussen mensen*. In haar paper *The Reasons We Can Share* (Korsgaard 1996c) maakt ze dit duidelijk. Bij moraliteit, zegt ze, gaat het niet primair om situaties 'in which I do something to you or you do something to me', zoals moraalfilosofen aldoor hebben aangenomen; het gaat om situaties waarin mensen samen iets doen (Korsgaard 1996c, 275). Waarden zijn *intersubjectief*. Ze maken onderdeel uit van de structuur van persoonlijke relaties (Korsgaard 1996c, 276).

Ook Korsgaard's fundering van de morele wet, die zich richt op de deelbaarheid van redenen, weerspiegelt deze focus op de relatie tussen mensen. En elders zegt ze: niet 'het Goede' maar de mens is de bron van normatieve claims (Korsgaard 1996b, 301).

Een andere verwoording die ze ook vaak kiest, sprekend over moraliteit, is de waarde van onze menselijkheid. Daarbij verwijst ze naar Kant's Formule van Menselijkheid: behandel de menselijkheid als een doel in zichzelf, niet alleen als een middel. Menselijkheid is de ongeconditioneerde doelstelling, en daarmee de conditie voor de goedheid van alle andere doelstellingen (Korsgaard 1996b, 125). Die waarde van menselijkheid ligt voor haar verankerd in onze relaties met de ander.

Wat Korsgaard doet met haar gerichtheid op relaties is de bron van moraliteit een ander plaats geven. Niet universaliseerbare wetten *boven* ons vormen onze bron voor morele claims. Nee, die bron vinden we veel dichterbij, in de relatie tussen de ander en mij. En bij die relatie gaat het niet alleen om respect en erkenning van de waardigheid van de ander, het gaat ook om genegenheid.

En daarmee krijgt ook de maatstaf voor het moreel beoordelen van iemands gedrag een andere inkleuring. Ook bij die beoordeling spelen naast respect voor de waardigheid ook het betrokken zijn op de ander, het rekening houden met de redenen van de ander, een belangrijke rol.

Met deze positie doet ze geen afstand van Kant. Wel geeft ze een sterke eigen inkleuring aan de morele wet, een inkleuring die afwijkt van de interpretatie van anderen, zoals Darwall.

4. Kritiek

In de vorige twee paragrafen heb ik een overzicht gegeven van enkele hoofdlijnen in Korsgaard's theorie over handelen en moraliteit. In deze paragraaf wil ik me richten op drie in mijn ogen meer controversiële onderdelen binnen haar theorie. Ik sta achtereenvolgens bij de volgende drie vragen stil. (1) Is er in Korsgaard's model van rationeel handelen plaats voor immorele redenen? (2) In hoeverre is het, wanneer morele redenen conflicteren met andere redenen, *mogelijk* om altijd te kiezen voor moraliteit? (3) Hoe houdbaar is Korsgaard's claim omtrent de deelbaarheid van redenen, als fundament voor de morele wet?

4.1 Rationaliteit en immoraliteit

Op welke wijze past immoreel gedrag binnen Korsgaard's theorie? Wanneer iemand iets slechts doet, heeft hij daar dan ook redenen toe? En zo ja, zijn dat dan dus immorele redenen? Of laat Korsgaard's model van rationeel handelen immorele redenen niet toe. Zijn redenen naar hun aard niet-immoreel, en dient immoreel gedrag te worden verklaard als 'on-redelijk' gedrag, als irrationeel gedrag in de zin dat dit gedrag niet voortvloeit uit zelfreflectie?

Een argumentatie om deze tweede optie (immoreel gedrag is irrationeel gedrag, immorele redenen bestaan niet) te verdedigen langs de logica van haar eigen model zou als volgt kunnen lopen.

Een actor ontleent zijn redenen aan zijn verschillende praktische identiteiten. Daaruit kunnen in een concrete situatie meerdere, conflicterende redenen uit voortvloeien. Als één van die redenen immoreel is dan zal deze conflicteren met een (morele) reden die uit zijn identiteit als mens voortvloeit, en die vraagt om respect voor de menselijkheid van de ander. En zijn plicht voor het goede, voor respect voor de ander, is *dieper* dan zijn immorele redenen (SoN 258). Dat zou betekenen dat iemands 'all things considered' reden, zijn conclusie na afweging van verschillende redenen, nooit immoreel is.

De vraag is hoe houdbaar deze interpretatie is. In de commentaren die onderdeel uitmaken van *Sources of Normativity* confronteert Cohen haar met een Maffioso die niet bewogen wordt door morele principes maar door een code van kracht en eer (SoN 183 en 184). Cohen wil met dit voorbeeld laten zien dat een proces van zelfreflectie en reflexieve acceptatie niet tot het adopteren van morele principes hoeft te leiden. In Korsgaard's reactie (en met name SoN 257 en 258) bevestigt ze dat er hier voor de Maffioso sprake kan zijn van echte verplichtingen, niet psychologisch maar normatief van aard. Ze erkent dat hij tot zijn redenen komt langs de weg van zelfreflectie, en dus (mijn conclusie) rationeel handelt. Ze stelt vast: 'ik denk niet dat alle verplichtingen moreel zijn'. Gegeven het voorbeeld dat ze bespreekt lijkt geen andere interpretatie mogelijk dan dat ze hier aangeeft dat immorele redenen c.q. verplichtingen mogelijk zijn. Ze rond haar reactie af met de vaststelling dat de Maffioso in zijn reflectie op zijn beweegredenen dieper terug had moeten gaan, tot het ongeconditioneerde. Dan was hij bij moraliteit uitgekomen. Ze concludeert: 'Zijn verplichting om een goed persoon te zijn is daarom dieper dan zijn verplichting om aan zijn code vast te houden' (SoN 258).

Een essentiële vraag hier is hoe dit 'dieper' begrepen moet worden. Ten eerste is de vraag of het begrip normatief moet worden gelezen. Als voor mij verplichting A dieper is van verplichting B, betekent dit dan dat verplichting A voor mij *zou moeten* prevaleren? Dit lijkt mij niet goed mogelijk. Daarmee zou ze haar eigen bewijsvoering op de kop zetten. Via een proces van zelfreflectie weeg ik redenen af om tot een eindoordeel over mijn handelen te komen. En volgens Korsgaard (en Kant) leidt dit proces van rationeel afwegen mij tot de conclusie dat morele verplichtingen zwaarder wegen. De normatieve autoriteit van moraliteit vloeit voort uit dit proces van zelfreflectie. Die claim zou worden ondergraven wanneer, haast via een achterdeur, bepaalde verplichtingen (morele verplichtingen) als dieper gelden dan andere in de zin dat deze zwaarder *horen* te wegen. Dan geraken we terug bij af: op grond waarvan wordt de 'mate van diepte' van een verplichting vastgesteld?

Dus 'diep' verwijst niet naar een norm, het verwijst naar een feitelijke eigenschap van een reden of een verplichting. Dan blijft nog de vraag overeind: wat betekent het dat de ene reden of verplichting dieper is dan de ander? De gekozen term 'diep' suggereert een andere, of in ieder geval ruimere betekenis dan 'sterk'. 'Reden A is voor mij dieper dan reden B' lijkt niet zonder meer equivalent aan 'Reden A is voor mij sterker dan reden B'. 'Diep' verwijst in mijn lezing naar de mate waarin de reden of verplichting samenhangt met mijn identiteit, met wie ik ben. Redenen vloeien voort uit één van mijn praktische identiteiten en de daaraan verbonden principes waarmee ik me vereenzelvig. Een reden is dan *dieper* naarmate het niet omarmen van die reden (omdat ik een andere conflicterende reden laat prevaleren) mij toch een grotere verloochening brengt van wie ik ben, van mijn identiteit. Als deze interpretatie klopt dan mag je aannemen dat diepere verplichtingen ook sterker zijn: een mens zal zo lang mogelijk vermijden om zijn eigen identiteit te ondermijnen, en dus redenen sterker laten wegen naarmate ze dieper zijn.

Korsgaard zegt, wanneer we haar uitspraak over de Maffioso algemener formuleren, dat mijn morele verplichtingen altijd dieper zijn dan mijn andere verplichtingen, voortvloeiend uit één van mijn contingente identiteiten. Als we mijn interpretatie zoals net weergegeven volgen dan betekent dit dat ik me meer identificeer met mijn algemene, morele identiteit als mens, als rationeel wezen, dan met mijn andere rollen, als vader, Nederlands burger, vriend.

De vraag is waarom dit altijd het geval zou moeten zijn. Dit blijkt niet uit haar model, en de Maffioso-case lijkt juist een voorbeeld dat de tegenoverstelde situatie illustreert. In die case blijken de verplichtingen voortvloeiend uit de code van kracht en eer voor de Maffioso dieper dan zijn morele verplichtingen – dat vormde de kern van de case. Je zou kunnen zeggen: het opgeven van de morele wet betekent voor de Maffioso minder zelfverloochening, tast minder zijn eigen identiteit aan, dan het opgeven van zijn code van kracht en eer. In die zin lijkt haar conclusie ('morele verplichtingen zijn dieper') dus feitelijk onjuist. Het lijkt er meer op dat Korsgaard eigenlijk wilde zeggen dat de morele verplichtingen dieper *zouden moeten zijn*. Hij heeft eigenlijk een *plicht* om redenen voortvloeiend uit respect voor de menselijkheid, dus morele redenen, zwaarder te laten wegen dan andere redenen. Maar dat kan zij niet zeggen want dan ontstaat weer hetzelfde probleem met de opbouw van haar argumentatie. Waar ligt dan de bron, de rechtvaardiging van die plicht om morele afwegingen te laten prevaleren? In Korsgaard's model gaat de plicht niet vooraf aan zelfreflectie, hij is er het resultaat van.

Korsgaard's reactie op de case waarmee Cohen haar uitdaagt is dus niet dat ze het handelen van de Maffioso irrationeel verklaart. Integendeel, ze bevestigt zijn rationaliteit: 'hij is een mens, die via reflectie tot zijn redenen komt' (SoN 257). De Maffioso handelt uit redenen, en gegeven de beschreven case: uit immorele redenen.

Mijn interpretatie uit deze reactie is dat Korsgaard immorele redenen toelaat in haar model, in de zin dat rationeel handelen kan samenvallen met handelen uit immorele redenen. En deze conclusie leidt tot een eerste knelpunt. Het verzwakt haar centrale claim dat onze zelfreflexieve aard ons de weg wijst naar onze morele plichten. Blijkbaar is zelfreflectie, het beoordelen van onze neigingen en impulsen, en het vervolgens daaruit kiezen op basis van ons zelf gestelde wetten of principes, niet voldoende.

4.2 Vrijheid van identificatie

Binnen Korsgaard's handelingsmodel zou je twee bronnen van handelen kunnen onderscheiden. De eerste wordt gevormd door onze contingente praktische identiteiten, de vele rollen die we hebben, die we 'zijn', ontleend aan ons praktisch leven. Die rollen dragen elk principes in zich, waarden, die in een concrete situatie mede leidend zijn voor de afwegingen die we maken, en die tot verschillende redenen kunnen leiden. Die redenen, zo hebben we hiervoor gezien, hoeven niet moreel neutraal te zijn, ze kunnen ook immoreel van aard zijn.

De tweede bron is onze ongeconditioneerde praktische identiteit als mens, die ons tot de morele wet brengt, respect voor de menselijkheid van de ander.

Die tweede bron, moraliteit, heeft waar nodig een corrigerende werking op de eerste bron. Wanneer één van mijn praktische rollen mij immorele redenen tot handelen geeft, wijst reflectie vanuit mijn rol als mens me op de waarde van de menselijkheid van de ander, en brengt me ertoe om me ('all things considered') niet door die immorele afwegingen te laten leiden. Het eerste net behandelde knelpunt, 'rationaliteit en immoraliteit', ging over de kracht van dit corrigerend mechanisme, over de vraag in hoeverre Korsgaard's model het succes hiervan verzekert.

Een andere vraag richt zich ook op situaties waarin morele redenen botsen met redenen voortvloeiend uit een contingente identiteit, en luidt: in hoeverre is het *mogelijk* om in geval van conflicterende redenen altijd te kiezen voor moraliteit, en daarmee dus een deel van je contingente identiteit te ontkennen.

Onze waarden, zegt Korsgaard, zijn nauw verbonden met de verschillende concepties van onze praktische identiteit. Die waarden zijn voor een belangrijk deel geen product van een bewuste keuze, van reflectie, maar het directe gevolg van de gemeenschappen waarvan we deel uitmaken: de cultuur waarin we worden geboren, het geloof van onze ouders, en een veelheid van andere contingente omstandigheden.

Toch, zegt Korsgaard, is ook bij deze waarden, inherent verbonden aan onze verschillende identiteits-concepties, sprake van een vrije keuze. 'Omdat deze concepties contingent zijn kun je ze van je afschudden' (SoN 120). En even later: 'Rationele reflectie kan je ertoe brengen om een bepaalde manier van denken van je praktische identiteit als dom of kleinzielig te beschouwen'. En in een voetnoot voegt ze toe: of immoreel.

Korsgaard claimt hiermee dus dat we op een bepaalde manier vrij zijn in het bepalen van onze identiteit. Weliswaar is die vrijheid niet absoluut, omdat we geen bewuste keuzes maken over de waarden die we meekrijgen en over de omgeving waarin we opgroeien en leven. Maar we kunnen er afstand van nemen wanneer we daar aanleiding toe zien, en in die zin onszelf beschermen tegen aspecten van onze identiteit, tegen waarden of principes, die we niet willen. Hoe verdedigbaar is die aanname? Laten we kijken naar twee argumenten, van Charles Taylor en Harry Frankfurt, die betrekking hebben op deze vrijheid van identiteitsbepaling. Beide denkers delen met Korsgaard een paar wezenlijke opvattingen. Ze wijzen het consequentialisme af en verbinden, net als Korsgaard, moreel oordelen met een proces van zelfreflectie.

Voor Taylor, uitgesproken communitarist, is de gemeenschap waarin we leven bepalend voor onze identiteit. Zijn argument is als volgt samen te vatten. In zijn model is een tweetal kenmerken bepalend voor wat het betekent een persoon te zijn. De eerste is dat dingen er voor een persoon toe doen. Dat betekent dat we een persoon doelstellingen, wensen, motieven toeschrijven. Die doelstellingen zijn origineel, authentiek in de zin dat de oorsprong van de wensen van een mens liggen binnen de mens zelf. Dit eerste kenmerk onderscheidt mensen en dieren van dingen: alleen mensen en dieren hebben authentieke wensen. Het tweede kenmerk onderscheidt mensen van (andere) dieren en gaat over het vermogen van de mens om zijn wensen te evalueren, oftewel, in termen van Korsgaard: zijn zelfreflexief vermogen. Taylor onderscheidt bij dit evalueren van wensen twee vormen: sterke en zwakke evaluaties, waarbij ons zelfreflexief vermogen voortvloeit uit sterke evaluaties. Een *sterke* evaluatie betreft de keuze tussen wensen op basis van de kwaliteit van die wensen. Wensen worden geclassificeerd in categorieën als hoger of lager, deugdelijk of immoreel, meer of minder zingevend. Een achterliggende vraag bij een sterke evaluatie is: wat voor soort persoon wil ik zijn. Zwakke evaluaties ontberen dit kwalitatieve element, en richten zich op de uitkomsten van de alternatieven, bijvoorbeeld bij de keuze tussen twee soorten vakanties. Mensen zijn voor Taylor per definitie 'strong evaluators'. Als we onze oriëntatie op wat voor ons waardevol is, onze morele raamwerken, kwijt zijn, weten we niet wie we zijn (Taylor 1989, pag. 29). We definiëren onszelf, onze identiteit, aan de hand van onze raamwerken, in die zin zijn we 'zelfinterpreterende dieren'.

Tot dusver is Taylor's argumentatie redelijk in lijn met het denken van Korsgaard. Belangrijkere verschillen ontstaan bij het vervolg. Om onze identiteit te kennen, om te weten wie we zijn, is een verwijzing naar en een articulatie van de morele raamwerken waarop we ons oriënteren niet toereikend. De mens wordt mede gedefinieerd door de gemeenschap waarvan hij deel uitmaakt. We definiëren onszelf in een taal en dus binnen de context van een taalgemeenschap. 'Een zelf kan nooit worden beschreven zonder referentie naar degenen die het omgeven' (Taylor 1989, 35). Immers, alleen langs de weg van taal vinden we toegang tot onze morele raamwerken. Die talige ontsluiting van raamwerken vindt plaats binnen de context van de gemeenschap waarin we leven. Zonder gemeenschap, zonder gesprekspartners is daarom een oriëntatie op morele raamwerken niet denkbaar. De notie om onszelf te definiëren los van enig sociaal netwerk leidt slechts tot illusies (Taylor 1989, 39).

Stel dat we Taylor hierin volgen, dan is daarmee Korsgaard's claim omtrent de vrijheid van identificatie nog niet weerlegd. Korsgaard ontkent niet het belang van de gemeenschap voor de identiteit. Korsgaard claimt dat we ons van waarden binnen onze identiteit kunnen ontdoen, wanneer we die niet meer willen.

In zijn paper *What is Human Agency* zoomt Taylor daar nader op in (Taylor 1985, 29-44).

We evalueren onze eerste-orde wensen op basis van onze waarden of principes, of in Taylor's termen op basis van onze morele raamwerken. Het probleem dat hierbij naar boven komt is: waaraan ontleen we onze morele raamwerken? Een invloedrijke stroming, zegt Taylor, claimt dat we onze principes zelf kiezen (Taylor 1985, 29). Het probleem met die positie is dat niet helder is *op grond waarvan* die keuzes zouden moeten worden gemaakt. Onze principes, onze waarden vormen de maatstaf voor ons handelen, ze vormen de norm om onze wensen te evalueren. Maar er is geen meta-maatstaf voorhanden om die principes zelf te beoordelen. Dat betekent dat zo'n keuze een wat Taylor noemt *radicale keuze* zou zijn, een keuze die niet gegrond is in redenen.

Toch brengt dit hem niet tot de conclusie dat we geen enkele verantwoordelijkheid hebben voor onze morele raamwerken, en voor onze eigen identiteit. Morele raamwerken zijn articulaties van onze doeleinden. Ze zijn gevormd in taal en kunnen minder goed of inconsistent geformuleerd zijn. Nieuwe inzichten kunnen leiden tot het veranderen van mijn morele raamwerken, en daarmee tot een radicale her-evaluatie (Taylor 1985, 39 en 40). Wel blijft het 'ongewoon moeilijk om te reflecteren op onze fundamentele evaluaties' (Taylor 1995, 42). Taylor acht dus het aanpassen van onze morele raamwerken niet onmogelijk, wel ongewoon moeilijk.

Ik geloof dat Taylor gelijk heeft, en dat onze vrijheid om onszelf de wet te stellen minder groot is dan Korsgaard aanneemt.

Aansluitend wil ik laten zien welk argument Frankfurt hier inbrengt, ontleend aan zijn paper *Rationality and the Unthinkable* (Frankfurt 1998, 177-190). Ook voor Frankfurt zijn wij mensen zelfreflexieve wezens, die worden gedefinieerd door onze wil. We zijn daarbij niet vrij in wat we willen. Althans, er zijn grenzen aan die vrijheid om zelf te kiezen wat we willen, als we onze autonomie niet willen verliezen. Wat geen grenzen heeft, heeft geen vorm. Net zo heeft een mens geen identiteit, tenzij hij door zijn wil (en daarmee: door wie hij is) wordt begrensd in zijn handelen (Frankfurt 1998, ix). Frankfurt richt zich in bovengenoemd paper op deze grenzen, en spreekt bij overschrijding van die grenzen over *het ondenkbare*. Het ondenkbare is iets anders dan het onmogelijke. Mensen kunnen niet in staat zijn om bepaalde dingen te doen bijvoorbeeld omdat de omstandigheden dit onmogelijk maken of omdat het hen ontbreekt aan de vereiste kracht of competenties. Maar soms kan iemand een bepaalde handeling niet uitvoeren alhoewel hij in een goede positie is om het doen, en mogelijk vindt dat hij de handeling zou moeten uitvoeren, of dat wenselijk vindt. En toch kan blijken dat die persoon niet in staat is om de handeling uit te voeren, omdat hij dit *niet kan willen*. Deze situatie duidt Frankfurt aan met het ondenkbare. Het voorbeeld dat Frankfurt geeft betreft een officier die nucleaire wapens moet lanceren, en op het beslissende moment daar niet toe in staat is, hij kon zichzelf er niet toe brengen om te handelen.

Frankfurt wil aantonen dat iemand zijn wil niet altijd naar believen kan bijsturen. Voor het denkende ik, dat meent het handelen in vrijheid van zelfreflectie te kunnen sturen, blijken soms onverwachte grenzen op te doemen. Blijkbaar kun je jezelf, je eigen identiteit, niet veranderen door een pure wilsvaad.

Frankfurt's punt over het ondenkbare sluit aan bij Taylor's punt over het radicale karakter van een keuze omtrent onze principes. Ze belichten, ieder vanuit een eigen invalshoek, het problematische karakter van het idee dat we zelf volledig grip hebben op onze eigen identiteit.

Laat ik over dit punt ter illustratie tot slot twee fictieve situaties schetsen. Het eerste betreft een Amerikaans officier, opgegroeid als zoon van een militair. Het dienen van zijn vaderland is hem met de paplepel ingegoten. Hij is gelegerd op de marinebasis Guantánamo Bay, en actief betrokken geweest bij het martelen van al dan niet vermeende terroristen, met als doel mogelijke nieuwe aanslagen te weten te komen en te verijdelen. Hij is overtuigd dat hij daarmee goed heeft gehandeld: hij dient zijn vaderland en beschermt het tegen mogelijke nieuwe terroristische aanslagen. Taylor's punt is: op grond van welk criterium zou deze officier moeten besluiten dat zijn eigen principes inferieur zijn aan andere en ze daarop opzij moeten schuiven?

Het tweede voorbeeld gaat over een in Nederland wonende Pakistaanse vader, wiens ongehuwde dochter zwanger is geworden. De eer van de familie is door haar geschonden, ook in de ogen van de Pakistaanse gemeenschap binnen de stad, die inmiddels weten van de zwangerschap. De vader besluit, na overleg met de familie, tot eerwraak: de dochter moet sterven. Hij heeft geen keuze, een andere handelswijze (hoe graag hij ook zou willen, het is zijn eigen dochter) is voor hem ondenkbaar. Frankfurt's punt is: deze vader zou waarschijnlijk niet iets anders *kunnen* willen.

Vanuit Korsgaard's model zou de reactie op deze situatie kunnen zijn: deze vader *hoort* anders te handelen, en ook anders te willen. De morele redenen om respect te hebben voor het leven van zijn dochter zijn zo diep, dat deze hem altijd zouden moeten weerhouden van geweld tegen zijn eigen dochter. Een antwoord daarop in de net geschetste lijn van denken van Frankfurt zou niet zijn om de morele plicht van de vader die Korsgaard schets te ontkennen. Een antwoord zou wellicht zijn: de genoemde morele redenen zijn inderdaad heel diep. Maar de redenen die de vader ontleent aan de noodzaak van eerwraak zijn zo sterk verbonden met zijn identiteit van gezinshoofd en lid van zijn lokale Pakistaanse gemeenschap, ze zijn zo bepalend voor wie hij is, dat het opzij schuiven daarvan zou betekenen dat hij het hart van zijn identiteit zou opofferen. Die redenen zijn voor hem daarom nog dieper dan de morele redenen.

Algemener gesteld: redenen die direct worden ingegeven om je eigen contingente identiteit niet kwijt te raken kunnen dieper zijn dan morele redenen. Dit is binnen Korsgaard's model problematisch in die zin dat daarmee de corrigerende werking van morele reflectie in bepaalde situaties wordt gefrustreerd.

De eerste kritiek, besproken in paragraaf 4.1, richtte zich op de vraag in hoeverre Korsgaard's model ruimte laat voor rationeel en tegelijk immoreel handelen. In deze paragraaf heb ik onderzocht in hoeverre de mens *vrij* is in het zelf kiezen, definiëren van zijn identiteit – en daarmee ook van zijn morele identiteit.

Beide argumenten lijken te convergeren tot de volgende conclusie. Korsgaard's model van handelen, gebaseerd op zelf-conceptie vanuit contingente praktische identiteiten, schetst het rationeel handelen van de mens zeer concreet, en brengt het heel dicht bij onze sociale werkelijkheid. Het wint daarmee aan kracht als verklarend model voor ons handelen.

Tegelijkertijd maakt deze benadering het haar moeilijker om vanuit zelfreflectie morele normativiteit af te leiden, wat het centrale doel was van haar project *Sources of Normativity*. De vraag op grond waarvan morele redenen voor een rationeel denkend wezen altijd zouden moeten prevaleren (en volgens de tweede kritiek: zouden *kunnen* prevaleren) boven andere redenen blijkt vanuit haar model moeilijk op overtuigende wijze te beantwoorden.

4.3 Fundering van de morele wet

De eerste twee genoemde knelpunten vloeien voort uit de spanning tussen morele redenen en andere redenen. Het derde knelpunt is van een andere orde en richt zich op Korsgaard's fundering van de morele wet.

Hiervoor is besproken dat Korsgaard het transcendentaal argument ontoereikend vindt als onderbouwing voor de morele wet. Dit argument toont wel aan dat mijn menselijkheid voor mij van belang is, en jouw menselijkheid voor jou. Het laat echter niet zien waarom jouw menselijkheid normatief zou zijn voor mij.

Om dit gebrek op te lossen wil ze aantonen dat onze redenen inherent een publiek karakter hebben, en dus dat we onze redenen met elkaar delen. Dat doet ze langs twee wegen, die ik heb aangeduid met het naturalistisch argument en het taalargument.

Van het naturalistisch argument heb ik aangegeven dat ik niet zeker weet of haar doelstelling was om dit als funderingsargument te gebruiken. Als dat wel het geval is denk ik dat het niet

succesvol is. Dit argument zegt dat de mens naar zijn aard sociaal is en als gevolg daarvan zich bekommert om de redenen van de ander. Laten we aannemen dat er voldoende onderbouwing is voor haar stelling dat wij mensen sociale dieren zijn. Wat doet dit argument? Het verklaart ons gedrag. Het legt een verband tussen onze natuur en ons gedrag. Maar dat is iets anders dan het *funderen* van morele claims, Korsgaard's streven. Het argument ondergraaft haar eigen handelingsmodel waarin ons handelen als rationeel wezen wordt gekenmerkt door onze zelfreflectie, onze wilsvrijheid, waarbij we onze verlangens beoordelen op basis van maximes. Dit argument zegt dat onze instincten en verlangens zelf moreel zijn, omdat we, als sociale dieren, niet anders kunnen. Daarmee is moreel gedrag wel verklaard, maar wel ten koste van het fundament van haar theorie, namelijk onze autonomie als zelfreflexief wezen. Het argument geeft dus geen onderbouwing van hoe ik me *zou moeten* gedragen. Het geeft geen verklaring voor de autoriteit, de bron van normativiteit. Het verklaart slechts waarom ik me gedraag zoals ik me gedraag.

Resteert Korsgaard's laatste argument, het taalargument.

Ze vergelijkt daarin de normativiteit van redenen met de normativiteit van betekenissen. Omdat *betekenis* een normatief begrip is moet het ook relationeel zijn, zo vat ze Wittgenstein's argument samen. En om precies dezelfde reden vormt ook *reden* een relationele notie die betrekking heeft op een relatie tussen entiteiten.

Het taalargument zegt dat voor een reden altijd twee mensen, in een onderlinge relatie, nodig zijn, net als bij taal. Maar, anders dan in het geval van taal kunnen bij een reden die twee entiteiten het denkende en het handelende zelf zijn. Met name wanneer iemand wordt gedreven door immorele redenen wordt dit evident. Stel een belangrijk doel in mijn leven is om rijk te worden. En tot een zekere grens kan dat gaan ten koste van anderen. Mijn streven naar rijkdom is een principe dat leidend is bij veel van mijn handelingen. Hoe publiek is deze reden? Stel ik deel deze reden met anderen. Er is een redelijke kans dat de meesten van hen behoedzaam zullen worden om met mij samen activiteiten te ondernemen, teneinde te voorkomen dat zij zelf nadeel ondervinden van mijn project. In ieder geval zullen zij niet gauw mijn ambitie ondersteunen. Dit beseffende is het voor mij verstandiger om die reden voor me te houden, en deze niet met anderen te delen. Mijn reden is dus niet inherent publiek en niet normatief voor anderen – een voorbeeld dat conflicteert met Korsgaard's claim.

Als mijn hiervoor geschetste conclusie dat binnen Korsgaard's theorie redenen immoreel kunnen zijn juist is dan lijkt Korsgaard's taalargument niet op alle redenen toepasbaar te zijn, namelijk niet op immorele. Dat was ook ongetwijfeld niet haar bedoeling. Je zou dit probleem kunnen ondervangen door Korsgaard's claim te interpreteren zoals ze het vermoedelijk heeft bedoeld: namelijk immorele redenen uitsluitend. Dan ontstaat echter een probleem bij de manier waarop ze het begrip *moreel* inhoud geeft.

Korsgaard spreekt van moraliteit wanneer we niet alleen onze eigen menselijkheid waarderen, maar ook die van de ander. De onconditionele waarde van de menselijkheid van de ander vloeit voor Korsgaard voort uit het idee dat redenen naar hun aard deelbaar zijn. Jouw redenen zijn in beginsel ook voor mij redengevend. Dus *respect voor de menselijkheid van de ander* betekent voor Korsgaard *respect voor de redenen van de ander*. Korter gezegd: moraliteit betekent respect voor de redenen van de ander.

Als we hier nu de beperkende interpretatie toepassen over de deelbaarheid van redenen dan zouden we dus moeten lezen 'moraliteit betekent respect voor de niet-immorele redenen van de ander'. Maar daarmee ontstaat een circulaire definitie van moraliteit. En dat brengt ons nog verder van huis.

Kijkend naar het thema *deelbaarheid van redenen* verdienen ook de *implicaties* van haar argumentatie aandacht.

Stel dat we het taalargument en het publieke karakter van redenen accepteren. We nemen dan aan dat ik me door de redenen van een ander op een vergelijkbare wijze laat leiden als door mijn eigen verlangens. Net als bij verlangens ga ik niet zonder meer op de redenen van de ander in, maar evalueer ze. En als er geen redenen zijn om ze af te wijzen accepteer ik ze als redenen voor mijzelf – net zoals ik dat doe bij mijn verlangens. Als dat waar zou zijn dan lijkt moreel gedrag geen verdere rechtvaardiging nodig te hebben. Korsgaard zegt dit zelf ook met zoveel woorden: ‘Dit [namelijk: het publieke karakter van redenen] zou er haast, maar niet helemaal, op neer komen dat moraliteit geen rechtvaardiging nodig heeft’ (SoN 135). De toevoeging ‘maar niet helemaal’ wordt niet verder toegelicht. Deze consequentie van het taalargument lijkt moeilijk denkbaar en lastig te rijmen met de dagelijkse sociale werkelijkheid waarin we leven.

Al met al lijkt Korsgaard’s onderbouwing van de morele wet met behulp van de deelbaarheid van redenen problematisch. Die deelbaarheid is moeilijk voorstelbaar voor immorele redenen, en het uitsluiten van immorele redenen uit haar deelbaarheids-claim zou tot problemen leiden met haar definitie van moraliteit. Daarnaast zou deelbaarheid van redenen impliceren dat moraliteit geen rechtvaardiging nodig heeft – een implicatie die Korsgaard’s eigen project overbodig zou maken.

5. Conclusies

Ik heb in dit paper me verdiept in Korsgaard’s moraalfilosofisch denken, waarbij ik me met name heb gericht op de vraag: hoe overtuigend heeft zij de normativiteit aangetoond van morele redenen, waaraan ontleen deze redenen hun gezag? Dit is de vraag naar de fundering van haar ethisch model.

Ik heb me hierbij niet alleen gericht op kwetsbaarheden in haar model, op componenten in haar argumentatie die in mijn ogen minder overtuigend zijn, of ter discussie zouden kunnen worden gesteld. Ik heb ook gekeken naar vernieuwingen en sterke punten, die in mijn ogen bijdragen aan een verdieping of verrijking van ons morele gedachtengoed.

Mijn analyse mondt uit in drie punten van kritiek en drie nieuwe inzichten. Die vat ik in deze concluderende paragraaf samen.

Kijkend naar de kwetsbaarheden heb ik me in de eerste plaats gericht op de vraag in hoeverre Korsgaard binnen een context van rationeel denken ruimte laat voor immoraliteit. Zijn redenen naar hun aard niet-immoreel, of kan een rationeel handelend mens ook gedreven worden door immorele redenen? Uit de wijze waarop ze de Maffioso-case behandelt die Cohen haar presenteert concludeer ik dat binnen haar model handelen uit immorele redenen en rationeel handelen kunnen samenvallen. Crux hierbij is dat ze niet helder maakt waarom morele verplichtingen altijd zouden moeten prevaleren boven andere verplichtingen. Haar argument dat morele verplichtingen ‘dieper’ zijn (of horen te zijn) is in mijn ogen niet overtuigend.

Mijn tweede kritiekpunt betreft Korsgaard’s aanname dat wij in belangrijke mate zelf onze waarden, doeleinden, maximes kunnen kiezen, en daarmee altijd in staat zijn om in een situatie van conflicterende redenen te kiezen voor moraliteit. Ik heb aan de hand van Taylor’s noties omtrent sterke evaluaties en radicale keuzes, en van Frankfurt’s denken over de grenzen van het ondenkbare, getracht om helder te maken dat zo’n keuze over je eigen waarden gelijk staat aan het maken van keuzes over wie je bent. Verplichtingen die voortvloeien uit een contingente identiteit kunnen zo diep zijn dat verloochening daarvan tot ernstige zelfverloochening zou leiden, en daardoor zwaarder wegen dan morele verplichtingen.

Deze eerste twee kritiekpunten raken het hart van Korsgaard’s model. Ze stellen de normatieve kracht van morele redenen, als altijd prevalerend boven andere redenen, ter discussie.

Mijn derde punt van kritiek betreft Korsgaard’s onderbouwing van de morele wet, deze is in mijn ogen problematisch. Ze geeft hiervoor drie varianten. Het eerste, het transcendentale

argument van Kant, wijst ze zelf af. Het tweede argument, wat ik heb genoemd het naturalistisch argument, zegt dat we elkaars redenen als normatief beschouwen omdat we naar onze aard sociale dieren zijn. Dit argument ondergraaft Korsgaard's eigen handelingsmodel waarin ons handelen als rationeel wezen wordt gekenmerkt door onze zelfreflectie, onze wilsvrijheid, waarbij we onze verlangens beoordelen op basis van maximes. Het argument verklaart wel ons moreel gedrag, maar het geeft geen onderbouwing voor hoe we ons *zouden moeten gedragen*, het verklaart niet de autoriteit van morele claims.

Het derde argument, het taalargument, claimt dat al onze redenen inherent publiek zijn en daarmee voor iedereen normatief. Volgens Wittgenstein heeft taal inherent een publiek karakter. Betekenis is een relationeel begrip, het is een normatief concept. Net zo, zegt Korsgaard, zijn, in het praktische domein, redenen relationeel en normatief.

Dit argument is in mijn ogen discutabel, omdat het publieke karakter van immorele redenen moeilijk denkbaar is. Daarnaast zou, indien deze claim waar is, moraliteit een vanzelfsprekendheid worden, moreel gedrag heeft dan geen verdere rechtvaardiging nodig.

Naast deze kwetsbaarheden zijn een drietal punten opvallend die in mijn ogen een andere, nieuwe interpretatie bieden van Kant's denken, en die ons nieuwe inzichten bieden omtrent de bronnen en de inhoud van moraliteit.

Het eerste nieuwe inzicht betreft Korsgaard's introductie, binnen het raamwerk van Kant, van het concept praktische identiteit. Met dit begrip biedt Korsgaard een instrument om de complexiteit van onze redenen tot handelen te verklaren, en om morele redenen, morele normativiteit en morele identiteit te plaatsen binnen de bredere context van de vele andersoortige redenen die we ontleen aan onze lokale, contingente identiteiten. Hiermee heeft ze een model geconstrueerd dat zich gemakkelijk laat toepassen op werkelijke situaties van menselijk handelen, waarin ook bijna altijd sprake van een complex van redenen die voortvloeien uit onze verschillende rollen.

Een tweede nieuwe inzicht dat zij inbrengt vloeit direct voort uit haar model van praktische identiteiten en betreft de principiële onoplosbaarheid van vele morele dilemma's. Zij brengt dit punt (voor zover ik heb kunnen nagaan) slechts één keer aan de orde (SoN 126). En die ene keer geeft ze het impliciet aan. Ze zegt dat iemand meerdere verplichtingen kan hebben die tegelijk ongeconditioneerd zijn en onderling conflicterend. Dat dit ook (of met name) morele dilemma's betreft en dat dit betekent dat deze dilemma's dus onoplosbaar zijn laat ze aan de lezer over om te concluderen. Toch acht ik het een belangrijke claim. Kijkend naar de onderbouwing vloeit de claim logisch voort uit haar model met praktische identiteiten en is daarmee overtuigend. Kijkend naar de impact zegt Korsgaard hiermee impliciet dat van een ethische theorie, en dus ook van haar theorie, niet verwacht kan worden dat deze voor alle morele vraagstukken een eenduidig, bevredigend antwoord kan bieden. Ethiek, moraalfilosofie, is door de aard van zijn aandachtsgebied beperkt in zijn vermogen om oplossingen te bieden voor de vragen waarop het een antwoord wil geven.

Een derde inzicht dat Korsgaard biedt betreft haar inhoudelijke invulling van moraliteit. Bij moraliteit gaat het voor haar om de vraag hoe mensen zich tot elkaar verhouden. Niet universaliseerbare wetten *boven* ons vormen onze bron voor morele claims. Nee, die bron vinden we veel dichterbij, in de relatie tussen de ander en mij. En bij die relatie gaat het niet alleen om respect en erkenning van de waardigheid van de ander, het gaat ook om genegenheid. Daarmee geeft ze een nieuwe, menselijker, interpretatie van Kant dan vele andere filosofen.

Literatuur

- Bagnoli, Carla, *Constructivism in Metaethics*, The Stanford Encyclopedia of Philosophy (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/constructivism-metaethics/>.
- Darwall, Stephen, *The Second-Person Standpoint. Morality, respect and accountability*, HUP, 2006
- Frankfurt, Harry, *The Importance of What We Care About*, Cambridge University Press, 1998
- Korsgaard, Christine, *Sources of Normativity*, Cambridge University Press, 1996a
- Korsgaard, *Creating the Kingdom of Ends*, Cambridge University Press, 1996b
- Korsgaard, Christine, *The Reasons We Can Share: an attack on the distinction between agent-relative and agent-neutral values*, in *Creating the Kingdom of Ends*, Cambridge University Press, 1996c
- Korsgaard, Christine, *The Constitution of Agency; Essays on Practical Reason and Moral Psychology*, Oxford University Press, 2008
- Rorty, Richard, *Is Truth a Goal or an Inquiry?*, in *The Philosophical Quarterly*, Vol. 45, 1995
- Street, Sharon, *What is Constructivism in Ethics and Metaethics?*, in *Philosophy Compass*, vol. 4, 2009
- Taylor, Charles, 'What is Human Agency', in *Philosophical Papers 1, Human Agency and Language*, Cambridge: Cambridge University Press, 1985
- Taylor, Charles, 'Sources of the Self; the Making of the Modern Identity', Cambridge: Harvard University Press, 1989