

Zelf-constitutie als grond voor integriteit van de organisatie

Jeroen Flamman



TILBURG UNIVERSITY
School of Humanities

MASTER FILOSOFIE
Ethiek van Bedrijf en Organisatie

Zelf-constitutie als grond voor integriteit van de organisatie

*Een voorstel voor het denken over integriteit van de organisatie
op basis van de zelf-constitutietheorie van Christine M. Korsgaard*

Student: Jeroen Flamman (ANR 547691)
Course: Masterthesis (799402)

Supervisor: Wim Dubbink
Second Reader: Ruud Welten

Breda, Tilburg: Juli 2016

Front image: Ian Edwards, Creation of Self (1974)

“[...] whenever you choose an action – whenever you take control of your own movements – you are constituting yourself as the author of that action, and so you are deciding who to be. Human beings therefore have a distinct form of identity, a norm-governed or practical form of identity, for which we are ourselves responsible. As a rational being, as a rational agent, you are faced with the task of making something of yourself, and you must regard yourself a success or a failure insofar as you succeed or fail at this task.”

(Korsgaard, Self-constitution. Agency, Identity and Integrity, 2009, pp. xi-xii)

Voorwoord

Met deze thesis sluit ik met trots, plezier en opluchting een intensieve periode van studeren af aan de masteropleiding Filosofie van Tilburg University. Trots omdat ik een langgekoesterde wens eindelijk heb gerealiseerd, en daartoe behoorlijk wat obstakels, pijn en moeite heb moeten overwinnen. Plezier omdat het, ondanks die inspanning, een fantastische tijd was, waarin ik van inspirerende docenten college en begeleiding mocht krijgen en mezelf heb kunnen ontwikkelen van ‘geïnteresseerde’ in de filosofie tot een gediplomeerd filosoof. En opluchting omdat ik na zoveel tijd ontzettend toe ben aan ‘ja’ zeggen op uitnodigingen en verzoeken van familie, vrienden en vooral van Piron en Juliette, in plaats van ‘nee, ik ga studeren’.

Als het auteurschap van een resultaat wordt bepaald door de mate van inspanningen die zijn verricht om dat resultaat te bereiken, dan zou onder deze thesis niet alleen mijn naam mogen staan. Mijn eigen inspanningen zouden weinig waard zijn zonder de inspanningen van mijn ouders, Ger en Ankie, mijn schoonouders, Ton en Marijke, mijn kleintje Juliette en eerst en vooral van mijn fantastische vrouw, Piron. Lieverd, dit resultaat is net zozeer van jou als van mij, simpelweg omdat het zonder jouw inspanningen er niet was geweest.

De ontvangen begeleiding van Wim Dubbink in het schrijven van de thesis heb ik ervaren als waardevol: efficiënt, gelijkwaardig en inspirerend; het was niet alleen leuk, maar het heeft me ook echt geholpen mijn denken aan te zwengelen, te structureren en door te ontwikkelen. Waarvoor dank.

De thesis is geschreven voor lezers die reeds over behoorlijk wat (moreel)filosofische kennis beschikken; de geïnteresseerde ‘ik’ die aan de vooravond van de masteropleiding stond zou er lang niet alles van begrijpen. Veel concepten en begrippen worden als bekend verondersteld en daarom niet nader uitgelegd. Het motief daartoe is louter egoïstisch: het is wat mij betreft niet zozeer de leesbaarheid, als wel de schrijfbaarheid ten goede gekomen; het had me simpelweg teveel moeite en bladzijden gekost. Desalniettemin denk ik een goed en leesbaar stuk gerealiseerd te hebben. Ik sta er in ieder geval *als een eenheid* volledig achter.

Jeroen Flamman

Breda, juni 2016

Samenvatting

Korsgaard claimt dat de functie van een actie de zelf-constitutie is van de actor, en dat het kiezen van acties daarom aan bepaalde constitutieve principes moet voldoen, namelijk de categorische en de hypothetische imperatief. De hypothetische imperatief maakt de actor doeltreffend, de categorische imperatief maakt de actor autonoom. Volgens haar wordt, zodra er een aansporing tot actie ontstaat, de ziel gesplitst in verschillende delen, en kunnen deze door het volgen van de imperatieven weer tot een eenheid worden gebracht. Dat is wat zij verstaat onder integriteit.

Op basis hiervan doe ik een voorstel voor het denken over integriteit van de organisatie, waarbij ik de organisatie als een normatief idee hanteer, dat stelt dat een organisatie pas een organisatie is als het zich als instrument inbedt in de samenleving. Met het voorstel claim ik dat een actie namens de organisatie pas een actie namens de organisatie kan zijn, als het de organisatie constitueert tot een doeltreffende en autonome organisatie. Acties dienen daarom deliberatief, conform de hypothetische en categorische imperatief, gekozen te worden, zodanig dat de organisatie als eenheid daar achter kan staan. De constitutie van de organisatie levert deze constitutieve c.q. normatieve principes.

Omdat de functie of praktische identiteit van de organisatie betrekking heeft op de wijze waarop de organisatie is ingebed in de samenleving, veronderstelt dat dat het zich als organisatie inbedden in de samenleving niet zozeer een materiële keuze is, maar een (derde) normatief principe waaraan acties zich dienen te conformeren, wil er *überhaupt* sprake zijn van een organisatie.

De organisatie is daadwerkelijk in staat de normatieve principes te integreren in een constitutie. Dat vermogen wordt gemanifesteerd met de interne besluitvormingsstructuur, de CID-Structure, die als het ware de constitutie van de organisatie representeert, en waarin een bepaalde vorm van redelijkheid is geborgd op basis waarvan voorstellen tot acties kunnen worden getoetst op hun waarde als actie-ter-realisatie-van-een-bepaald-doel. Die redelijkheid binnen de structuur krijgt vooral een plaats in enerzijds een identificatie met en begrip van de functie die organisatierepresentanten binnen de organisatie vervullen, in relatie tot de functie van de organisatie als geheel: de ingebedde functie-rationaliteit. En anderzijds in de mate waarin een zelfcorrigerende of reflectieve werking binnen de structuur is geborgd, zodanig dat de interne communicatie, de hantering van conflicten en dialogen over meningsverschillen, in feite de deliberatie gestalte geven.

Het voorstel draagt op een tweeledige manier bij aan het huidige debat over de organisatie als morele actor. Het laat allereerst zien dat het niet nodig is om over het vermogen tot intentionaliteit te beschikken wil die actor een morele actor zijn. Wat nodig is, is het vermogen om te *delibereren* over een te verrichten actie, conform principes die de actor constitueren tot doeltreffende en autonome eenheid aan wie de actie kan worden toegeschreven. En de CID-Structure representeert dat vermogen van de organisatie. Het voorstel laat ook zien dat, waar verondersteld wordt dat er actorschap nodig is om intenties en keuzes voor acties reflectief te kunnen bekrachtigen, dit nu juist wordt omgedraaid: het actorschap wordt *geconstitueerd* door het verrichten van de actie die op basis van deliberatie conform de imperatieven is verkozen. Het is kortom niet de organisatie als actor die aan bepaalde kenmerken moet voldoen wil zij autonoom in staat zijn bepaalde acties te verrichten; het is de manier van delibereren over de acties namens de organisatie die aan bepaalde kenmerken moet voldoen wil het in staat zijn de organisatie te constitueren tot een doeltreffende en autonome actor, en dus tot een morele actor.

Maar wel tot een *specifiek* soort morele actor, die zich op een essentieel punt onderscheidt van het moreel actorschap van de mens. De organisatie dient als organisatie noodzakelijkerwijs niet alleen de imperatieven, maar ook het normatieve principe van de inbedding in de samenleving te hanteren in het waarderen van te verrichten acties. Pas dan is er sprake van de integere organisatie.

Inhoudsopgave

1.	Inleiding.....	1
2.	Zelf-constitutie als grond voor moraliteit bij mensen.....	3
2.1.	Inleiding.....	3
2.2.	We ervaren een dwang om onszelf te constitueren tot integere eenheid.....	3
2.3.	De morele waarde van een actie ligt besloten in de vorm van de actie.....	4
2.4.	Een actie constitueert een actor als eenheid.....	5
2.5.	Constitutieve principes als formele principes hebben normatieve kracht.....	6
2.6.	De constitutieve principes van een actie zijn de imperatieven.....	7
2.7.	Het zelfbewustzijn creëert de delen van de ziel.....	8
2.8.	Deliberatie conform de constitutieve principes constitueert eenheid.....	9
2.9.	Actie conform de imperatieven constitueert integere eenheid.....	10
3.	Zelf-constitutie als grond voor integriteit van de organisatie.....	12
3.1.	Inleiding.....	12
3.2.	Methodologische kanttekening.....	12
3.3.	Het normatieve idee van de organisatie volgens Scott.....	13
3.4.	Van normatief idee van de organisatie naar de functie van de organisatie.....	18
3.5.	Een contingent normatief principe: streef naar inbedding in de samenleving.....	19
3.6.	Acties namens de organisatie constitueren de organisatie tot eenheid.....	20
3.7.	De constitutieve principes van een actie namens de organisatie zijn de imperatieven.....	21
3.8.	De mogelijkheid tot integriteit bij organisaties op grond van zelf-constitutie.....	22
3.9.	Eisen aan de constitutie van de integere organisatie.....	24
4.	De plaats van de integere organisatie in het debat over de organisatie als morele actor.....	25
4.1.	Inleiding.....	25
4.2.	Hoofdpijnen van het debat over de organisatie als morele actor.....	25
4.3.	De organisatie moet in staat zijn een intentie te hebben om een actie te verrichten.....	26
4.4.	De organisatie moet in staat zijn een intentionele actie autonoom te kiezen.....	29
4.5.	De integere organisatie als morele actor.....	31
5.	De organisatie als een specifiek soort morele actor.....	33
5.1.	Inleiding.....	33
5.2.	De mens als actor onderscheidt zich van het dier in het kiezen van principes.....	33
5.3.	De organisatie als actor onderscheidt zich van de mens in het inbedden in de samenleving.....	33
6.	Conclusie.....	35
	Bibliografie.....	37

1. Inleiding

Wangedrag van of bij organisaties en bedrijven is aan de orde van de dag. Denk bijvoorbeeld aan grote zaken als Volkswagen en het gemanipuleer met de software die de uitstoot van schadelijke stoffen van dieselmotoren reguleert, de NS en het Fyra-debacle, de corruptieschandalen bij de FIFA, de Rabobank en de Liborfraude, maar ook (relatief) kleine gebeurtenissen als sjoemelende autohandelaren, ernstig tekortschietende ouderenzorgcentra, winkelmedewerkers die hun klanten voorliegen, politiemensen die onnodig geweld gebruiken, etc..

De manier waarop buitenstaanders of betrokkenen reageren op het wangedrag van of bij organisaties doet lijken alsof het *morele* misstappen zijn die begaan worden. Als ergens de bliksem inslaat en veel schade aanricht dan vinden we dat vervelend, maar snappen we ook dat zo iets *als natuurverschijnsel* nou eenmaal kan gebeuren; we houden de bliksem niet moreel verantwoordelijk. Maar als ergens in een magazijn een grote voorraad vuurwerk ontploft met veel schade als gevolg, zijn we verontwaardigd, en vinden we dat dat niet had mogen gebeuren, dat een organisatie – of mensen binnen die organisatie - zich niet goed gedragen hebben.

We doen kortom een moreel appèl op organisaties, voornamelijk vanuit een behoefte om morele verantwoordelijkheid toe te kunnen schrijven. Binnen de organisatie-ethiek wordt ook veelal impliciet aangenomen dat organisaties een morele verantwoordelijkheid hebben. Dit komt deels door de behoefte om 'de organisatie' te kunnen kwalificeren als een eenheid om haar acties ethisch te kunnen analyseren. (Collier, 1998) En deels komt het door de praktische behoefte om, daar waar morele verantwoordelijkheden niet aan de organisatiedeelnemers of –representanten toegeschreven kunnen worden, ze toch ergens een plaats te kunnen geven. (Werhane, 1989)

Het is weinig controversieel om te stellen dat we morele verantwoordelijkheid aan een *individuele persoon* toe kunnen schrijven, omdat die persoon in staat is moreel te handelen. Op dezelfde manier wordt vaak verondersteld dat de organisatie *als organisatie* alleen moreel verantwoordelijk kan zijn voor haar acties, als de organisatie als organisatie in staat is moreel te handelen. Het veronderstelt dat de organisatie als *morele actor* gedacht kan worden, en in dat verband wordt veelal de vergelijking gemaakt met de noodzakelijke condities voor de individuele persoon als morele actor, en onderzocht in hoeverre die ook van toepassing zijn op de organisatie.

Maar de organisatie is niet hetzelfde als een persoon. De relatie tussen de noodzakelijke condities voor moreel actorschap bij personen en de mogelijkheid tot toeschrijving van morele verantwoordelijkheid geldt daarmee niet automatisch voor de organisatie. Het debat over moraliteit bij de organisatie geeft dan ook blijk van een worsteling met beide concepten, die soms door elkaar heen worden gebruikt, soms elkaar veronderstellen of soms aan elkaar gelijk worden gesteld.

Met mijn thesis beoog ik een bijdrage aan dat debat te leveren door, *losgekoppeld* van het denken over de toeschrijving van morele verantwoordelijkheid aan de organisatie, een antwoord te formuleren op de vraag of de organisatie als organisatie *überhaupt* als morele actor gedacht kan worden.

Daarbij baseer ik me op de zelf-constitutietheorie van Christine M. Korsgaard (Korsgaard, 2009). Zij presenteert een theorie over moraliteit die juist niet uitgaat van bepaalde noodzakelijke condities waar een individueel *persoon* aan moet voldoen wil die persoon in staat zijn moreel te handelen. Zij claimt dat de noodzakelijke condities te vinden zijn in *de manier waarop te verrichten acties worden gekozen*, door te stellen dat, als de acties op een bepaalde manier worden gekozen, daarmee juist een moreel actorschap *geconstitueerd* wordt, dat wil zeggen een integere eenheid die in staat is moreel te handelen. En volgens Korsgaard is dat niet enkel voorbehouden aan een persoon: “*this is not yet to say whether the agent [...] needs to be a separately existing entity – as I will explain, I don't think that.*” (Korsgaard, 2009, p. 18) En het is precies daarom dat ik mogelijkheden zie om ook de organisatie als morele actor te denken.

In het volgende hoofdstuk zal ik derhalve Korsgaards zelf-constitutietheorie verder uitwerken, aan de hand van haar belangrijkste claims die relevant zijn in het licht van de thesis.

Op basis van deze claims probeer ik in hoofdstuk 3 een voorstel te formuleren voor het denken van de organisatie als morele actor. Dat kan alleen door eerst te onderzoeken wat we eigenlijk onder het idee van een organisatie verstaan, waarna ik vervolgens uiteenzet in hoeverre de condities die Korsgaard stelt aan de wijze waarop acties tot stand moeten komen willen zij de morele actor constitueren, van toepassing kunnen zijn op dat normatieve idee van de organisatie.

In hoofdstuk 4 zet ik het geformuleerde voorstel voor het denken van de organisatie als morele actor af tegen het huidige debat over moreel actorschap bij de organisatie. Enerzijds om te laten zien hoe mijn voorstel zich verhoudt tot de argumenten van zowel voor- als tegenstanders van het idee van de organisatie als morele actor. Anderzijds om te laten zien op welke wijze met mijn voorstel een bijdrage aan dat debat geleverd kan worden.

Tot slot geef ik in hoofdstuk 5 kort weer wat, ondanks de vele overeenkomsten tussen de individuele persoon en de organisatie als morele actor, het wezenlijke verschil tussen beiden is. Hiermee wil ik laten zien dat de organisatie niet op dezelfde manier als een persoon als morele actor gedacht moet worden, maar als een *specifiek soort* morele actor, in de hoop dat het debat over moreel actorschap van de organisatie daarmee een nieuwe wending krijgt.

2. Zelf-constitutie als grond voor moraliteit bij mensen

2.1. Inleiding

Ik begin met een algemene weergave van de filosofie van Korsgaard, waarin ze laat zien hoe integriteit als functie van het menselijk handelen tot stand komt door je te committeren aan moraliteit (Korsgaard, 2009). Ik doe dat aan de hand van de belangrijkste claims die zij doet en die in mijn ogen relevant zijn in het licht van de thesis, en wellicht aanknopingspunten bieden om in het volgende hoofdstuk tot een voorstel te komen voor het denken over integriteit bij de organisatie.

2.2. *We ervaren een dwang om onszelf te constitueren tot integere eenheid*

Mensen zijn volgens Korsgaard gedoemd tot keuze en actie. Ook kiezen voor niet handelen is een vorm van actie. Daarmee is niet gezegd dat je niet kunt falen met je handelen, want dat kan natuurlijk wel. Je kunt in slaap vallen achter het stuur, verlamd raken door angst etc.. In dat geval mislukt je handeling, maar dat is niet iets wat je kunt *ondernemen*. Je *kies*t er niet voor om al rijdend achter het stuur in slaap te vallen. Doe je dat wel, dan is dat op zichzelf weer een actie. Dus voor zover je de baas bent over je handelingen *moet* je handelen. Je hebt geen andere keuze dan te kiezen, en te handelen naar die keuze.

Actie is dus noodzakelijk. Volgens Korsgaard is dit geen logische of causale noodzakelijkheid. Het gaat niet om *iets* dat op jou inwerkt, of je er nu weet van hebt of niet, zoals een oorzaak, maar om een noodzakelijkheid waarmee je wordt *geconfronteerd*. Het is ook niet te duiden als een rationele noodzakelijkheid, in termen van: als je dit doel wil realiseren, dan moet je die handeling verrichten. Want daar zit een als-dan constructie in, en de noodzakelijkheid van actie is onvoorwaardelijk. De noodzakelijkheid van kiezen en handelen is in plaats daarvan een *toestand (plight)* waarin we noodzakelijkerwijs verkeren: een simpel onoverkomelijk feit van het mens-zijn.

Binnen die noodzakelijkheid worden we geconfronteerd met een ander soort noodzakelijkheid. We leven onder de druk van een groot spectrum aan wetten, plichten, verwachtingen en regels, die ons vertellen wat we moeten doen. Sommige daarvan zijn ongetwijfeld betwifelbaar, imaginair of selectief, maar van anderen voelen we dat we echt gebonden zijn ze te gehoorzamen; denk aan professionele verplichtingen, seksuele trouw, alledaagse etiquette. Soms voelen we weerstand aan dergelijke eisen en wetten te voldoen, en zijn we ook in staat ze daadwerkelijk te weerstaan: we reageren bijvoorbeeld niet op een vervelende voicemail. Maar veel vaker mislukt het om weerstand te bieden, en zien we onszelf doen wat we denken *behoren* te doen, terwijl er geen duidelijk aanwijsbare kracht is die ons ertoe dwingt: we trekken erop uit in regenachtig weer om te stemmen, we bellen familieleden die we liever niet spreken, etc.. Het komt kortom vaak genoeg voor dat we specifieke acties kiezen te verrichten, die van ons vereist worden, terwijl we ze niet willen verrichten.

Waar komt die normativiteit vandaan? Wat is de grond of de autoriteit ervan? Volgens Korsgaard is de normativiteit van de verplichting, onder andere, een psychologische kracht. En omdat normativiteit een vorm van noodzakelijkheid is, noemt ze de werking ervan in onszelf – de manifestatie ervan als psychologische kracht – *necessitation*, wat vertaald zou kunnen worden als de ervaring van een *dwang*¹.

Om te duiden *waarom* en *hoe* we die dwang ervaren, plaatst Korsgaard allereerst twee posities tegenover elkaar, die beiden denken dat normatieve standaarden *buiten* de persoon te vinden zijn. Ze stelt enerzijds dat David Hume denkt dat normatieve standaarden überhaupt geen psychologische krachten zijn die in en op onszelf inwerken, maar dat morele concepten gegenereerd worden vanuit het gezichtspunt van de toeschouwer, op basis van sympathie en daaruit voortvloeiende sentimenten. En anderzijds stelt ze dat

¹ Dit is gebaseerd op Kants Nötigung: “*De verhouding van objectieve wetten tot een niet volledig goede wil wordt weliswaar voorgesteld als de bepaling van de wil van een redelijk wezen door redenen van de rede, waaraan echter deze wil vanwege zijn natuur niet noodzakelijk gehoorzaamt.*” (Kant, 1785/2008, p. 413;85) Kant lijkt hier echter te suggereren dat dwang enkel ervaren wordt door een imperfecte rationele wil (‘niet volledig goede wil’), alsof de ervaring van dwang enkel toebehoort aan personen met wie iets mis is. Korsgaard wil dat juist weerleggen.

dogmatisch rationalisten denken dat de juistheid van een actie een objectief feit is over de externe wereld, waar je met de rede bij kunt komen en waaraan vervolgens de overtuigingen en acties worden geconformeerd; dogmatisch rationalisten transformeren volgens Korsgaard externe standaarden naar de rede zelf.

Als we vanuit een derde persoonsperspectief mensen irrationeel of ondeugdelijk noemen, bedoelen we dat ze gefaald hebben om zich aan bepaalde standaarden te conformeren. Maar dat falen is een uiterlijke manifestatie van een innerlijke conditie, en deze posities vertellen niet wat die innerlijke conditie is. Ze vertellen niet *waarom* en *hoe* we dwang ervaren. Volgens Korsgaard onderschatten en misvatten beide posities de rol van de ervaring van dwang in ons psychische leven.

Daarom zoekt ze aansluiting bij Socrates, die in *Politeia* over een term als ‘zelfbeheersing’ spreekt als een op het oog vreemd begrip, dat suggereert dat een sterkere zelf een dwang oplegt aan een zwakkere zelf, alsof “*in het binnenste van een mens ergens in verband met zijn psyche een verschil in kwaliteit bestaat en dat daar twee afzonderlijke elementen tegenover elkaar staan van verschillend niveau.*” (Plato, 2010, pp. 430e; 143-144). Hij suggereert dat een term als ‘zelfbeheersing’ als een spoor of aanwijzing is die de deugd heeft achtergelaten in de taal. De ervaring van dwang, denkt hij, legt iets belangrijks over de menselijke natuur bloot, over de constitutie van de menselijke ziel.

En dat is precies waar de filosofie van Korsgaard zich op richt, namelijk dat *de bron van normativiteit ligt in het menselijke project van zelf-constitutie*. In het morele leven is er sprake van een inspanning – een soort worsteling –, en degene die daarmee succesvol is noemen we goed of rationeel. Maar we worstelen volgens Korsgaard niet zozeer om een goed persoon te zijn, maar om *integer* te zijn, d.w.z. een psychische eenheid te worden, om een *eenheid als actor* te worden. Korsgaard betoogt daarom dat de normatieve standaarden de principes zijn op basis waarvan we die eenheid bereiken en actorschap mogelijk maken. En dat proces om een psychische eenheid te bereiken, wat we ervaren als een dwang, noemt zij zelf-constitutie.

2.3. *De morele waarde van een actie ligt besloten in de vorm van de actie*

Als we willen weten wat acties slecht of goed maakt, moeten we beginnen met ons af te vragen wat acties zijn, en wat hun functies zijn. In tegenstelling tot bijvoorbeeld utilitaristen als John Stuart Mill en Jeremy Bentham – die zeggen dat een actie in essentie *productie* is en de functie van een actie dus is het realiseren van een bepaald *doel*, op basis waarvan geoordeeld kan worden of een actie goed of slecht is – claimt Korsgaard aan de hand van Aristoteles (en Kant) dat het doel in de actie zelf ligt besloten. Een actie is altijd een *handeling-ter-realisatie-van-een-doel*.

Volgens Aristoteles zijn productie en actie twee verschillende dingen: “*De productie heeft namelijk wel een doel buiten zichzelf maar het handelen (actie) kan dat onmogelijk hebben: het geslaagde handelen is immers zelf het doel.*” (Aristoteles, 2009, p. 1140b5;185) Acties, in ieder geval goede acties die getuigen van deugd, worden gekozen ten behoeve van zichzelf, en niet voor iets dat zij produceren.

Van belang voor Korsgaard is de manier waarop Aristoteles het concept van een goede actie die getuigt van deugd omschrijft, namelijk als een actie die het juiste principe belichaamt – de *orthos logos*: een actie gedaan op het juiste moment, op de juiste manier, richting het juiste object en met het juiste doel. Het doel is opgenomen in de beschrijving van de actie, en het is die actie als geheel, inclusief het doel, die een actor verkiest te verrichten.

Korsgaard laat zien dat de manier waarop Aristoteles acties omschrijft overeenkomt met hoe Kant dat doet. Kant omschrijft een actie doorgaans met een maxime, die een structuur heeft van ‘*deze-handeling-verrichten-ter-realisatie-van-dat-doel*’. Hij test met de categorische imperatief niet zozeer of de handeling of het doel *afzonderlijk* als universele wet gewild moet worden, maar *de maxime als geheel*. Het is de maxime als geheel die al dan niet morele waarde heeft en al dan niet vanuit plicht verricht moet worden, op dezelfde manier als hoe bij Aristoteles de deugdzaamheid in de actie als geheel ligt besloten, dus inclusief het doel.

Op basis hiervan maakt Korsgaard een definiërend onderscheid tussen *handeling* en *actie*, wat ik de rest van de thesis zal proberen te volgen: een *actie* (action) is voor Korsgaard een *handeling-ter-realisatie-van-een-doel*, dus een omschrijving van de handeling inclusief het doel. Bij het begrip van de *handeling* (act) afzonderlijk is het doel ervan losgekoppeld. Het doen van een valse belofte teneinde wat extra geld te vergaren, is daarmee een *actie*. Het doen van een valse belofte, is daarmee een *handeling*.

Waarin ligt dan de morele waarde van een actie besloten? Als we vragen naar de reden voor een actie, zijn we vaak geneigd enkel de *bedoeling* ervan uit te leggen. Waarom reis je van Breda naar Amsterdam? Om mijn moeder te bezoeken. Dat leidt tot verwarring, omdat daarmee de schijn wordt gewekt dat de bedoeling van een actie ook de reden is die actie te verrichten. Maar als we er vanuit gaan dat een reden *normatief* zou moeten zijn, en de bedoeling van een actie wordt aangezien als de reden, dan volgt daaruit dat het de bedoeling is die normatief is voor de actor. Maar dat hebben noch Kant noch Aristoteles voor ogen. Bij Aristoteles gaat het om de deugdzaamheid, en bij Kant om de plicht, van een *actie als geheel*. Het doel op zichzelf kan nooit de bron zijn voor de reden; als het doel de actie als geheel niet waard maakt om te doen, ook al wordt het doel met de handeling gerealiseerd, dan vinden we het geen goede actie. Een actie als ‘van Breda naar Amsterdam reizen om een pakje zakdoekjes te kopen’ laat dat zien. Voor Aristoteles wordt een deugdzaam actie voor het juiste doel, op de juiste manier, op het juiste moment, en richting het juiste object gedaan. En volgens Kant ontstaat de normativiteit uit de autonomie, uit het onszelf wetten opleggen; we kiezen niet eerst een bepaalde bedoeling, die we tot een wet ombuigen, waar we vervolgens proberen aan te voldoen, omdat we dat nu verplicht zijn. Wat we als wetten willen zijn maxims, acties als geheel, waar bedoelingen of doelen onderdeel van uitmaken. Acties worden dus gekozen omwille van zichzelf.

Echter, zo lijkt het alsof elke actie om dezelfde reden wordt gekozen, namelijk omdat het gezien wordt als een actie die het omwille van zichzelf waard is te doen. Maar dat is niet wat Korsgaard stelt. In plaats daarvan claimt ze, aan de hand van Kant en Aristoteles, dat de morele waarde van een actie ligt besloten in de *vorm* van een actie als geheel, dus in de *wijze waarop* handeling en doel zijn *samengesteld*.

De categorische imperatief zoals Kant die voor ogen heeft is gericht op een specifieke actie die een actor van plan is te verrichten, en vraagt of de maxime ervan de vorm heeft van een universele wet. Als Kant het over vorm heeft, zo claimt Korsgaard, dan heeft hij het over vorm op een Aristotelische manier. Bij Aristoteles is de vorm van een ding de manier waarop de delen ervan zijn samengesteld en gearrangeerd zodat het in staat wordt gesteld te doen wat het doet, zijn functie kan vervullen. De categorische imperatief bij Kant vraagt als het ware of de delen van de maxime (voornamelijk handeling en doel) zo zijn gearrangeerd dat de maxime als geheel kan dienen als universele wet. En dat betekent dat die ‘wettelijkheid’ (lawfulness) een *interne eigenschap* is van de maxime, en niet een of andere externe of onafhankelijke wet waarmee de actie toevalligerwijs overeenstemt. Op dezelfde manier is bij Aristoteles de vorm van de *logos* bepalend voor die eigenschap. De onderdelen – handeling, doel, tijd, plaats en manier – zijn zodanig gearrangeerd dat ze gezamenlijk een harmonieus geheel vormen, en maken de actie als geheel waard om te verrichten. Een goede actie is daarom een goed gevormde actie.

2.4. Een actie constitueert een actor als eenheid

In relatie tot bovenstaand concept van een actie doet Korsgaard twee claims. De eerste is dat het essentieel is voor een actie dat de actie wordt verricht door een *actor*. Een actie vereist een actor aan wie we de betreffende beweging kunnen toeschrijven als auteur ervan. De tweede claim is dat een dergelijke actor een *eenheid* moet zijn: om een beweging van mijn geest of mijn lichaam als *mijn actie* te beschouwen, moet ik de beweging zien als een uitdrukking van mijzelf als een geheel, in plaats van als product van bepaalde krachten die op mij of in mij aan het werk zijn. In dat laatste geval overkomen mij de dingen, in het eerste geval ben ik de auteur ervan: “*For a movement to be my action, for it to be expressive of myself in the way that an action must be, it must result from my entire nature working as an integrated whole.*” (Korsgaard, 2009, pp. 18-19)

Het lijkt logisch te veronderstellen dat een actie een expressie of resultaat is van een eenheid van de actor die al aanwezig is, een integriteit die al bereikt is. Eerst bereik je het soort eenheid of integriteit dat nodig is om bewegingen aan jou toe te kunnen schrijven, en daarna zijn de keuzes voor acties die je maakt uitdrukking van die reeds aanwezige eenheid. Maar Korsgaard draait dit juist om. Zij beargumenteert dat er geen eenheid is *voorafgaand* aan de keuzes en acties, omdat die eenheid (op een vrij letterlijke manier) *geconstitueerd* worden door de keuzes en acties.

Wat ons mensen volgens Korsgaard onderscheid van andere dieren, is dat wij een (zelf)bewustzijn hebben van de *gronden op basis waarvan we kiezen en handelen*. En dat maakt dat we ze kunnen sturen. Als je bewust besluit welke soorten effecten je in de wereld wil brengen, besluit je daarmee ook bewust wat voor soort oorzaak je zult zijn. En dat betekent dat je besluit *wie* je zult zijn. Daarom zijn we allemaal geconfronteerd met de taak van het construeren van een specifieke individuele identiteit, een praktische identiteit, waar we ook zelf verantwoordelijk voor zijn. Onder het concept praktische identiteit verstaat Korsgaard zoiets als een beschrijving waaronder je jezelf waardeert en het leven waard vindt geleefd te worden en jouw acties waard om te ondernemen. Onze conceptie van onze praktische identiteit stuurt onze keuzes voor acties, omdat het jezelf in een bepaalde rol waarderen tegelijkertijd betekent dat je sommige handelingen voor bepaalde doelen waardeert, en andere niet. En als eigenaar van een praktische identiteit ben je de auteur van jouw acties, en er daarom verantwoordelijk voor. Maar tegelijkertijd met het kiezen van jouw acties *creëer* je die praktische identiteit: met het kiezen van jouw acties constitueer je jezelf als de auteur van jouw acties.

Omdat we volgens Korsgaard als rationele wezens *moeten* handelen, heeft elk menselijk wezen dus een praktische identiteit nodig om redenen te hebben voor het handelen. Met het onszelf tot specifiek persoon constitueren committeren we ons dus aan onze rol als rationele actor (zie o.a. Korsgaards *The Sources of Normativity* (1996, pp. 113-125; 132-145)). En in *Self-constitution* laat Korsgaard zien dat vanuit dat noodzakelijke commitment de hypothetische en categorische imperatief de twee principes zijn op basis waarvan we acties moeten kiezen, willen de acties ons tot een actor als eenheid constitueren.

2.5. *Constitutieve principes als formele principes hebben normatieve kracht*

De principes van de praktische rede dienen dus om ons tot actor als eenheid te constitueren. En daarin ligt hun normativiteit besloten. Dit baseert Korsgaard voornamelijk op de teleologie van Aristoteles.

Volgens Aristoteles is dat wat een object tot een object maakt of wat het zijn identiteit geeft, datgene wat het doet: zijn functie. Dit is duidelijk te zien bij gebruiksvoorwerpen, waarbij de functie helder is gedefinieerd en het materiaal ervan (de verschillende delen) zodanig is gearrangeerd dat het die functie kan vervullen. Het zijn van een object betekent op een bepaalde manier teleologisch georganiseerd zijn, en als zodanig een eenheid zijn, in plaats van een *'mere heap'* van materialen. Een hoopje bakstenen en dakpannen maakt nog geen huis. Dit betekent dat de teleologische organisatie van een object zich conformeert aan bepaalde normatieve standaarden, op basis waarvan gezegd kan worden dat het object zijn functie kan verrichten. De bakstenen en dakpannen moeten zodanig gearrangeerd worden dat het bijvoorbeeld een ruimte creëert dat de regen buiten houdt. Doet het dat niet, dan is het geen goed object in het licht van de functie die het moet verrichten. Dergelijke normatieve standaarden zijn daarmee tegelijkertijd ook constitutieve standaarden. En waar het niet objecten, maar *activiteiten* betreft, kunnen het constitutieve *principes* worden genoemd, die gevolgd moeten worden wil er überhaupt van activiteit sprake zijn. Alle activiteit wordt op de een of andere manier gericht of gestuurd door degene die de activiteit onderneemt, en de principes die beschrijven hoe degene dat moet doen zijn de constitutieve principes voor die activiteit. Volg je die principes niet, dan voer je die activiteit niet uit. Wie een huis probeert te bouwen zonder de constitutieve principes daarvan in acht te nemen – bijvoorbeeld: er moet een waterdicht dak op – is niet bezig met de activiteit van het bouwen van een huis, maar met iets anders. Toch zien we regelmatig slechte huizen gebouwd worden. Volgens Korsgaard is dat dezelfde activiteit, maar slecht uitgevoerd. Het gaat haar erom dat je je met de uitvoering van de activiteit laat leiden door de constitutieve principes, en soms lukt dat beter dan andere keren. Laat je je niet door de principes leiden, dan ben je niet met die activiteit bezig.

Korsgaard wil hiermee laten zien dat een normatief principe daadwerkelijk autoriteit heeft, opdat ze kan laten zien dat de wetten van de praktische rede de normatieve en constitutieve principes zijn voor onze acties: als we ze niet volgen, dan verrichten we geen acties. Maar omdat we genoodzaakt zijn acties te verrichten, ontkomen we er niet aan.

Ook een persoon heeft een vorm, een functie: een persoon als levend ding is zo ontworpen dat het zichzelf in stand houdt en reproduceert. Een persoon is zijn eigen doel, zijn functie is te zijn wat het is. En dat maakt dat het zijn van een persoon niet zozeer een toestand is, maar een *activiteit*. Wat het is om een persoon te zijn, een rationele actor, is niets anders dan het voortdurend bezig zijn jezelf tot persoon te constitueren. Dit

is niet alsof een of andere ambachtsman op mysterieuze wijze zichzelf als zijn eigen product maakt, maar een voortdurend zelf-constitutief proces als essentiële activiteit van het leven. En om die activiteit goed te verrichten, dien je de betreffende constitutieve principes te volgen. Dus als Korsgaard zegt dat actie zelf-constitutie is, dan bedoelt ze hiermee dat we met de acties die we kiezen te verrichten onze eigen praktische identiteiten constitueren, evenals ons actorschap.

Deze ogenschijnlijk paradoxale claim is afhankelijk van de veronderstellingen dat voor actie een actorschap nodig is, en dat voor actorschap een eenheid nodig is. Met acties kan eenheid gerealiseerd worden, maar met sommige acties gaat dat beter dan met andere. De acties die het beste in staat zijn eenheid te realiseren, zijn de acties die zich conformeren aan de principes van de praktische rede, die we op z'n minst moeten proberen te volgen, willen we überhaupt kunnen handelen. En dat verklaart waarom ze ons binden, wat hun normativiteit is.

Van belang hierbij is om een onderscheid aan te brengen in *formele* en *materiële* principes. Een materieel moreel principe identificeert moraliteit in termen van de inhoud; een formeel moreel principe doet dat in termen van een *manier* om tot een bepaalde inhoud te komen. We zijn vaak geneigd over moraliteit te denken in materiële termen (je beloftes nakomen, de waarheid vertellen, anderen helpen). De morele waarde wordt dan gebaseerd op inhoudelijke gronden. Maar die kunnen ter discussie gesteld worden. Bij formele principes gaat het juist om de manier van redeneren waarop een inhoud tot stand komt. Het duidelijkste voorbeeld hiervan is Kants categorische imperatief. Moreel handelen betekent volgens Kant ervan verzekerd zijn dat jouw maxime als universele wet gewild kan worden. Moraliteit zit hem dus niet in een set aan overtuigingen die inhoudelijk geladen zijn, maar in een manier van delibereren. Het gaat er bij Kant dus ook niet zozeer om *wat* een actor besluit te doen, maar *hoe* hij tot dat besluit komt.

2.6. *De constitutieve principes van een actie zijn de imperatieven*

De vraag is *hoe* de principes van de praktische rede helpen tot een actor als eenheid te komen en *waarom* ze dat normatief maakt. Korsgaard laat dit zien aan de hand van Kants uitleg van de hypothetische imperatief. De hypothetische imperatief is in feite het gebod dat aangeeft dat als je een doel wilt, je de middelen wilt om dat doel te realiseren. Kant zegt daarover:

Wie het doel wil, wil (voor zover de rede beslissende invloed heeft op zijn handelingen) ook het daartoe op onontbeerlijke wijze noodzakelijke middel dat binnen zijn macht ligt. Deze positie is, wat het willen betreft, analytisch, want in het willen van een object als mijn effect is reeds mijn causaliteit als handelende oorzaak, dat wil zeggen: het gebruik van middelen, begrepen en de imperatief haalt uit het begrip van het willen van dit doel reeds het begrip van tot dat doel noodzakelijke handelingen naar voren [...]
(Kant, 1785/2008, p. 417;90)

Wat dit volgens Korsgaard vooral laat zien is dat de hypothetische imperatief een constitutief principe is voor het willen. Wat het willen iets anders maakt dan iets verlangen of wensen, is dat de persoon die een doel wil, zichzelf bepaalt tot oorzaak van dat doel. Hij *constitueert zichzelf* tot oorzaak van dat doel. De reden waarom ik me moet conformeren aan de hypothetische imperatief is dat als ik dat niet doe, als ik altijd mezelf toelaat me af te laten leiden van de weg naar het doel, ik dat doel niet echt wil. Er is dan geen onderscheid meer tussen de impulsen die op mij inwerken en mijn wil; ik, als actor, besta dan niet. Conformiteit aan de hypothetische imperatief is dus constitutief voor het hebben van een wil.²

² Overigens claimt Korsgaard dat bij nadere beschouwing de hypothetische imperatief niet een afzonderlijk principe naast de categorische imperatief is. Daar in haar optiek een actie altijd een handeling-ter-realisatie-van-een-bepaald-doel is, is elke keuze voor een actie *uiteindelijk* gebaseerd op de categorische imperatief. Toch is het nuttig de hypothetische imperatief afzonderlijk te beschouwen en te gebruiken, omdat in de claim dat 'actie betekent het jezelf constitueren tot oorzaak van een doel', de hypothetische imperatief betrekking heeft op het jezelf constitueren tot *oorzaak* van een doel, terwijl de categorische imperatief betrekking heeft op het *jezelf* constitueren tot oorzaak van een doel.

Als actie zelf-constitutie is, dan is daarnaast ook de categorische imperatief een constitutief principe voor actie. Het jezelf bepalen als oorzaak van een doel is namelijk niet hetzelfde als door 'iets' in jou bewogen worden, bijvoorbeeld een wens of een impuls (of zoals Kant het noemt: een *aansporing*) dat als oorzaak functioneert. Wanneer je delibereert, wanneer je je eigen causaliteit bepaalt, is het alsof er iets is dat al jouw aansporingen overstijgt, een *jij*, die kiest op basis van welke aansporingen gehandeld zal worden. Dus wanneer je jouw causaliteit bepaalt, *moet* je opereren als een geheel. En om dit te kunnen moet je jouw maximes als universele wet willen.

Korsgaard legt dit onder andere uit door te stellen dat als onze redenen niet universeel hoeven te zijn, ze volledig particularistisch zouden kunnen zijn; dan zou het mogelijk zijn een reden te hebben die enkel voor jou in een specifieke situatie geldt, zonder implicaties voor andere situaties. Maar, als particularistisch willen niet mogelijk is, dan *moet* willen universeel zijn, d.w.z. dan moet een maxime, om gewild te kunnen worden, gewild worden als een universele wet.

En volgens haar *is* particularistisch willen niet mogelijk. Je kunt enkel de keuze voor een actie als *jouw* keuze voor een actie beschouwen wanneer je het principe op basis waarvan je hebt gekozen beschouwt als een expressie of een representatie van jezelf. Om iets echt particularistisch te willen dien je je volledig te identificeren met de aansporing op basis waarvan je kiest te handelen; die aansporing zou op dat moment gelden als jouw wet die jouw actorschap en wil definieert. Maar daardoor verdwijnt het onderscheid tussen de persoon en de aansporing op basis waarvan hij handelt, en is er geen ruimte meer voor actorschap die zijn zelfbepaalde wil is als onderscheidend van de veelheid aan aansporingen die op hem inwerken. Je bent geen heel persoon, maar een *mere heap* van willekeurige impulsen. Er is in die zin geen verschil tussen iemand die een particularistische wil heeft en iemand die helemaal geen wil heeft. Het subject ontbreekt, de persoon die de oorzaak is van zijn acties.

Kortom: omdat particularistisch willen onmogelijk is, moet je zodra je een maxime wilt, deze als universeel willen. Doe je dat niet, dan bepaal je jezelf niet tot het willen van een actie, en wil je niet. De categorische imperatief is dus constitutief voor actie, omdat het je conformeren aan de categorische imperatief constitutief is voor de wil, voor de wilsbepaling van een persoon *zelf*, in plaats van voor de wilsbepaling door iets in hem.

2.7. *Het zelfbewustzijn creëert de delen van de ziel*

Wat ons mensen volgens Korsgaard onderscheid van andere dieren is dat wij over een specifiek soort zelfbewustzijn beschikken. Wij zijn ons bewust, niet alleen van het feit dat we bepaalde zaken verlangen of vrezen, maar ook dat we op bepaalde manieren zullen moeten handelen op basis van die verlangens en angsten. We zijn ons bewust van de potentiële gronden van onze acties, de principes waarop onze acties zijn gebaseerd. En dat geeft ons een probleem dat andere dieren niet hebben. Want zodra we ons bewust zijn dat we op een bepaalde manier zullen moeten handelen op grond van een bepaalde aansporing, worden we geconfronteerd met een te maken keuze, namelijk: zou ik dat moeten doen of niet? Dit betekent dat een dergelijke vorm van zelfbewustzijn een soort bevrijding van de dierlijke instincten inhoudt. Instincten *bepalen* niet hoe we moeten reageren op aansporingen, maar doen een *voorstel* hoe we zouden kunnen reageren. En dankzij het zelfbewustzijn ontstaat de vraag of de aansporingen ons *redenen* geven voor actie. En om die vraag te beantwoorden hebben we *principes* nodig die bepalen wat we als reden kunnen duiden. Hier wordt dus de rede geboren, die ons naar onszelf doet kijken, en die ons confronteert met normatieve problemen over wat we moeten doen voor welk doel.

Maar het zelfbewustzijn introduceert ook nog iets anders, namelijk wat Korsgaard (naar Plato³) noemt 'de delen van de ziel'. Wanneer we ons bewust worden van de werking van een aansporing in ons, wordt die aansporing niet zozeer als een noodzakelijkheid ervaren, maar als een voorstel, iets waar we een besluit over moeten nemen. Omdat we losgeweekt zijn van de dwingende werking van de instincten, moeten we principes formuleren die ons vertellen hoe we moeten omgaan met de aansporingen die we ervaren. En de ervaring van het maken van een keuze of een besluit, is een andere ervaring dan die van de werking van de

³ Zie daartoe onder andere het deel *Elementen in de Psyche* in *De Ideale Staat* (Plato, 2010, pp. 435-445/149-163).

aansporing zelf. Zelfbewustzijn creëert dus enerzijds de behoefte aan principes op basis waarvan iets besloten kan worden, maar anderzijds transformeert het de aansporingen tot uiteindelijke *drijfveren* (*inclinations*), als zijnde de responsen van iemand op de aansporingen, waarover we nu moeten besluiten daar al dan niet gehoor aan te geven. Zo is het zelfbewustzijn niet alleen de bron van de rede, maar ook van de drijfveren: het zelfbewustzijn produceert de delen van de ziel.

Hiermee is overigens nog niet gezegd dat het de drijfveer is die je een reden geeft om een actie te verrichten, want, zoals gezegd, dat is iets waarover je een *besluit* zult moeten nemen, op basis van de rationele principes die aangeven waar een actie aan moet voldoen wil die actie het waard zijn om te verrichten, en wil dus de drijfveer een reden zijn voor een actie.

Kortom: zodra we een zelfbewustzijn hebben, heeft onze ziel verschillende delen, en om een actie te kunnen verrichten zullen we die delen weer bijeen moeten brengen tot een eenheid. Tegelijkertijd is een praktische deliberatie noodzakelijk, omdat de rede niet is gedwongen een drijfveer te volgen: we moeten besluiten wat te doen. En dit betekent dat de deliberatie niet enkel de functie heeft te besluiten wat je zult doen, maar ook dat het jou weer tot een eenheid maakt, omdat jouw beweging geen actie zal zijn als de beweging niet is toe te schrijven aan jou als eenheid.

2.8. *Deliberatie conform de constitutieve principes constitueert eenheid*

Hoe vindt de deliberatie die leidt tot actie en eenheid plaats? Korsgaard formuleert hiertoe het 'Constitutioneel Model', zich baserend op zowel Kant als Plato. Het Constitutioneel Model ziet de actor niet als een afzonderlijke entiteit die ervoor kiest zich te identificeren met een van de delen van de ziel, maar als iets dat boven die delen uitstijgt, zoals een constitutie van een stad boven zijn bewoners en ambtenaren uitstijgt. Als de actor zich conformeert aan de rede, dan is dat niet omdat de actor zich identificeert met de rede, maar met zijn constitutie, die zegt dat de rede moet heersen. Volgens Korsgaard kunnen we dit model goed gebruiken om actie te verklaren, omdat we ermee kunnen uitleggen hoe een actor de eenheid bereikt die het mogelijk maakt zijn bewegingen aan hem toe te schrijven als auteur ervan.

Ze baseert dit model op Plato, die de menselijke ziel vergelijkt met de constitutie van een *polis* of een stad. Hij identificeert de deugden van de ideale, rechtvaardige stad met bepaalde eigenschappen van de verschillende delen van de stad. De *wijsheid* van de stad correspondeert met de wijsheid van de regeerders, de koning-filosofen, die zodanig regeren dat ze streven naar het goede voor de stad als geheel, en niet naar een eigen belang (ze hebben dan ook geen bezittingen); de *moed* van de stad correspondeert met de moed van de helpers, die erop toezien dat de wetten van de regeerders worden nageleefd en de stad beschermen. De werkers en burgers laten zich leiden door de *begeerte*. De *gematigdheid* van de stad, de *sophrosyne*, correspondeert met de overeenstemming tussen alle klassen over wie er zou moeten regeren en wie er geregeerd zou moeten worden. En de *rechtvaardigheid* van de stad schuilt erin dat elke klasse zijn eigen werk doet en zich niet bemoeit met het werk van de andere klassen. Hij laat vervolgens zien dat dezelfde delen ook in de menselijke ziel zijn terug te vinden. De *rede* correspondeert met de regeerders, en de functie ervan is te besturen ten behoeve van het goede van de hele persoon. *Geestdrift* correspondeert met de helpers, en de functie ervan is de orders van de rede uit te voeren. En de *begeerte* correspondeert met de werkers, met als functie de hele persoon te voorzien in wat hij nodig heeft.

Korsgaard ziet een overeenkomst tussen de drie delen van de ziel en de delen van deliberatieve actie. Deliberatieve actie begint met het feit dat we bepaalde dingen willen, begeren of wensen. Daar zijn we ons bewust van, en ze nodigen ons uit om bepaalde acties te verrichten. Maar omdat we zelfbewust zijn, handelen we niet automatisch op basis van onze begeerten en wensen, maar nemen we daar een besluit over, waarna we de actie daadwerkelijk verrichten. Kortom: de begeerte doet een voorstel, de rede neemt er een besluit over en de geestdrift voert het besluit uit. Dit komt overeen met Plato's analogie tussen de ziel en de stad. Ook een stad verricht deliberatieve acties; politieke besluiten bijvoorbeeld komen tot stand doordat bewoners iets wensen en daarom een voorstel doen, waar de bestuurders een besluit over nemen. Als ze instemmen gaan de helpers aan de slag om het besluit uit te voeren. Alleen als deze procedure wordt gevolgd kunnen we de actie aan de stad toeschrijven. Het is de constitutie van een stad die deze procedures aangeeft. De constitutie maakt het als het ware mogelijk voor een groep burgers – die zonder constitutie als een *mere*

heap van individuen zouden zijn – te functioneren als een ééngeworden actor. En dat is precies waaruit de normatieve kracht van de constitutie bestaat. Zowel voor de stad, als voor de ziel: “[p]las wanneer we [...] ons volledig tot een eenheid, een harmonische, evenwichtige persoonlijkheid hebben ontwikkeld, dan pas mogen wij handelen en doen wat ons te doen staat. [...] Elke handeling die ons innerlijk evenwicht verstoort is immoreel.” (Plato, 2010, pp. 443;160-161)

De moeilijkheid is dat, zoals Plato aangeeft, een onrechtvaardig of immoreel persoon helemaal niet kan handelen, omdat een onrechtvaardig of immoreel persoon niet tot eenheid is gebracht, maar er in plaats daarvan intern een soort strijd om de macht gaande is tussen de delen van de ziel. En op dezelfde wijze lijkt Kant aan te geven dat alleen een autonome actie, bestuurd door de categorische imperatief, een echte vrije actie is, terwijl slechte of heteronome actie meer gedrag is dat wordt veroorzaakt door verschillende verlangens en aansporingen in ons. Waarom zouden we dat überhaupt nog als acties moeten zien?

Het antwoord van Korsgaard hierop is dat, als de principes *constitutieve* principes zijn, dan zijn acties die aan die principes voldoen goed *als actie*. Op dezelfde manier zijn acties die niet aan die principes voldoen niet zozeer geen actie, maar slecht *als actie*: defect. Een defecte actie slaagt er niet in zijn actor tot eenheid te brengen, en slaagt er dus ook niet in de actor een doeltreffende en autonome auteur van zijn bewegingen te laten zijn.

Maar kan een actie defect zijn, en toch een actie? Volgens Korsgaard biedt het Constitutioneel Model daar zelf een antwoord op. Een actie van een stad kan formeel of procedureel constitutioneel zijn, maar toch niet materieel rechtvaardig: ze kan bestuurd worden op basis van de verkeerde wet. En dat geeft volgens Plato en Kant weer hoe defecte actie gedacht kan worden. Kant spreekt wat dat betreft over een slecht persoon als iemand die, door het principe van *zelfliefde*, ervoor *kies*t te handelen op basis van de drijfveren; hij neemt zijn drijfveren, zonder verdere reflectie, als redenen voor actie: “[w]anneer hij deze drijfveren in zijn maxime zou opnemen als volstonden ze op zich alleen *ter bepaling van de willekeur, zonder zich om de morele wet (die hij toch in zich draagt) te bekommeren, dan zou hij moreel slecht zijn.*” (Kant, 1793/2010, p. 33;42) En Plato onderscheidt vijf manieren waarop de ziel bestuurd kan worden, te vergelijken met vijf manieren waarop een stad bestuurd kan worden, waarvan er één de rechtvaardige manier is, de *aristocratie*, en de overige onrechtvaardig, namelijk de *timocratie*, de *oligarchie*, de *democratie* en de *tirannie*. Bij deze onrechtvaardige vormen neemt een deel van de ziel het werk over van de rede, en zet zo een principe neer dat enkel goed is voor dat deel van de ziel, in plaats van goed voor het geheel. Er is nog steeds sprake van actie, maar het is slechte actie, omdat de actie wordt gekozen volgens principes die niet uitgaan van het goede voor het geheel. Daarmee is de eenheid die wel wordt bereikt instabiel en contingent, omdat de omstandigheden toevalligerwijs zodanig zijn dat ze de strijd tussen de delen niet blootleggen en de doeltreffendheid niet ondermijnen. Maar het is geen integere eenheid.

2.9. *Actie conform de imperatieven constitueert integere eenheid*

Je kunt je afvragen waarom we er niet voor zouden kiezen om dan maar geen eenheid te zijn, en ons gewoon laten leiden door bijvoorbeeld de impulsen die op ons afkomen, of de eer die we zoeken, als dat niet per se leidt tot een uiteenvallen van de delen van de ziel. Zolang we kunnen handelen en de doeltreffendheid daarvan niet noodzakelijkerwijs wordt ondermijnd, lijkt er geen echt probleem te zijn.

Korsgaard geeft hier antwoord op door te verwijzen naar Glaucon, die Socrates uitdaagt een antwoord te geven op de vraag waarom het beter voor iemand is om rechtvaardig te zijn in plaats van onrechtvaardig, zelfs in een situatie waarin een rechtvaardig persoon door iedereen gezien en behandeld wordt als was hij onrechtvaardig, en een onrechtvaardig persoon door iedereen gezien en behandeld wordt als was hij rechtvaardig. Korsgaard laat zien dat deze uitdaging vooral een *praktische* vraag is. Het gaat niet zozeer over de vraag of het rechtvaardige leven *beter* is dan het onrechtvaardige leven, maar of het rechtvaardige leven *te verkiezen* is boven het onrechtvaardige leven. En als je ervoor kiest een rechtvaardig persoon te zijn, dan kies je daarvoor zelfs al zou je door iedereen behandeld worden als een onrechtvaardig persoon. Je kunt er niet naar streven rechtvaardig te willen zijn onder bepaalde voorwaarden; de rechtvaardigheid schuilt in de *manier* waarop je je committeert, niet in de eventuele gevolgen of consequenties ervan. Dit betekent natuurlijk niet

dat je niet kunt falen om je streven naar rechtvaardigheid te realiseren, maar dat maakt je nog niet onrechtvaardig, want het is niet iets waar je op voorhand voor koos.

Volgens Plato houdt een rechtvaardige actie de innerlijke harmonie in stand; bij alles wat wij doen “*mogen wij een handeling alleen als rechtvaardig en goed beschouwen en zo noemen, als zij bijdraagt tot de handhaving en vervolmaking van die innerlijke toestand.*” (Plato, 2010, pp. 443;160-161) Het principe van rechtvaardigheid geeft ons de richting om die acties te verrichten die onze willende eenheid realiseren en in stand houden. En volgens Plato is die willende eenheid noodzakelijk om als een eenheid, te handelen. Dus het principe van de rede is simpelweg het principe van het handelen op een manier die jou tot eenheid constitueert, op dezelfde manier als Kants categorische imperatief het principe is op basis waarvan we onszelf tot eenheid constitueren. Dat maakt het volgens Korsgaard een *formeel* principe van deliberatieve actie: “*deliberative action by its very nature imposes unity on the soul. When you deliberate about what to do and then do it, what you are doing is organizing your appetite, reason, and spirit, into a unified system that yields an action that can be attributed to you as a person.*” (Korsgaard, 2009, p. 179)

Rechtvaardigheid is daarmee niet een of andere randvoorwaarde waar een actie aan moet voldoen wil het de eenheid van de ziel realiseren en behouden. Rechtvaardigheid *is* het realiseren en behouden van de eenheid van de ziel. En op dezelfde manier is de categorische imperatief niet een of ander principe dat ons in staat stelt autonoom te handelen, maar het handelen volgens de categorisch imperatief *is* autonoom handelen.

Toch is er nog een probleem. Volgens Kant vereist actie autonomie, en dat vereist universaliseerbaarheid. Volgens Plato vereist actie constitutionele eenheid, en dat vereist procedurele rechtvaardigheid in de ziel. Beiden zeggen dat deze eigenschappen enkel volledig gerealiseerd worden in de ziel van de moreel goede persoon. Maar hoe komen we van de genoemde formele eigenschappen (universaliseerbaarheid en procedurele rechtvaardigheid) naar een commitment aan materiële moraliteit?

Volgens Korsgaard wordt de relatie tussen het bewaren van de eenheid van de ziel en het vertellen van de waarheid, je beloftes nakomen e.d., zichtbaar door te kijken naar de relatie tussen een naar buiten gerichte rechtvaardigheid (outward justice) en een innerlijke rechtvaardigheid (inward justice). De claim is dat de innerlijke rechtvaardigheid *voorwaardelijk* is voor een naar buiten gerichte rechtvaardigheid.

Korsgaard baseert zich hiertoe op Kants theorie over interactie. Elke persoon interacteert met anderen zoals hij met zichzelf interacteert. Daarom moeten redenen voor acties publiek zijn, in de zin dat er een normatieve kracht van uit moet gaan die door alle rationele wezens gedeeld kan worden, inclusief door jou zelf op een ander moment of in een andere situatie. Dit betekent dat je in de interactie met een ander, de redenen van die ander als jouw redenen moet beschouwen. Het vereist het respecteren van de mensheid van de ander. Wanneer we samen interacteren, maken we als het ware samen wetten die ons samen binden. En omdat wanneer we zelf handelen we in feite interacteren met onszelf, dienen we in het delibereren over onze acties in ons zelf de mensheid te respecteren, en onze eigen redenen te beschouwen als publieke redenen met een normatieve kracht: “*Without respect for the humanity in your own person, it is impossible to will the laws of your own causality, to make something of yourself, to be a person; [...] unless you constitute yourself as a person, it will be impossible for you to act at all.*” (Korsgaard, 2009, p. 204)

Het is niet mogelijk om als actor een publieke wet te willen die alleen voor hemzelf opgaat, om alleen zijn eigen mensheid te respecteren: “*[w]hat is your own, in the individual sense of your own, is not your humanity but what you make of it, your practical identity, and the existence of that depends on your respect for humanity in general.*” (Korsgaard, 2009, p. 204) De innerlijk rechtvaardige persoon is daarmee ook rechtvaardig naar anderen. Het jezelf wetten opleggen is hetzelfde als het opleggen van universele wetten; doe je dat niet, dan houden ze je jou niet bijeen als een eenheid.

In het proces van het bijeenbrengen van de delen van jouw ziel, constitueer je *jezelf*. De manier waarop je jezelf tot een actor maakt, een persoon, is het jezelf tot een *specifiek* persoon maken, met een eigen praktische identiteit. En de manier waarop je jezelf tot een specifiek persoon maakt, die kan interacteren met zichzelf en met anderen, is om consistent te zijn, een eenheid – om *integriteit* te hebben. En als je jezelf goed constitueert, als je goed bent in het zijn van een persoon, dan ben je een goed persoon. De morele wet is de wet van zelf-constitutie.

3. Zelf-constitutie als grond voor integriteit van de organisatie

3.1. *Inleiding*

Op basis van de zelf-constitutietheorie van Korsgaard is het volgens mij goed mogelijk tot een voorstel te komen voor het denken over integriteit van de organisatie. Dat voorstel zal ik hieronder uitwerken, waarna ik dit in het volgende hoofdstuk zal afzetten tegen de belangrijkste argumenten die in het huidige debat over de organisatie als morele actor geformuleerd worden.

Het voorstel komt neer op de claim dat, als de functie van een actie de zelf-constitutie van de actor is, dan is de functie van een *actie namens de organisatie*, de constitutie van de *organisatie als actor*, in het licht van zijn functie. En waar de functie van persoon-zijn is het zijn van een persoon, is de functie van het zijn van een organisatie het zijn van een in de samenleving ingebed instrument. Dit betekent dat acties namens de organisatie de organisatie constitueren tot een in de samenleving ingebed instrument.

Volgens Korsgaard kan er alleen maar sprake zijn van een actie die aan een actor kan worden toegeschreven, als die actie de actor *doeltreffend* en *autonoom* maakt. Dit betekent dat acties pas acties namens de organisatie zijn, als de acties de organisatie constitueren tot een doeltreffende en autonome organisatie. Acties dienen daarom deliberatief, conform bepaalde constitutieve principes, gekozen te worden, zodanig dat de organisatie als eenheid daar achter kan staan. Hiertoe moet zij zichzelf zodanig inrichten dat ze *kan* delibereren over de te verrichten acties zodanig dat dit tot eenheid leidt. En daartoe dient de constitutie van de organisatie, die zegt *hoe* die deliberatie plaats moet vinden, hoe de te verrichten acties gekozen moeten worden. De constitutie van de organisatie levert als het ware de normatieve principes op basis waarvan te verrichten acties gekozen kunnen worden. En dat moeten de hypothetische en categorische imperatief zijn.

Omdat de functie of praktische identiteit van de organisatie betrekking heeft op de wijze waarop de organisatie is ingebed in de samenleving, dient de organisatie zodanig teleologisch georganiseerd te zijn dat de organisatie die functie kan vervullen. En dat veronderstelt dat het zich als organisatie inbedden in de samenleving niet zozeer een materiële keuze is, maar een (derde) normatief principe waaraan acties zich dienen te conformeren, wil er *überhaupt* sprake zijn van een organisatie.

Bovendien volgt vanuit dat normatieve principe van het inbedden in de samenleving dat de organisatie *interacteert* met de samenleving, en dus dat redenen voor acties namens de organisatie *publieke* redenen moeten zijn, in de zin dat er een normatieve kracht vanuit moet gaan die door alle rationele wezens gedeeld kan worden, inclusief door de organisatie zelf op een ander moment of in een andere situatie: de maximes voor de acties dienen universeel gewild te kunnen worden.

De organisatie *is* ook daadwerkelijk in staat de normatieve principes te integreren in een constitutie. Dat wordt geïllustreerd met de interne besluitvormingsstructuur, waarin een bepaalde vorm van *redelijkheid* is geborgd op basis waarvan voorstellen tot acties kunnen worden getoetst op hun waarde als actie-ter-realisatie-van-een-bepaald-doel. Die redelijkheid krijgt vooral een plaats in enerzijds een identificatie van organisatierepresentanten met de ingebedde functie-rationaliteit, en anderzijds in de mate waarin een zelfcorrigerende werking binnen de structuur is geborgd.

De waarde toetsen van een actie-ter-realisatie-van-een-doel betekent de maxime ervan toetsen aan de hypothetische en categorische imperatief, én aan het normatieve principe dat uit de praktische identiteit van de organisatie volgt, de inbedding in de samenleving. Alleen dan constitueren acties namens de organisatie de organisatie tot een doeltreffende en autonome eenheid, die als een integer geheel bijdragen aan de samenleving levert. En daarmee goed is in het zijn van een organisatie.

3.2. *Methodologische kanttekening*

Als ik de zelf-constitutietheorie van Korsgaard als basis wil hanteren voor een voorstel voor het denken over integriteit bij organisaties, dan dien ik allereerst hetzelfde perspectief te hanteren als Korsgaard. Korsgaard beargumenteert nadrukkelijk vanuit een *teleologisch* perspectief dat een persoon – ontologisch gezien – een bepaalde functie heeft, namelijk persoon-zijn. Vanuit die functie kan zij stellen dat een persoon voortdurend

bezig is om die functie te vervullen, dus voortdurend bezig is met de activiteit van het zijn van een persoon. De verrichte acties moeten in die zin aan bepaalde principes voldoen willen ze de persoon als autonome en doeltreffende actor van die acties constitueren.

Om op basis hiervan tot een voorstel te komen voor het denken over integriteit bij de organisatie, dien ik dus vanuit een teleologisch perspectief naar de organisatie te kijken, door te stellen dat de organisatie een bepaalde functie heeft. Mijn eerste taak wordt dan te komen tot de formulering van wat die functie van de organisatie is, waarbij ik de organisatie niet zozeer als een empirisch fenomeen, maar als een ontologisch idee hanteer (dat is nadrukkelijk niet hetzelfde als een beschrijving van de kenmerken van alle bestaande organisaties, maar een *normatieve* beschrijving waar een organisatie aan moet voldoen wil er überhaupt sprake zijn van een organisatie). Vanuit dat normatieve idee van de organisatie kan ik vervolgens beargumenteren wat constitutief is voor de organisatie, en welke constitutieve principes daaruit voortvloeien waar acties namens de organisaties aan moeten voldoen.

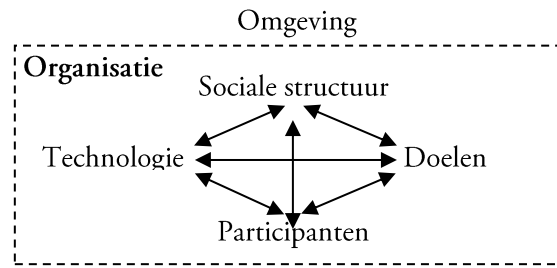
Ik zou deze taak kunnen vervullen door een ontologische studie naar ‘het wezen van de organisatie’ te verrichten, maar dat is een te grote opgave binnen de reikwijdte van deze thesis. Ik zou ook de onderscheidende kenmerken die alle organisaties gemeen hebben kunnen inventariseren, en van daaruit een wezenlijke functie van de organisatie trachten te beargumenteren. Maar ik denk enerzijds dat dat voor deze thesis eveneens een te grote opgave is, en anderzijds dat dat methodologisch kwetsbaar is, in die zin dat met een opsomming van onderscheidende kenmerken nog niet vanzelfsprekend alle kenmerken benoemd zijn, laat staan de belangrijkste kenmerken en het wezenlijke ervan. Bovendien *veronderstelt* een dergelijke aanpak reeds een bepaald idee van de organisatie. Het is als het ware eenzelfde exercitie als proberen het wezen van een boom te formuleren op basis van de kenmerken van alle bomen. Zodra we bezig zijn met het zoeken naar de kenmerken van alle bomen, vragen we onszelf al af wat we wel en wat we niet als boom mogen beschouwen; het ‘wezen van de boom’ is als het ware al als normatief idee verondersteld in onze zoektocht naar de kenmerken. En hetzelfde geldt voor de zoektocht naar de kenmerken van de organisatie; in elke organisatietheorie die dergelijke kenmerken beschrijft ligt reeds een bepaald normatief idee van de organisatie besloten, soms expliciet, soms impliciet.

Ik dien kortom een andere aanvliegroute te gebruiken. En dat kan volgens mij ook: als in de formulering van gangbare organisatietheorieën reeds een bepaald normatief idee van de organisatie verondersteld wordt, dan is het wellicht mogelijk om op basis van die organisatietheorieën te zoeken naar dat veronderstelde normatief idee van de organisatie dat daaronder ligt. Dat is wat ik hieronder zal doen. Daarbij stel ik voornamelijk de ideeën van W. Richard Scott centraal, enerzijds omdat hij zeer invloedrijk is in het denken over organisaties, en anderzijds omdat hij zich juist baseert op de belangrijkste organisatietheorieën zoals die zich ontwikkeld hebben sinds men überhaupt begon te denken over organisaties.

3.3. Het normatieve idee van de organisatie volgens Scott

W. Richard Scott geeft aan dat de meeste organisatieanalisten organisaties zien als “*social structures created by individuals to support the collaborative pursuit of specified goals*” (Scott, 1981/2003, p. 11). Gegeven deze conceptie, zien alle organisaties zich geconfronteerd met een aantal algemene problemen: elke organisatie moet haar doelen definiëren (en herdefiniëren); elke organisatie moet participanten aantrekken om aan het bereik van die doelen bij te dragen; elke organisatie moet dergelijke bijdragen coördineren en beheersen; bronnen moeten vergaard worden uit de omgeving en producten of diensten geleverd; organisatierepresentanten moeten geselecteerd, getraind en vervangen worden; en er dienen bepaalde verstandhoudingen met betrokkenen in de omgeving bereikt te worden (Scott, 1981/2003, p. 11).

Dit veronderstelt volgens Scott dat alle organisaties, hoe verschillend ze ook kunnen zijn, in de basis een aantal centrale eigenschappen of kenmerken met elkaar gemeen hebben. Hij tracht deze te groeperen in het volgende model, ontleend aan Leavitt (Leavitt, 1965, p. 1145):



Figuur 1 – Elementen van een organisatie (Scott, 1981/2003, p. 18)

Scott geeft aan dat Leavitt oorspronkelijk vier “interne” elementen van de organisatie identificeerde, maar daarbij de *organisatie-omgeving* niet als afzonderlijke factor opnam. Scott beziet de omgeving van de organisatie echter als “*an indispensable ingredient in the analysis of organizations, and one that reframes all the other elements.*” (Scott, 1981/2003, p. 18) De organisatie kan niet los van haar omgeving gedacht worden. Dit blijkt ook uit de wijze waarop hij elk element van de organisatie verder uitwerkt.

3.3.1. Technologie

Een organisatie is volgens Scott een plaats waar een bepaald type werk wordt verricht, waar energie wordt aangewend om bepaalde materiële of immateriële bronnen te verwerken tot iets anders, een mechanisme waar inputs worden verwerkt tot outputs. Bij sommige organisaties gaat het om de verwerking van materiële input om een bepaalde materiële output te fabriceren; andere organisaties ‘verwerken’ mensen, bijvoorbeeld een vergroting van hun kennis en expertise in onderwijsinstellingen, of een herstel van hun gezondheid in ziekenhuizen; weer andere organisaties verwerken meer symbolische materialen, zoals informatie en muziek. Maar elke organisatie bezit een bepaalde technologie om dat te doen, een bepaalde set aan machines, kennis en vaardigheden die nodig zijn om bepaalde input tot bepaalde output te verwerken.

In die zin hanteert Scott een normatief idee van de organisatie als een *instrument* dat is ingericht om dat specifieke verwerkingsproces zo optimaal mogelijk uit te voeren, waarbij de omgeving in dit verband wordt verondersteld als zowel de *bron* van de input die door middel van de technologie wordt verwerkt tot output, als de *ontvanger* van die output, daar producten verkocht moeten worden, diensten geleverd, en eventueel afval moet worden vernietigd of opgeslagen.

3.3.2. Sociale structuur

De sociale structuur verwijst volgens Scott naar de gereguleerde aspecten van de relaties tussen de participanten binnen een organisatie. Hij benoemt drie componenten van een dergelijke sociale structuur. De *normatieve* structuur omvat normen, waarden en rolverwachtingen die het gedrag van de participanten reguleren. De *cultureel-cognitieve* structuur is een meer symbolische orde die een kennisraamwerk biedt waarbinnen participanten collectief hun wereld kunnen begrijpen. De *gedrags*structuur laat het feitelijke gedrag zien binnen de organisatie in plaats van dat wat normatief wordt voorgeschreven en dat wat wordt begeleid door cognitieve patronen.

Volgens Scott verhouden in elke sociale structuur de drie structuren zich voortdurend op een dynamische voet tot elkaar: de normatieve structuur legt regels en beperkingen op aan de gedragsstructuur, waardoor het gedrag wordt gevormd en gekanaliseerd. De cultureel-cognitieve structuur biedt een algemeen raamwerk dat helpt om verantwoordelijkheid te nemen voor de gedragspatronen die gevolgd worden. Maar veel gedrag wijkt af van dergelijke modellen en richtlijnen, en zulke afwijkingen zijn weer een bron voor aanvullingen op en veranderingen van de structuur. Gedrag beïnvloedt normen en overtuigingen net zoals normen en overtuigingen gedrag beïnvloeden. Daarbij is ook de omgeving van invloed op de sociale structuur, wat blijkt uit Scotts constatering dat de sociale structuur varieert naargelang de mate waarin het is geformaliseerd: binnen een formele sociale structuur zijn de sociale posities en onderlinge relaties expliciet gespecificeerd en

gedefinieerd, onafhankelijk van de persoonlijke eigenschappen en relaties tussen de participanten die deze posities bezetten. In een informele sociale structuur echter verdwijnt dat onderscheid.

Het concept van de sociale structuur verwijst daarom volgens Scott niet zozeer naar een statische, maar naar een dynamische toestand, in de zin dat een sociale structuur enkel kan bestaan voor zover de participanten met hun gedrag haar blijven produceren en reproduceren. De sociale structuur van een organisatie wordt als het ware gevormd of geconstitueerd door het gedrag van haar participanten: “[w]ithout the ongoing participation of specific individual actors, there is no social structure, no organization.” (Scott, 1981/2003, p. 21). Tegelijkertijd wordt het gedrag van de participanten beïnvloed door die structuur die zij zelf vormen. Een sociale structuur is niet zozeer een begrenzing waarbinnen sociale actoren handelen, maar zowel een *resultaat van* als een *instrument voor* dat handelen van sociale actoren⁴.

3.3.3. Participanten

Scott blijft enigszins onduidelijk in het duiden van wie en wat die organisatieparticipanten nou eigenlijk zijn. Volgens hem zijn organisatieparticipanten “[...] those individuals who, in return for a variety of inducements, make contributions to the organization.” (Scott, 1981/2003, p. 21) Zo gezien is elke klant, elke leverancier, elke werknemer, kortom iedereen die iets bijdraagt aan de organisatie in ruil voor iets anders, een organisatieparticipant. Scott signaleert dat zij vaak gezamenlijk worden geduid als de *stakeholders* van een organisatie, om te benadrukken dat vele partijen beïnvloed worden door en legitieme claims hebben op de organisatie. Hij sluit hierbij aan op de wijze waarop Ancona en collega’s het concept van het stakeholder-model formuleren: “The stakeholder model of the organizational environment extends the political perspective’s focus on power, interests, influence, coalition-building and negotiation beyond the formal boundaries of the organization to provide a way to assess the environment’s influence – and political influence – on the organization.” (Ancona, Kochan, Scully, Maanen, & Westney, 1996, p. 11)

Hiermee lijkt Scott een schimmig onderscheid te maken tussen een organisatieparticipant *als deelnemer* aan de organisatie en een stakeholder als iemand die geen deel van de organisatie uitmaakt, maar er wel bepaalde belangen bij of invloed op de organisatie heeft. Scott biedt geen *criterium* op basis waarvan dat onderscheid gemaakt kan worden, omdat de mate van incorporatie (de mate waarin een actor als deel van de organisatie beschouwd moet worden) volgens hem afhankelijk is van het *type* organisatie en haar functie of rol. Er is wat dat betreft volgens hem bijvoorbeeld een verschil te duiden tussen een priester binnen een geloofsgemeenschap en een toevallige klant in een supermarkt. Een geloofsgemeenschap zal meer facetten van (het gedrag van) de priester als deel van de organisatie beschouwen dan een supermarkt dat van de klant doet, volgens Scott omdat het type organisatie anders is. Maar hier lijkt Scott wel een criterium te *veronderstellen* om het onderscheid te kunnen hanteren, namelijk de vraag in hoeverre van iemand gezegd kan worden *namens* de organisatie te handelen. Een klant die iets koopt in de supermarkt handelt niet *namens* die supermarkt, maar heeft wel invloed op het functioneren van de supermarkt. En dat van de priester meerdere facetten worden geïncorporeerd, is niet zozeer op grond van het type organisatie, maar op grond van dat hij *namens* de geloofsgemeenschap handelt; de priester *representeert* de geloofsgemeenschap.

In het denken over de organisatie als normatief idee is het daarom wat mij betreft van belang om niet zozeer over participanten te spreken, maar over *organisatierepresentanten*, die zich nadrukkelijk onderscheiden van andere stakeholders omdat de organisatierepresentanten *namens* de organisatie handelen, en zich daarmee “within the formal boundaries of the organization” bevinden.

⁴ Scott baseert dit o.a. op Giddens: “Every process of action is a production of something new, a fresh act; but at the same time all action exists in continuity with the past, which supplies the means of its initiation. Structure thus is not to be conceptualized as a barrier to action, but as essentially involved in its production, even in the most radical processes of social change.” (Giddens, 1979, p. 70)

3.3.4. Doelen

Scott geeft aan dat, hoewel elke organisatie doelen nastreeft, de wijze waarop organisatiedoelen kunnen worden gebruikt, enorm kunnen verschillen. Grof gezegd zijn er volgens hem vijf functies van organisatiedoelen, die niet met elkaar samenvallen, niet alleen omdat ze verschillende doeleinden dienen, maar ook omdat ze voortkomen uit verschillende bronnen en niveaus binnen de organisatie.

Allereerst is er de *cognitieve* functie van organisatiedoelen. Doelen leveren criteria voor het genereren en selecteren van handelingswijzen. Ze geven een richting aan en een beperkend kader voor de te nemen beslissingen en te ondernemen acties. Doelen kunnen ook een *motiverende* functie vervullen. Ze dienen dan als een bron voor identificatie en motivatie voor organisatievertegenwoordigers en andere stakeholders. Organisatiedoelen kunnen ook een *symbolische* functie vervullen. Waar de cognitieve en motiverende eigenschappen van doelen hun effecten op organisatieparticipanten benadrukken, wijst het symbolische aspect van doelen naar de betekenis en meerwaarde van de organisatie voor buitenstaanders. Denk aan publiek, belastingbetalers, belangengroeperingen, et cetera. De symbolische doelen die een organisatie zegt te dienen, hebben belangrijke effecten op het vermogen van de organisatie om legitimiteit te verwerven, bondgenootschappen te creëren, bronnen te vergaren en personeel aan te trekken. Organisatiedoelen kunnen ook dienen om reeds verrichtte acties achteraf te *rechtvaardigen*. Gedrag wordt dan niet zozeer gestuurd door de doelen, maar gaat eraan vooraf, en doelen worden achteraf geconstrueerd om het gedrag op acceptabele wijze te verantwoorden. Tot slot kunnen organisatiedoelen aangewend worden om het gedrag van participanten of van de gehele organisatie te *evalueren*; ze leveren dan criteria voor het identificeren en het op waarde schatten van geselecteerde aspecten van het functioneren van de organisatie. Dergelijke criteria om het gedrag te evalueren kunnen al dan niet overeenkomen met de criteria om het gedrag te sturen.

Scott geeft voorts aan dat in het denken over wat organisaties willen bereiken vaak de voorkeur gegeven aan het denken in termen van de *strategie* van de organisatie. Hij haalt hierbij Chandler aan, die *strategie* definieert als “*the determination of the basic long-range goals and objectives of an enterprise, and the adoption of courses of action and the allocation of resources necessary for carrying out these goals.*” (Chandler, 1962, p. 13) Deze definitie identificeert volgens Scott verschillende onderscheidende kenmerken van strategie als concept, waarvan volgens hem de belangrijkste is dat de primaire focus ligt op *externe aangelegenheden*: de verbinding van de organisatie met haar omgeving. (Scott, 1981/2003, pp. 293-294).

Het normatief idee van de organisatie dat Scott hier lijkt te veronderstellen is dat de organisatie als het ware als instrument gericht is op het onttrekken van bepaalde bronnen aan (een deel van) de samenleving om deze op een bepaalde manier te herverdelen en te verwerken, *teneinde iets anders aan de samenleving bij te kunnen dragen*. Een geformuleerde strategie tracht dat te expliciteren; symbolische doelen trachten daarvan de legitieme waarde en het belang te laten zien; en cognitieve, motiverende, evaluatieve en rechtvaardigende doelen zijn er als het ware verdere operationalisaties van, gericht op het realiseren van de bijdrage.

De organisatie als normatief idee is een door individuen geconstrueerd instrument. En iets kan alleen een instrument zijn als het iets doet wat door de gebruiker als nuttig wordt beschouwd. Dat betekent de organisatie niet gedacht kan worden los van een bepaalde mate van legitimiteit in de samenleving, die als het ware aangeeft dat het een nuttig ding is. In de bijdrage die de organisatie levert dient dus op de een of andere manier aangesloten te worden bij bepaalde waarden die in de samenleving gelden; doet ze dat niet, dan *is* er geen bijdrage of nut, en dus geen organisatie.

3.3.5. Meer over de totstandkoming van organisatiedoelen

Scott maakt duidelijk dat, hoewel alle organisaties bepaalde doelen nastreven, daarmee niet is gezegd dat ook alle organisatievertegenwoordigers diezelfde organisatiedoelen nastreven. Er valt in plaats daarvan juist een belangrijk onderscheid te duiden tussen de doelen van de organisatie en de individuele doelen van de organisatievertegenwoordiger. Individuele doelen van de organisatievertegenwoordiger begeleiden de beslissing van het individu om de organisatie te vertegenwoordigen, en om dat al dan niet te blijven doen; van organisatiedoelen wordt verwacht dat ze de beslissingen en het gedrag van de individuen *als organisatievertegenwoordiger* begeleiden. Dat

relatief heldere (theoretische) onderscheid gaat in de praktijk echter nauwelijks op. In veel organisaties introduceren organisatierepresentanten ook hun eigen doelen, belangen en voorkeuren en laten deze mede van invloed zijn op hun beslissingen en gedrag als organisatierepresentant. Organisiatierepresentanten laten zich kortom niet altijd leiden door de doelen van de organisatie.

Scott introduceert het concept van de *dominante coalitie* (Scott, 1981/2003, pp. 296-303) om te verklaren dat organisatiedoelen niet met de individuele doelen van organisatierepresentanten samenvallen, zonder dat we de organisatie als een verstoffelijkte doelgerichte actor moeten zien met antropomorfe eigenschappen. De organisatie wordt daarmee gezien als samengesteld uit coalities: groepen van individuen die dezelfde doelen nastreven of belangen hebben. Elke groep tracht de eigen voorkeuren en doelen op het grotere systeem op te leggen en onderhandelt daartoe met andere groepen. Al deze groepen maken samen deel uit van de dominante coalitie, en elke onderhandelde overeenstemming binnen de coalitie geeft richting aan de organisatie en geeft beperkingen aan dat wat als acceptabele handelingswijzen beschouwd kan worden. Zo komen organisatiedoelen tot stand als zijnde “*complex preference statements that summarize the multiple conditions any acceptable choice must satisfy*” (Scott, 1981/2003, p. 296). De organisatiedoelen zijn daarmee verschillend ten opzichte van de doelen van haar organisatierepresentanten. Ze zijn ook geen statisch gegeven, maar voortdurend onderhevig aan nieuwe invloeden, daar er voortdurend nieuwe belangen en doelen ontstaan in de interacties tussen individuen en groepen, niet alleen binnen de organisatie, maar ook daarbuiten; de dominante coalitie kan ook de belangen en doelen van stakeholders omvatten die buiten de formele grenzen van de organisatie liggen.

Op het eerste gezicht lijkt het gebruik van Scott van de dominante coalitie enigszins vreemd, omdat het resultaat van de dominante coalitie uiteindelijk een resultante is van een onderhandelings- of machtsstrijd. De vraag is dan hoe een organisatie, gedacht als instrument dat bijdragen aan de samenleving levert, doelen kan nastreven die resultaat zijn van een dergelijke onderhandelingsstrijd. De veronderstelling lijkt hier te zijn dat de geformuleerde doelen die vanuit een dominante coalitie voortvloeien, pas daadwerkelijk *organisatiedoelen* kunnen zijn wanneer ze een legitimiteit hebben in de omgeving waarin ze een instrument wil zijn. Immers, als die legitimiteit ontbreekt, als die doelen niet van een bepaald nut of een bepaalde bijdrage aan de samenleving blijken te geven, dan is het resultaat van de dominante coalitie niets anders dan particularistisch willen, en kan het dus niet de organisatie *als organisatie* zijn die die doelen nastreeft. Anders gezegd: een dominante coalitie zou niet eens een dominante coalitie *kunnen* zijn wanneer haar geformuleerde doelen geen legitimiteit kennen, dus wanneer haar doelen geen betrekking hebben op het zijn van de organisatie als instrument dat van nut is voor de samenleving.

3.3.6. *Omgeving*

Volgens Scott bestaat elke organisatie in een specifieke fysieke, technologische, culturele en sociale omgeving waar het zich op de een of andere manier naar *vormt*. Geen enkele organisatie is zelfvoorzienend, maar afhankelijk van de type relaties die het aangaat met grotere systemen waar ze een onderdeel van uitmaakt. De omgeving is daarmee meer dan een afzonderlijke categorie aan variabelen, maar een alomtegenwoordige invloed. En tegelijkertijd is het de omgeving die als het ware een fiat geeft aan de organisatie, omdat de organisatie zelf ook voortdurend invloed op die omgeving uitoefent, via bijvoorbeeld het onttrekken van bepaalde bronnen, het leveren van de beoogde bijdragen, of de productie van afval of een bepaald (al dan niet negatief) neveneffect: de doelen die de organisatie nastreeft dienen daarom, vanuit het perspectief van de samenleving, blijken te geven van een bepaald nut voor de samenleving. En dit betekent dat de organisatie aansluiting zal moeten vinden bij bepaalde waarden in de samenleving, wil het als organisatie haar functie kunnen verrichten.

En hoewel Scott hier min of meer empirische uitspraken doet over de relatie tussen organisatie en omgeving, lijkt daar de veronderstelling onder te liggen dat het de organisatie als *idee* is, dat niet gedacht kan worden los van haar omgeving. De organisatie als normatief idee zoals Scott hanteert is niet zomaar een instrument dat iets bijdraagt aan de samenleving, het is altijd een in de samenleving *ingebed* instrument; de

organisatie dient zich op een bepaalde manier te *vormen* naar de samenleving om er een bepaalde bijdrage aan te kunnen leveren. Zonder die inbedding kunnen we de organisatie helemaal niet denken.

Daarmee geldt voorts dat de organisatie, vanwege de voortdurende invloed op de samenleving, nooit gedacht worden kan als een *neutraal* instrument; in het onttrekken van bepaalde bronnen aan de samenleving, in het nastreven en realiseren van een bepaalde bijdrage aan de samenleving, en in de eventuele al dan niet voorziene neveneffecten die daarbij optreden, is de organisatie niet louter een neutraal instrument, maar is het ten opzichte van die samenleving *normatief geladen*; ze kan het goed of niet goed doen. De organisatie zou vanuit het perspectief van de samenleving bijvoorbeeld de verkeerde bronnen kunnen onttrekken, ongewilde bijdragen kunnen leveren en ongewenste neveneffecten kunnen produceren. De organisatie is dan niet goed *als organisatie*, wat blijkt geeft van een normatief idee van de organisatie die als instrument de *taak* of *functie* heeft datgene te doen wat door de samenleving gewaardeerd wordt.

3.3.7. Scotts normatief idee van de organisatie

Onder Scotts beschrijvingen van de onderscheidende kenmerken van organisaties ligt een normatief idee van de organisatie besloten, en het is dat idee waarvan ik geprobeerd heb de bouwstenen stap voor stap te expliciteren, zoekend en redenerend vanuit de (empirische) beschrijvingen van de elementen van de organisatie. Wanneer we die afzonderlijke bouwstenen per element samenvoegen, komen we tot een normatief idee van de organisatie waarin de organisatie wordt gedacht als een door meerdere individuen gevormd *instrument*, dat, *omdat* het een instrument is, *moet* streven naar bepaalde *bijdragen* aan de samenleving en zich zodoende dient *in te bedden* in de samenleving. De individuen organiseren zich in – of constitueren zich tot – een bepaalde *structuur*, waarin het handelen gericht op het realiseren van de *doelen* op een bepaalde manier wordt gecoördineerd; ze hanteren een bepaalde *technologie* om die doelen zo optimaal mogelijk te realiseren en ze formuleren in een bepaalde *interactie* met elkaar en met de omgeving die doelen die ze met de organisatie willen nastreven, zodanig dat ze een *legitimiteit* binnen de samenleving hebben. De organisatie kan daarom niet gedacht worden als een neutraal instrument, in die zin dat de organisatie zich als instrument altijd op een bepaalde manier *vormt naar* – of inbedt in – de samenleving, en haar invloed op de samenleving afstemt op dat wat in de samenleving als waardevol wordt geacht.⁵

Het is dit normatieve idee van de organisatie dat ik verder zal hanteren in het formuleren van het voorstel voor denken over integriteit bij de organisatie. Dit betekent dat waar ik in het vervolg over de organisatie spreek, ik over het normatieve idee van de organisatie spreek, en niet over de daadwerkelijk bestaande organisaties. Sterker, wanneer we bestaande organisatieverbanden (die we in het dagelijkse taalgebruik organisaties noemen) zouden afzetten tegen het normatieve idee van de organisatie, zal wellicht blijken dat veel van die verbanden helemaal geen organisatie zijn conform het normatieve idee van de organisatie.

3.4. Van normatief idee van de organisatie naar de functie van de organisatie

Als ik bovenstaand normatief idee van de organisatie hanteer, en de organisatie denk als een in de samenleving ingebed instrument teneinde iets aan de samenleving bij te dragen, dan kan ik daarmee bij de functie of *telos* van de organisatie komen.

Laat ik om dit te beargumenteren nog even terug gaan naar de wijze waarop Korsgaard de teleologie van Aristoteles gebruikt om aan te tonen dat constitutieve principes als normatieve principes daadwerkelijk autoriteit hebben: Korsgaard laat zien dat volgens Aristoteles datgene wat een object tot een object maakt, of datgene wat een object zijn *identiteit* geeft, datgene is wat het *doet*: de functie of *telos* van het object. Dit is duidelijk te zien bij gebruiksvoorwerpen, waarbij de functie helder is gedefinieerd en het materiaal ervan (de verschillende delen) zodanig is gearrangeerd dat het die functie kan vervullen. Het zijn van een object betekent op een bepaalde manier *teleologisch georganiseerd* zijn, en *als zodanig een eenheid* zijn, in plaats van een *'mere heap'* van materialen. En dit betekent dat de teleologische organisatie van een object zich

⁵ Dit normatief idee lijkt zo vooral op te gaan voor *publieke* organisaties, die nadrukkelijk instrumenteel moeten zijn ten behoeve van de samenleving, maar zoals in 3.5. zal blijken, geldt dit evengoed voor *private* organisaties.

conformeert aan bepaalde normatieve standaarden, op basis waarvan gezegd kan worden dat het object zijn functie kan verrichten. En daar waar het niet objecten, maar activiteiten betreft, kunnen het normatieve *principes* worden genoemd, die gevolgd moeten worden wil er überhaupt van activiteit sprake zijn.

Als we dit teleologisch perspectief toepassen op het idee van de organisatie, dan kunnen we stellen dat als het de functie of *telos* van de organisatie is dat de normatieve standaarden of principes levert op basis waarvan we kunnen vaststellen of iets wel of niet een organisatie is, en het hierboven vastgestelde normatieve idee van de organisatie levert die normatieve standaarden of principes, dan *is* het normatieve idee van de organisatie de functie of *telos* van de organisatie.

En dit betekent dus dat de organisatie zodanig *teleologisch georganiseerd* moet zijn, (en *als zodanig een eenheid*), dat het zich conformeert aan de normatieve standaarden of principes op basis waarvan gezegd kan worden dat de organisatie functioneert als een in de samenleving ingebed instrument, in de vorm van een doelgerichte sociale structuur, belichaamd door individuen, dat bepaalde bronnen aan de samenleving onttrekt en verwerkt teneinde iets aan de samenleving bij te dragen.

3.5. *Een contingent normatief principe: streef naar inbedding in de samenleving*

Veel (met name) economen hebben een specifieke opvatting over de wijze waarop of mate waarin de organisatie zich dient in te bedden in de samenleving. Zo claimt bijvoorbeeld Milton Friedman dat bedrijven in een kapitalistisch systeem enkel bestaan om winst te genereren voor hun investeerders: *"there is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud."* (Friedman, 1970)

Maar daarmee laat hij juist zien dat – ook al streeft een organisatie enkel naar (meer) winst – er altijd sprake moet zijn van een bepaalde mate van inbedding van de organisatie in de samenleving. Een op winst geïntendeerde organisatie als structuur maakt altijd onderdeel uit van een grotere (kapitalistische) structuur, en vervult daarin een bepaalde functie; in minimale zin is dat het in stand houden en bekrachtigen van die structuur, en in andere zin bijvoorbeeld het scheppen van werkgelegenheid, het bijdragen aan de economische welvaart, het produceren van producten die in de samenleving gewenst worden. Friedman hanteert de zogenaamde 'onzichtbare hand benadering' vanuit de overtuiging dat het kapitalistische vrije marktsysteem in zichzelf alle waarden herbergt om van waarde te zijn voor de samenleving: *"In an ideal free market resting on private property, no individual can coerce any other, all cooperation is voluntary, all parties to such cooperation benefit or they need not participate. There are no values, no "social" responsibilities in any sense other than the shared values and responsibilities of individuals."* (Friedman, 1970) Streven naar winst in een vrije markt rechtvaardigt in die zin volgens Friedman alles, zolang het binnen de kaders van de wet en de basale ethische normen van de samenleving plaatsvindt. Bovendien geldt dat *succesvol* streven naar winst een organisatie alleen maar kan doen *omdat* het op een bepaalde manier aansluit op waarden die in de samenleving gelden. Zoals we bij Korsgaard al zagen, het is immers nooit *enkel het doel* van een handeling dat op waarde wordt geschat, maar altijd de actie als geheel, dus de *handeling-ter-realiseren-van-een-doel*. En juist daarin *kan* een organisatie niet anders dan een bepaalde mate van legitimiteit binnen de samenleving zoeken. Een samenleving zal hooguit de schouders ophalen als een bedrijf zegt naar winst te streven, maar zal protesteren als het bedrijf zegt naar winst te streven door de leden van de samenleving uit te buiten. In die zin heeft de organisatie de taak zich op een bepaalde manier te vormen naar de samenleving.

Uit de functie van de organisatie volgt kortom dat de *praktische identiteit* van de organisatie altijd betrekking moet hebben op de wijze van inbedding in de samenleving; de *vorm* van de praktische identiteit van de organisatie moet zodanig zijn dat er sprake is van inbedding in de samenleving. En volgens Korsgaard is een praktische identiteit een bron voor redenen voor bepaalde acties: een bron voor aansporingen en voor principes op basis waarvan we bepaalde voorstellen voor acties accepteren of verwerpen. Vanuit een praktische identiteit kunnen als het ware contingente vormen van plicht en deugd gedefinieerd worden. En dit betekent dat, als uit het normatieve idee van de organisatie volgt dat de praktische identiteit van de organisatie altijd betrekking moet hebben op een inbedding in de samenleving, daaruit een (contingent)

constitutief en dus normatief principe volgt wat de organisatie zichzelf dient op te leggen wil ze überhaupt een organisatie zijn, namelijk: *streef naar inbedding in de samenleving*.

Dus waar ik zocht naar een normatief idee van de organisatie om te komen tot het idee van zelf-constitutie als grond voor integriteit van de organisatie, en dus de hypothetische en categorische imperatief als constitutieve en normatieve principes, blijkt nu dat ik op basis van dat normatieve idee vooraleerst op een derde principe stuit.

Dit is geen bezwaar. Enerzijds omdat dit principe dezelfde werking heeft als de categorische imperatief ('ik kan dat niet doen, want dat zou verkeerd zijn' is in die zin van dezelfde structuur als 'ik kan dat niet doen, want daarmee kan ik me niet inbedden in de samenleving'). Anderzijds omdat het principe als zodanig *geïntegreerd* kan worden in de hypothetische imperatief: voor zover we de hypothetische imperatief van toepassing laten zijn op acties namens de organisatie, geldt het principe van de inbedding in de samenleving als toegevoegde standaard: de organisatie(representant) toetst de waarde van een middel in relatie tot een doel niet alleen op grond van doeltreffendheid, maar ook op grond van het daarmee bijdragen aan de inbedding in de samenleving.

Derhalve, waar ik in het vervolg spreek over toepassing van de hypothetische imperatief op de organisatie, veronderstel ik daarbij de integratie van het principe van de inbedding in de samenleving.

3.6. *Acties namens de organisatie constitueren de organisatie tot eenheid*

Mijn claim is dat, als de functie van een actie de zelf-constitutie van de actor is, dan is de functie van een *actie namens de organisatie*, de constitutie van de *organisatie als actor*. Acties namens de organisatie hebben op dezelfde wijze als bij personen en dieren de functie om de actor te constitueren, in het licht van zijn functie als actor. En waar de functie van persoon-zijn is het zijn van een persoon, is de functie van het zijn van een organisatie het zijn van een in de samenleving ingebed instrument dat daar iets aan bijdraagt. Dit betekent dat acties namens de organisatie de organisatie constitueren tot een in de samenleving ingebed instrument dat iets aan de samenleving bijdraagt.

Een moeilijkheid hierbij is dat het altijd *individuen* zijn die de acties namens de organisatie verrichten. Toch is het juist dat handelen namens de organisatie – *als organisatierepresentant* – dat aanknopingspunten biedt om te stellen dat acties namens de organisatie constituerend zijn voor de *organisatie* als actor. Daartoe moet echter allereerst vastgesteld worden of er zoiets *is* als handelen namens de organisatie. Kan dat überhaupt gedacht worden? Dat wordt problematisch als we het handelen namens de organisatie interpreteren als een vorm van handelen dat volledig is losgeweekt van het individu dat die handelingen verricht. Maar daar is geen sprake van: een individu 'zet zichzelf niet even in de wacht' zodra hij namens de organisatie handelt, maar meet zich een bepaalde praktische identiteit aan in het handelen namens de organisatie, namelijk de praktische identiteit van organisatierepresentant. Individuen, ook als organisatierepresentant, blijven de individuele taak van zelf-constitutie behouden. Maar het is juist *in* het bevestigen van en handelen conform de waarden en principes ontleend aan die praktische identiteit als organisatierepresentant, dat er ook *namens* de organisatie wordt gehandeld. We mogen er immers vanuit gaan dat de rol van de organisatierepresentant, met de daaruit voortvloeiende waarden en principes, op de een of andere manier ten dienste staat of deel uitmaakt van het realiseren van de organisatiedoelen en van het vervullen van de functie van de organisatie. En dit betekent dat, zodra een individu vanuit zijn praktische identiteit als organisatierepresentant handelt conform de daaruit voortvloeiende waarden en principes, hij als het ware *ook* altijd handelt ten dienste van of onderdeel uitmakend van het realiseren van de doelen van de organisatie en de functie van de organisatie. En als we het handelen als organisatierepresentant, dus het handelen namens de organisatie, zo interpreteren, is het wel degelijk mogelijk te stellen dat er namens de organisatie gehandeld kan worden.

Hoe komen we dan van dit handelen namens de organisatie bij de claim dat acties namens de organisatie *constituerend* zijn voor de organisatie als actor? Daar is een eenvoudig antwoord op te formuleren, met een complexe uitleg. Het antwoord is dat het handelen namens de organisatie en het handelen als constituerend voor de organisatie *hetzelfde* is. In of door het handelen namens de organisatie wordt de organisatie als geheel

geconstitueerd. De uitleg hiervan is complexer: zodra een organisatierepresentant namens de organisatie handelt, dus conform de waarden en principes die voortvloeien uit die specifieke praktische identiteit als organisatierepresentant, is de organisatie als geheel, *als organisatie*, bezig haar functie te vervullen. Dat kan niet anders, omdat het uiteindelijk die functie van de organisatie als geheel is van waaruit de waarden en principes ontstaan voor de acties die namens de organisatie door de organisatierepresentanten verricht worden: alle activiteiten binnen en namens de organisatie, alle rollen, functies, verantwoordelijkheden, doelen, technologieën, besluitvormingsprocedures, protocollen e.d., alles is in principe zodanig gestructureerd, gecoördineerd, geselecteerd, geformuleerd en uitgevoerd – *teleologisch georganiseerd* – dat ze gezamenlijk ertoe bijdragen dat de organisatie als geheel haar functie kan vervullen. En dit betekent dat de waarden en principes die daaruit voortvloeien niet alleen normatief zijn voor de organisatierepresentant, wil hij überhaupt namens de organisatie kunnen handelen, maar *tegelijkertijd* constitutief zijn voor de organisatie als geheel, voor haar praktische identiteit in het licht van haar functie. Immers: conformeert een organisatierepresentant zich niet aan die waarden en principes, dan handelt hij niet namens de organisatie, en is de organisatie als geheel dus niet bezig haar functie te vervullen. Doet hij dat wel, dan *bekrachtigt* hij als het ware zijn rol als organisatierepresentant in het licht van de functie van de organisatie als geheel. Een actie namens de organisatie heeft de *functie* de organisatie als geheel te constitueren.

Een manier om dit zichtbaar te maken is aan de hand van het Constitutioneel Model zoals Korsgaard dat omschrijft. Vanuit dit model kan gedacht worden over hoe acties namens de organisatie de verschillende delen van die organisatie tot een eenheid brengen. Net als de menselijke ziel, en de stad, kan de organisatie gezien worden als een entiteit die uit verschillende delen bestaat, die elk bepaalde eigen belangen en doelen kennen en van daaruit bepaalde aansporingen in de organisatie inbrengen (denk hierbij aan het concept van de dominante coalitie zoals Scott deze verwoordde). En net zoals in de menselijke ziel en in de stad de verschillende delen tot een eenheid gebracht moeten worden om de bewegingen toe te kunnen schrijven aan de actor als auteur ervan, dienen in de organisatie de verschillende delen tot een eenheid gebracht te worden om de bewegingen aan de organisatie *als organisatie* toe te kunnen schrijven als auteur ervan. Een actie is pas een actie namens de organisatie zodra de actie conform de constitutie van de organisatie is verricht, en daarmee constitueert een actie namens de organisatie de organisatie tot eenheid.

3.7. De constitutieve principes van een actie namens de organisatie zijn de imperatieven

Volgens Korsgaard kan er alleen sprake zijn van een actie die aan een actor kan worden toegeschreven, als die actie de actor *doeltreffend* en *autonoom* maakt, wat betekent dat acties de hypothetische en categorische imperatief als normatieve principes dienen te volgen. Als we dat betrekken op de organisatie dan betekent dat dat acties namens de organisatie pas acties namens de organisatie zijn, als de acties de organisatie constitueren tot een doeltreffende en autonome organisatie. En het met de acties bepalen van de *organisatie* als oorzaak van een doel, is niet hetzelfde als door 'iets' of iemand binnen de organisatie bewogen worden, een aansporing, dat als oorzaak functioneert. Acties dienen daarom deliberatief, conform bepaalde constitutieve principes, gekozen te worden, zodanig dat de organisatie als eenheid daar achter kan staan. Dus wanneer de organisatie haar causaliteit bepaalt door zich als eenheid te scharen achter een actie, *moet* zij opereren als een geheel, als iets dat boven haar delen uitstijgt. En om dit te kunnen moet zij zichzelf zodanig inrichten dat ze *kan* delibereren over de te verrichten acties zodanig dat dit tot eenheid leidt. En daartoe dient de *constitutie* van de organisatie, die zegt *hoe* die deliberatie plaats moet vinden, hoe de te verrichten acties gekozen moeten worden. De constitutie van de organisatie levert de normatieve principes op basis waarvan te verrichten acties gekozen kunnen worden. En dat moeten de hypothetische en categorische imperatief zijn.

Elke actie namens de organisatie moet conform de imperatieven gewogen worden op haar waarde als actie-ter-realisatie-van-een-doel. Dat impliceert dat de verschillende delen van de organisatie zich niet moeten identificeren met zichzelf, maar met de constitutie van de organisatie die de principes levert op basis waarvan de organisatie kan handelen en zo haar praktische identiteit als organisatie vervult. De rol of functie van een organisatierepresentant, of van een bepaalde afdeling of deelbelang, kan nooit een doel *op zichzelf*

zijn, maar heeft een inbedding in de gehele organisatie die haar functie als organisatie tracht te vervullen. Er is altijd sprake van een functionele relatie tot andere rollen en deelbelangen, en dat vraagt dat de afzonderlijke aansporingen vanuit dergelijke rollen of deelbelangen nooit als *materieel* principe kunnen dienen voor de keuze om als organisatie tot acties te komen, maar dat ieder deel binnen de organisatie zich dient te identificeren met de *formele* constitutieve principes van de organisatie als geheel. Het enige alternatief zou immers zijn dat de principes op basis waarvan de organisatie kiest, een ander soort aansporing vanuit een van de delen van organisatie is, die sterker op de organisatie inwerkt dan andere aansporingen. De organisatie zou dan zelf meer een soort toeschouwer worden van de interne strijd tussen de aansporingen, en is daarmee niet in staat *de organisatie als eenheid* als de oorzaak te beschouwen van de bewegingen die voortkomen uit de actie.

De noodzakelijkheid van de categorische imperatief voor organisaties dient echter nog nadere aandacht. *Waarom* zou immers een organisatie een eenheid moeten zijn, autonoom moeten zijn? Kunnen er niet ook doeltreffende acties namens de organisatie worden verricht zonder dat deze de organisatie tot autonome actor constitueren? Een deel van het antwoord hierop is reeds gegeven: omdat het een *constitutief* principe is, betekent het niet volgen ervan dat er überhaupt niet van actie sprake is. Maar we hebben ook gezien dat, omdat de individuen namens de organisatie handelen conform waarden en principes die zijn ontleend aan de functie van de organisatie als geheel, zij daarmee de organisatie als geheel, haar *praktische identiteit*, constitueren. En die praktische identiteit was de bron voor het contingente imperatief van de inbedding in de samenleving, mede op basis waarvan de waarde van de acties als handelingen-ter-realiseren-van-een-doel afgewogen moeten worden. En juist die inbedding in de samenleving impliceert dat de organisatiedoelen niet willekeurig gekozen kunnen worden, maar altijd op de een of andere manier in *interactie* met de omgeving tot stand moeten komen. Met het concept van de dominante coalitie liet Scott al zien dat doelen namens de organisatie bemiddeld worden. Maar ze hebben ook altijd een bepaalde *legitimatie* vanuit de samenleving nodig, dus ze zullen op de een of andere manier moeten aansluiten bij bepaalde waarden die in de samenleving gelden. En dat veronderstelt dat de redenen voor acties namens de organisatie *publiek* moeten zijn, in de zin dat er een normatieve kracht vanuit moet gaan die door de samenleving, door alle rationale wezens gedeeld kan worden. De maximes voor de acties namens de organisatie dienen universeel gewild te kunnen worden. Wanneer de organisatie interacteert met de omgeving, maakt ze als het ware samen met die omgeving wetten die hen samen binden. En omdat wanneer *binnen* de organisatie gedelibereerd wordt over te verrichten acties, de organisatie als het ware met zichzelf interacteert, dient ze in dat delibereren over de acties de eigen redenen te beschouwen als publieke redenen met een normatieve kracht.

3.8. *De mogelijkheid tot integriteit bij organisaties op grond van zelf-constitutie*

Integriteit volgens Korsgaard *is* het met het verrichten van acties realiseren van een eenheid. Op dezelfde manier is organisatie-integriteit het (als een in de samenleving ingebed instrument) als eenheid achter de acties staan. Dit vraagt dat acties namens de organisatie conform de imperatieven worden gekozen. En dát vraagt weer dat vanuit de constitutie van de organisatie een normatieve kracht uitgaat op de afzonderlijke delen van de organisatie, zodanig dat zij conform die constitutie delibereren.

De vraag is *hoe* zo'n constitutie van gestalte kan krijgen. Een goede mogelijkheid hiertoe is dat wat Peter French duidt als de "*corporate internal decision (CID) structure*" (French, 1979). CID-Structures bieden volgens French inzicht in hoe verantwoordelijkheden binnen de organisatie gestructureerd zijn, en bepaalde regels waarmee genomen beslissingen zijn te herkennen als beslissingen *om organisatiele redenen*. De CID-Structure belichaamt als het ware een intern mechanisme op basis waarvan de beslissingen die door haar organisatierepresentanten worden genomen geïntegreerd en gestructureerd worden. French zegt daarover:

[T]he CID-Structure accomplishes a subordination and synthesis of the intentions and acts of various biological persons into a corporate decision. When viewed in another way [...], the CID-Structure licenses the descriptive transformation of events, seen under another aspect as the acts of biological persons [...],

to corporate acts by exposing the corporate character of those events. A functioning CID-Structure incorporates acts of biological persons. (French, 1979, p. 212)

De CID-Structure is te beschouwen als de structuur op basis waarvan beslissingen binnen de organisatie geduid kunnen worden als beslissingen *namens* de organisatie, en is als zodanig zowel richtinggevend als regelstellend voor de te nemen acties namens de organisatie. In die zin biedt de CID-Structure goede aanknopingspunten om te laten zien dat de constitutie in feite door een dergelijke structuur gerepresenteerd wordt, en dat het die structuur is die aan bepaalde constitutieve condities moet voldoen willen de acties de organisatie tot integer geheel constitueren⁶.

French duidt de CID-Structure als een set aan constitutieve regels analoog aan de regels van een spel (French, 1979, pp. 212, 213). Spelregels maken duidelijk wat wel en wat niet als actie binnen het kader van het spel geldt. Er gaat in die zin een constitutieve en dus normatieve kracht uit van spelregels. Als die regels niet worden gevolgd, valt het spel uit elkaar. Bovendien geldt dat, zodra iemand bijvoorbeeld aangeeft een partij te willen schaken, diegene in feite aangeeft de regels van het schaakspel te willen volgen. De spelers *identificeren* zich met de spelregels op het moment dat ze het spel spelen; ze laten zich niet leiden door afzonderlijke aansporingen en belangen, maar delibereren conform de regels van het spel, en willen op dat moment acties conform de spelregels verrichten.

Op dezelfde manier fungeert de CID-Structure binnen de organisatie; het levert als het ware de regels of principes die gevolgd moeten worden om überhaupt van de organisatie te kunnen spreken, en om de verrichtte acties te kunnen zien als verricht namens de organisatie en constituerend voor de organisatie. Wie aangeeft om als organisatierepresentant te willen handelen, geeft daarmee in feite aan conform de CID-Structure te willen handelen. En conform de CID-Structure willen handelen betekent vrij letterlijk bezig zijn met het spel, dus bezig zijn met de activiteit van het constitueren van de organisatie tot integere eenheid. Het betekent dus ook dat redenen voor acties geen persoonlijke redenen zijn, of redenen als zijnde aansporingen vanuit bepaalde delen, maar *deliberatieve* redenen, gebaseerd op de principes die de CID-Structure levert, waar de organisatierepresentant zich als organisatierepresentant mee identificeert.

Een bezwaar tegen een dergelijke werking van de CID-Structure, en de analogie ervan met de regels van een spel, is dat de CID-Structure wellicht de regels en principes biedt op basis waarvan gehandeld moet worden om ‘het spel te spelen’, maar dat daarmee nog onbepaald is *welk* spel gespeeld wordt en *waartoe* het spel gespeeld wordt. Dergelijke bezwaren zijn dan ook geuit om French’ betoog van de organisatie als moreel persoon met intentionaliteit te bekritisieren. Zo geeft Donaldson aan dat “*the rules of a game fail to tell us what the game itself intends – in fact, it makes little sense to say that the game intends anything – and one can argue that the same is true for corporations.*” (Donaldson, 1982, p. 22)

Deze bezwaren zijn wellicht houdbaar in het bekritisieren van organisatie-intentionaliteit, maar niet in het licht van zelf-constitutie. Het gaat er niet om dat de organisatie intentioneel moet kunnen handelen, het gaat er om dat de acties de organisatie tot integere eenheid constitueren. Bovendien geldt dat, zodra je wilt vaststellen welk spel gespeeld wordt en waartoe dat spel gespeeld wordt, je in feite al een *materiële* invulling wilt geven aan een CID-Structure, die pas constitutief voor de integere organisatie kan zijn wanneer het *formele* principes levert. Dat neemt echter niet weg dat de praktische identiteit van de organisatie, en het daaruit volgende imperatief van de inbedding in de samenleving, al wel een formele richting geeft aan ‘het spel’, en in die zin aangeeft welk spel waartoe gespeeld wordt.

⁶ French gebruikt de CID-Structure in zijn betoog om te beargumenteren dat een organisatie *intentioneel* kan handelen en zodoende als een moreel *persoon* kan worden beschouwd. Het is echter niet mijn doel aan te tonen dat de organisatie een moreel *persoon* kan zijn met bepaalde *intenties* en dat er daarom een grond bestaat om als organisatie moreel te handelen: het is mijn doel om een grond voor organisatie-integriteit te formuleren, oftewel de organisatie als eenheid, die we als morele actor kunnen denken (in het volgende hoofdstuk zal ik nader ingaan op de plaats van mijn voorstel binnen het debat over de organisatie als morele actor).

3.9. *Eisen aan de constitutie van de integere organisatie*

Wil een CID-Structure daadwerkelijk *constitutief* zijn voor de organisatie als integere eenheid, dan stelt dat eisen aan de inhoud en de vorm ervan, en aan de wijze waarop de organisatierepresentanten zich tot die structuur verhouden, willen zij daadwerkelijk conform die CID-Structure handelen.

Zoals gezegd dient de CID-Structure enkel *formele* principes te leveren wil er een constitutieve werking vanuit gaan die leidt tot een integere eenheid, en geen *materiële* principes⁷. Het dient te borgen dat binnen de organisatie gedelibereerd kan worden over de waarde van voorstellen voor acties op grond van de hypothetische en categorische imperatief, inclusief het principe van de inbedding in de samenleving. Dat vraagt dat, overeenkomstig de *koning-filosofen* in Plato's rechtvaardige stad en de *rede* binnen de individuele ziel, in de CID-Structure een zekere mate van *redelijkheid* geborgd wordt, die het mogelijk maakt dat de maximes van voorgestelde acties getoetst wordt op universaliseerbaarheid.

Die redelijkheid binnen de organisatie krijgt gestalte door te borgen dat een organisatierepresentant begrijpt *waarom* bij zijn of haar functie juist die specifieke taken, verantwoordelijkheden, bevoegdheden en te volgen procedures horen, *waartoe* deze dienen, en hoe die zich verhouden tot de andere functies en tot het geheel. Zodra daar sprake van is, genereert dat als het ware een identificatie met de ingebedde *functie-rationaliteit* op basis waarvan de organisatierepresentant voorstellen voor acties als geheel in het licht van zijn of haar betreffende functie op waarde kan schatten.

In dit verband is niet alleen de identificatie met de eigen functie-rationaliteit van belang, maar ook dat er vanuit de CID-Structure zelf een zelfcorrigerende of *reflectieve werking* uitgaat. De structuur moet borgen dat, zodra er een voorstel tot een actie wordt geopperd, de structuur een mechanisme biedt waarin dat voorstel van feedback wordt voorzien, erop gereflecteerd wordt, zodanig dat daarmee de deliberatie conform de principes gestalte krijgt. Het is als het ware het interacteren van de organisatie met zichzelf. En ook hier geldt weer dat elke organisatierepresentant, vanuit de identificatie met de eigen functie-rationaliteit, *begrijpt* waarom dergelijke feedback, reflectie en correctie vanuit het perspectief van de functie van de organisatie waardevol is, waartoe het dient, en zich er daarom aan conformeert. De organisatierepresentant identificeert zich kortom niet enkel met zijn functie-rationaliteit, maar ook met de CID-Structure *als zodanig* waarbinnen zijn functie een plaats heeft, en waarbinnen zodoende de deliberatie is geborgd.

De organisatieomgeving verandert voortdurend (denk aan nieuwe technologieën, scherpere regelgeving, maatschappelijke ontwikkelingen), dus het als organisatie blijven vervullen van de functie als ingebed instrument vraagt een voortdurende aanpassing aan die omgeving. En de CID-structure dient daarin noodzakelijkerwijs mee te bewegen. De organisatie zo gezien wordt daarmee een soort organisch geheel dat zichzelf voortdurend qua samenstelling vormt - teleologisch organiseert – naar dat wat nodig is om de functie te blijven vervullen. En enkel indien dat proces van zich vormen ook *conform de constitutie* van de organisatie plaatsvindt, is het in staat om zichzelf tot integere eenheid te blijven constitueren.

⁷ Een CID-Structure die bijvoorbeeld enkel en alleen is ingericht om als organisatie financiële winst te realiseren, en als zodanig dat als handelingsprincipe hanteert, is in die zin geen goede CID-Structure. Plato (en met hem Korsgaard) zou een dergelijke constitutie duiden als die van een tiran, die een slaaf is van zijn begeerten, zichzelf laat besturen door de verkeerde wet, en niet in staat is (of beter: omdat hij niet in staat is) te delibereren over de vraag of een actie-ter-realisatie-van-een-doel als geheel het waard is te verrichten; voor de tiran staat dat doel immers al vast, en elke actie daartoe is volgens zijn zelfbesturend principe het waard te verrichten (Plato, 2010, pp. 571-576;321-329).

4. De plaats van de integere organisatie in het debat over de organisatie als morele actor

4.1. *Inleiding*

Met het voorgaande voorstel voor het denken over organisatie-integriteit beargumenteer ik dat dóór moreel te handelen, namelijk conform de categorische en hypothetische imperatief (inclusief het principe van de inbedding in de samenleving), de organisatie geconstitueerd wordt tot een integere organisatie, die als een eenheid achter de acties namens de organisatie staat, en dus als moreel actor gedacht kan worden. Dat heeft consequenties voor het denken over de organisatie in relatie tot moraliteit. In die zin is het van belang te onderzoeken hoe het voorstel zich verhoudt tot het debat daarover. Dat is wat ik hieronder wil doen. Daarbij is het nu niet zozeer mijn doel te beargumenteren dát de organisatie als morele actor gedacht kan worden, maar te onderzoeken in hoeverre de voorgestelde *specifieke manier* waarop de organisatie als morele actor gedacht kan worden, stand houdt in en bijdraagt aan het debat over de organisatie als morele actor.

Wat direct opvalt in dat debat is, zoals we zullen zien, dat zowel voor- als tegenstanders niet altijd consequent zijn in het aanbrengen van een helder onderscheid tussen enerzijds het concept van moreel actorschap bij de organisatie – moraliteit in strikte zin –, en anderzijds het denken over de attributie van morele verantwoordelijkheid aan de organisatie. Beide lopen door elkaar, worden soms samengevoegd, of worden beschouwd als elkaar veronderstellend. Mijn voorstel voor organisatie-integriteit richt zich specifiek op het concept van de organisatie als morele actor, op dezelfde wijze als Korsgaard zich richt op moraliteit in strikte zin, maar gaat niet in op (de consequenties daarvan voor) de attributie van morele verantwoordelijkheid aan de organisatie.

4.2. *Hoofdpijnen van het debat over de organisatie als morele actor*

Het debat over de organisatie als morele actor vindt zijn oorsprong in Amerika, waar gedurende de jaren '70 en '80 verschillende moreelfilosofen begonnen te onderzoeken in hoeverre de organisatie wel of niet als morele actor gekwalificeerd kon worden, wat zij interpreteerden als een entiteit waaraan we op legitieme wijze morele verantwoordelijkheid kunnen toeschrijven (bijvoorbeeld met Ladd (1970); Donaldson (1982); French (1984); Werhane (1985)). Zodra er een gebeurtenis plaatsvindt in de context van het realiseren van organisatiedoelen kunnen we ons immers afvragen wie moreel verantwoordelijk is voor de consequenties daarvan. En waar het weinig controversieel is om te stellen dat organisaties *rechtspersonen* zijn aan wie *legale* verantwoordelijkheden toegeschreven kunnen worden, is de vraag of en in hoeverre organisaties ook *morele* actoren kunnen zijn aan wie *morele* verantwoordelijkheden toeschreven nog niet bevredigend beantwoord.

Toch wordt tegenwoordig binnen de organisatie-ethiek veelal impliciet aangenomen dát organisaties een morele verantwoordelijkheid hebben. Dit komt volgens o.a. Jane Collier (1998) deels door de behoefte om 'de organisatie' te kunnen kwalificeren als een eenheid om haar acties ethisch te kunnen analyseren. En deels komt het door – volgens o.a. Patricia H. Werhane (1989) – de praktische behoefte om, daar waar morele verantwoordelijkheden niet aan de organisatie-representanten toegeschreven kunnen worden, ze toch ergens een plaats te kunnen geven.

Het concept van de organisatie als morele actor impliceert derhalve dat er een onderscheid is tussen de organisatie als morele actor en haar representanten, niet zozeer in de zin dat de organisatie acties kan verrichten zonder haar representanten, maar wel dat de organisatie als een afzonderlijke eenheid moreel verantwoordelijk kan zijn op een manier die zich onderscheidt van haar representanten. Als zodanig dient het concept van de organisatie als morele actor dan ook gescheiden te worden van noties van *collectieve* morele verantwoordelijkheid, waarin morele verantwoordelijkheden wel worden toegeschreven aan organisatie-representanten, maar niet aan de *organisatie* als actor.

Het debat over de morele status van organisaties baseert zich op de aanname dat we alleen morele verantwoordelijkheden kunnen toeschrijven aan een morele actor. Gegeven die aanname zijn voornamelijk pogingen ondernomen om te bezien in hoeverre de organisatie al dan niet voldoet aan de noodzakelijke voorwaarden voor of eigenschappen van moreel actorschap. En in dat verband wordt veelal een vergelijking gemaakt met de noodzakelijke (metafysische) condities voor de *individuele persoon* als morele actor, en onderzocht in hoeverre die ook van toepassing zijn op de organisatie (vaak wordt er daarom ook gesproken over bijvoorbeeld *corporate moral personhood* (Phillips, 1992), of *the corporation as a moral person* (French, 1979)). Deze condities voor moreel actorschap op basis waarvan het debat wordt voortgestuwd, kunnen herleid worden tot de volgende twee *minimaal noodzakelijk* geachte condities waaraan voldaan moet zijn wil er sprake zijn van moreel actorschap:

1. *Er dient sprake te zijn van het vermogen tot het hebben van een intentie om een actie te verrichten*
2. *Er dient sprake te zijn van het vermogen een intentionele actie autonoom te kiezen*

Tegenstanders van het concept van de organisatie als morele actor bepleiten dat, omdat de organisatie niet aan één (of beide) van deze condities voldoet of kan voldoen, daaruit volgt dat de organisatie niet als morele actor gedacht kan worden op eenzelfde manier als individuen als morele actor gedacht kunnen worden. Voorstanders van het concept van de organisatie als morele actor beargumenteren dat de organisatie op de een of andere manier wel aan deze condities kan voldoen. Ook zijn er voorstanders van het concept die beargumenteren dat, hoewel de organisatie wellicht niet aan deze condities kan voldoen, de organisatie toch als morele actor gedacht kan worden, maar op een manier die anders is dan de individuele morele actor; de organisatie wordt geduid als een *speciale soort* morele actor. Het is deze laatste categorie waarbij ik met mijn voorstel aansluiting zoek. Daarbij richt ik me enkel op het concept van de organisatie als morele actor, los van de eventuele consequenties daarvan voor het toeschrijven van morele verantwoordelijkheid.

4.3. *De organisatie moet in staat zijn een intentie te hebben om een actie te verrichten*

Er wordt binnen het debat veelal verondersteld dat het vermogen tot het hebben van een intentie om een actie te verrichten, een noodzakelijke conditie is voor moreel actorschap. Zo claimt Arnold dat “[c]entral to the claim that corporations are moral agents is the claim that corporations have intentions.” (Arnold, 2006, p. 281) Ook argumenten over de bredere notie van collectieve morele verantwoordelijkheid (zie o.a. May (1987) en Pettit (2001)), of over gedeelde intenties (zie o.a. French (1996) en Bratman (1993) (2014)), nemen de noodzakelijke verbinding tussen intentionaliteit en verantwoordelijkheid als een belangrijk uitgangspunt.

French verdedigde oorspronkelijk het concept van intentionaliteit bij organisaties (en de legitimiteit daartoe van de CID-Structure) door het te gronden in de *‘desire-belief’*-theorie van intentionaliteit. Dat zegt in het kort dat als een actor een bepaalde intentie voor E heeft, die actor het verlangen heeft naar E en de overtuiging dat met bepaald gedrag in een bepaalde situatie het geïntendeerde E bereikt kan worden. French beargumenteerde op basis hiervan dat “when the corporate act is consistent with, an instantiation or an implementation of established corporate policy, then it is proper to describe it as having been done by a corporate desire coupled with a corporate belief and so, in other words, as corporate intentional.” (French, 1979, p. 213).

De kritieken hierop bepleitten voornamelijk dat, omdat French specifieke menselijke kwaliteiten aan de organisatie toeschreef die niet over dergelijke menselijke kwaliteiten *kan* beschikken, zijn verdediging van intentionaliteit bij organisaties geen stand houdt. Velazquez verwoordde dit als volgt:

[A]n act is intentional only if it is the carrying out of an intention formed in the mind of the agent whose bodily movements bring about the act. The intentions French attributes to corporations, then, do not mark out corporate acts as intentional because the intentions are attributed to one entity (the corporation) whereas the acts are carried out by another entity (the corporate members) [...] The underlying reason for corporate policies and procedures being unable to generate intentional action is that the concept of

intentional action, as I have suggested, is rooted in the concept of an agent with a certain mental and bodily unity that corporations do not have. (Velasquez, 1983, p. 8)

In zoverre French intentionaliteit bij organisaties dacht op een niet-metaforische, maar antropomorfe wijze, lijkt de kritiek daarop terecht; organisaties zijn nu eenmaal geen personen. French onderkende dat, en zocht daarom naar een manier om de organisatie als morele actor te gronden in een theorie over intentionaliteit die dat probleem omzeilt. In een later werk, *Integrity, Intentions and Corporations* (French, 1996), verwerpt hij dan ook zijn eerdere gebruik van de *'desire-belief'*-theorie over intentionaliteit, en zoekt in plaats daarvan aansluiting bij de door Bratman geformuleerde *'planning theory'* van intentionaliteit.

Wat volgens Bratman karakteristiek is aan intenties kunnen we volgens hem niet achterhalen door te kijken naar een geïsoleerde intentie en een geïsoleerde actie, maar door ze te zien als delen van een grotere structuur, die hij plannen noemt: *"I tried to understand intentions as, in the basic case, elements of larger and typically partial plans whose primary roles in our lives are one of coordination and organization, both cross-temporal and social"*. (Bratman, 2008, p. 29) En omdat intenties onderdeel uitmaken van plannen, zijn ze volgens Bratman onderhevig aan twee rationele normen: consistentie en coherentie. De intenties van een persoon, inclusief zijn overtuigingen, moeten *gezamenlijk* passen in een consistent model van zijn toekomst waar hij zijn plannen aan ontleent. Intenties zijn in die zin geen atomische eenheden die één voor één tot stand komen, maar onderdeel van het algehele systeem van plannen van de actor. Als intenties niet consistent met elkaar zijn, zou dat systeem er überhaupt niet zijn. De actor dient in die zin een eenheid te zijn. En verder dienen de intenties van de persoon middel-doel coherent te zijn, in de zin dat het niet zo kan zijn dat iemand de intentie heeft tot E, ervan overtuigd is dat E vraagt om middel M te intenderen, en toch niet de intentie heeft tot middel M. (dit is vergelijkbaar met de hypothetische imperatief: wie het doel wil, wil de middelen om dat doel te realiseren).

Bratman beargumenteert dat deze twee normen van consistentie en coherentie noodzakelijk werkzaam zijn in het praktisch redeneren van een actor, *omdat* het een plannende actor is, en Bratman wil laten zien dat deze normen ook daadwerkelijk als praktische normen op zichzelf staan (en niet zijn afgeleid van bepaalde theoretische normen). Zo kan een actor volgens Bratman, pragmatisch gezien, onmogelijk stoppen met een plannende actor te zijn (zelfs als dat metafysisch gezien wel mogelijk zou zijn). Bovendien is ons plannende actorschap zodanig organisch verweven met het weefsel van onze levens, zowel individueel als sociaal, dat stoppen met plannen zou betekenen een volledig ander soort leven te leiden. Belangrijkste is echter volgens Bratman dat het plannen onze (cross-temporele) integriteit en zelfbestuur constitueert. Zonder te plannen, en zonder intenties die consistent ten opzichte van elkaar zijn en coherent ten opzichte van de doelen en middelen, zou die integriteit en dat zelfbestuur uit elkaar vallen. Dat legt hij uit in termen van *'agential authority'*, wat zoveel betekent als het vermogen van een actor om zich te identificeren met een bepaalde gedachte of attitude, datgene op basis waarvan we kunnen zeggen dat een gedachte of attitude van de actor is. Bratman zegt daarover:

[N]orms of consistency and coherence of intention are aspects of the capacity of planning structures to solve this problem of agential authority. Agential authority involves a more-or-less unified framework that has that authority; and consistency of intention helps ensure that there is such a single, more-or-less unified framework. Agential authority also involves some degree of actual control of action; and means-end coherence helps ensure, given basic features of our non-divine agency, that our plans for ends do indeed tend effectively to shape our action. (Bratman, 2008, p. 59)

Dit laat nog niet zozeer zien dat dergelijke planstructuren noodzakelijk zijn voor actorschap (wat Bratman ook onderkend), maar wel dat we op basis hiervan onszelf, pragmatisch gezien, als plannende actor kunnen begrijpen. En omdat het zijn van een plannende actor onder andere betekent dat je je moet committeren aan de normen van consistentie en coherentie van je intenties, wordt het *waartoe* van dat commitment duidelijk: als we niet consistent en coherent zijn in onze plannen, dan valt onze integriteit en autonomie uit elkaar.

Het is dit concept van intentionaliteit dat French aanwendt om een antwoord te formuleren op de kritiek dat intenties gedacht als een 'desire-beliëf-structuur mentale toestanden zijn, en daarom niet aan de organisatie kunnen worden toegeschreven. In zijn paper *Integrity, Intentions, and Corporations* tracht hij een organisatie-intentionaliteit te beargumenteren, door intenties gedacht als een planstructuur toe te schrijven aan de CID-Structure. Hij stelt daartoe dat:

If we concentrate on the possible descriptions of events and acknowledge that there are distinctly corporate plans and policies that provide the reasons why corporations do the things they do, then we should not feel compelled to reduce statements about corporate actions to ones about the actions, reasons, plans, or interests of humans who happen to be agents of the corporate actor. [...] [C]orporations make plans, in fact, if anything, they are planning entities, designed to do just that. (French, 1996, p. 152)

Maar volgens critici als Rönnegard komt French hiermee nog steeds niet los van de intentie als mentale toestand, simpelweg omdat ook voor Bratman een intentie, ook al is het een planstructuur, een mentale toestand is. Volgens Rönnegard misbruikt French de planning theory, "because Bratman conceives of intentions as mental plans and because a CID-Structure clearly does not possess any mental states it cannot have any intentions in the sense of Bratman's theory." (Rönnegard, 2015, p. 23)

In plaats van Bratmans concept van *individuele* intentionaliteit is daarom gepoogd het later door Bratman ontwikkelde concept van *gedeelde* intentionaliteit te gebruiken als grond voor organisatie-intentionaliteit, onder andere door Arnold.

Bratman stelt dat het regelmatig voorkomt dat we intenties hebben die in belangrijke mate gedeeld zijn (we hebben de intentie samen een duet te zingen, samen een huis te verven of samen een reis naar New York te maken). Maar hij vraagt zich af wat een gedeelde intentie *is*, en wat de *condities* zijn om van een gedeelde intentie te spreken. Hij stelt enerzijds dat het voor een gedeelde intentie niet genoeg is om te stellen dat ieder individu dezelfde intentie heeft. En anderzijds stelt hij dat een gedeelde intentie niet een attitude is in de geest van een of andere superactor waarin die van de individuen zijn samengesmolten. Ook stelt hij dat wederzijds uitgesproken beloften om samen hetzelfde te intenderen nog niet voldoende is om van een gedeelde intentie te spreken, daar iemand niet oprecht kan zijn en niet de intentie heeft zijn belofte na te komen. Waar het Bratman dan wel om gaat, is dat we gedeelde intenties in de basis moeten begrijpen als "*a state of affairs consisting primarily of appropriate attitudes of each individual participant and their interrelations.*" (Bratman, 1993, p. 99)

Dit werkt hij verder uit door te kijken naar de *functie* van een gedeelde intentie, die hij in drieën splitst: een gedeelde intentie helpt onze intentionele acties te coördineren, helpt onze plannen te coördineren, en kan een raamwerk bieden waarbinnen we kunnen onderhandelen over hoe we ons doel willen bereiken. Bratman beargumenteert vervolgens dat, als individuen attitudes kunnen hebben die over dusdanige inhoud beschikken en op een dusdanige manier naar elkaar verwijzen, dat het complexe geheel van die attitudes de drieledige functie van de gedeelde intentie kan vervullen, dan kunnen we dat complexe geheel identificeren als een gedeelde intentie. Individuen dienen daartoe zowel (1) een gemeenschappelijke intentie te hebben (die consistent en coherent is), als (2) op elkaar ingrijpende (deel)plannen ten behoeve van die intentie en (3) kennis hebben van (1) en (2).

Op basis van dit idee van de gedeelde intentie ontwikkelt Arnold een nieuw concept van organisatie-intentionaliteit. Volgens Arnold leunt dat niet meer op een intentionaliteit als een individuele mentale toestand, maar op een gedeelde intentionaliteit die tot stand komt wanneer de attitudes van individuen binnen de organisatie aan bepaalde voorwaarden voldoen, zowel qua inhoud als qua relatie tot elkaar. En in navolging van French voert Arnold de CID-Structure op om te laten zien dat dit binnen organisaties goed mogelijk is. Niet zozeer door te stellen dat een CID-Structure *garandeert* dat alle actoren binnen de organisatie dezelfde relevante intenties met elkaar delen, maar wel door te stellen dat, omdat een CID-Structure een *prescriptieve* en geen *descriptieve* functie vervult, "[i]t tells the agents of the corporation how they

ought to act.” (French, 1995, p. 31) De CID-Structure zegt kortom dat “[e]mployees should have the appropriate intention, and if they do not they fail to properly represent the corporation.” (Arnold, 2006, p. 289) Er kan in die zin volgens Arnold ook een corrigerende werking uitgaan van de CID-Structure om te borgen dat de relevante actoren de juiste, d.w.z. de relevante gemeenschappelijke intenties hebben. En volgens Arnold kan met de CID-Structure ook aan de andere twee condities van gedeelde intentionaliteit worden voldaan: “[t]he meshing subplans condition is satisfied by the coordination of various subplans in the different departments and divisions of the company [...] The knowledge condition is satisfied insofar as each person is aware of the subplans of other members of the group necessary for the execution of the plan.” (Arnold, 2006, p. 289) Zodoende kunnen we volgens Arnold spreken van organisatie-intentionaliteit, zonder het als mentale toestand aan de organisatie te moeten toeschrijven.

4.4. *De organisatie moet in staat zijn een intentionele actie autonoom te kiezen*

De vraag is echter in hoeverre een dergelijk beargumenteerd vermogen tot organisatie-intentionaliteit nog steeds helpt om de organisatie als morele actor te zien. Voor Arnold lijkt dat voor de hand liggend: “for corporations to be properly regarded as moral agents, a further condition must also be satisfied. Corporations must be capable of reflectively endorsing corporate intentions [...] There can be little doubt that such agents are commonplace.” (Arnold, 2006, p. 291) Maar hij lijkt hier de bocht af te snijden, omdat het juist nog maar de vraag is of een organisatie *als organisatie* op een reflectieve manier de organisatie-intenties kan bekrachtigen. Dubbink en Smith zeggen hier bijvoorbeeld over:

Reflective endorsement implies more than simply identifying and accepting a consideration in favor (or against) a course of action. It implies a unified understanding of why that consideration is valid in light of a plurality of other considerations. Reflective endorsement, in other words, requires an actor to compare and evaluate its considerations in favor of certain actions relative to an overarching practical identity. [...] CID-Structures do not provide a perspective from which corporations can be said to understand their capacity to act on reasons as such. (Dubbink & Smith, 2011, p. 237)

Hier gaat kortom het probleem van de autonomie een rol spelen. Wellicht kunnen we aan de organisatie gedeelde intenties toeschrijven, maar is de organisatie ook in staat om *als organisatie* autonoom een intentionele actie te kiezen? Dat vermogen wordt veelal als een noodzakelijke conditie voor moreel actorschap beschouwd, omdat we pas een actie aan de actor kunnen toeschrijven wanneer die actie *zijn* keuze is, en wanneer het *zijn* keuze is.

Om het concept van autonomie te duiden worden vaak termen gebruikt als ‘vrije wil’, ‘zelfbepaling’ en ‘zelfbestuur’. Deze termen baseren zich op de intuïtie dat autonomie betrekking heeft op het vermogen van de actor om vanuit een bepaald punt, dat zich onderscheidt van zijn directe verlangens, te kunnen kiezen naar welk verlangen de actor zal handelen. Autonomie zo gezien betreft daarmee niet alleen het vermogen te kiezen, maar gaat ook over *hoe* keuzes worden gemaakt. In dat verband wordt vaak Harry Frankfurt aangehaald, die het onderscheid aanbracht tussen ‘first-order intentionality’ en ‘second-order intentionality’ (Frankfurt, 1971). Daarmee liet hij zien dat een autonome actor die bepaalde directe verlangens kent (bijvoorbeeld: “ik heb nu zin om vlees te eten”: een first-order intention), daarop moet kunnen *reflecteren* (bijvoorbeeld: “maar ik wil niet dat daarvoor dieren gedood worden, dus ik verlang ernaar niet naar vlees te verlangen”: een second-order intention), zodat hij op basis van die reflectie tot keuzes voor acties komt. De actor is zodoende in staat onafhankelijk keuzes voor acties te maken, in plaats van louter direct te reageren op zijn directe verlangens. Autonomie als zodanig betreft dan het vermogen van de actor om reflectief bepaalde verlangens te bekrachtigen, en daarmee begeleid te worden door verlangens die daadwerkelijk van hemzelf zijn, en maakt het zo mogelijk dat we de actor als de eigenaar van een intentionele actie kunnen zien.

Onafhankelijk een intentionele actie kiezen zou ook niet mogelijk zijn als het een actor ontbrak aan het vermogen om ook *iets anders* te kiezen. Dit betekent echter niet noodzakelijkerwijs dat er ook daadwerkelijk andere keuzeopties beschikbaar moeten zijn. Frankfurt liet daartoe zien dat, ook al zijn er voor de actor geen

andere keuzeropties beschikbaar, de keuze voor een actie van de actor toch diegene kan zijn die de actor zou kiezen als er wel andere keuzeropties beschikbaar waren. In een dergelijke situatie is de actor toch moreel verantwoordelijk, en is het irrelevant of de alternatieven waaruit de actor had kunnen kiezen ook daadwerkelijk beschikbaar waren (Frankfurt, 1969).

Tegenstanders van het idee van de organisatie als morele actor (bijvoorbeeld Rönnegard (2015)) bepleiten vaak dat de organisatie *als organisatie* niet in staat is tot het reflectief bekrachtigen van bepaalde verlangens, intenties of voorstellen voor acties. Het ontbreekt de organisatie zagezegd aan het vermogen tot second-order intentionality, omdat de organisatie als organisatie zich nooit *bewust* kan zijn van gemaakte keuzes. En dat betekent dat, als dat vermogen bij de organisatie ontbreekt, de organisatie als organisatie niet in staat is autonoom keuzes te maken voor intenties en acties, en zodoende niet de eigenaar kan zijn van de acties die verricht worden.

Philip Pettit betoogt juist dat het *wel* mogelijk is om autonomie aan de organisatie toe te schrijven, door te laten zien dat met bepaalde beslisprocedures binnen een doelgerichte organisatie het denkbaar is dat er een beslissing wordt genomen die anders is dan wat elk lid van die groep verlangde. En het is daarom dat we een doelgerichte organisatie moeten beschouwen als intentionele en autonome actor. Hij zegt:

There is a type of organization found in certain collectivities that makes them into subjects in their own right, giving them a way of being minded that is starkly discontinuous with the mentality of their members. This claim in social ontology is strong enough to ground talk of such collectivities as entities that are psychological autonomous and that constitute institutional persons. (Pettit, 2002, p. 167)

Pettit legt hiertoe een situatie voor waarin een groep werknemers in een fabriek een besluit moeten nemen of ze het wel of niet eens zijn met de betaling van extra kosten om de veiligheid van hun werksituatie te verbeteren. Hierin zijn drie vraagstukken van belang: 1. Bestaat er onder de huidige werkomstandigheden een serieus veiligheidsrisico? 2. Zullen de voorgestelde maatregelen voor verbetering effectief zijn? 3. Zijn de te maken extra kosten draaglijk? Ervan uitgaande dat de groep uit drie werknemers bestaat, legt Pettit de volgende matrix voor:

	<i>Serieus veiligheidsrisico?</i>	<i>Effectieve maatregelen?</i>	<i>Draaglijke kosten?</i>	<i>Betaling van extra kosten?</i>
<i>Werknemer A</i>	<i>Ja</i>	<i>Nee</i>	<i>Ja</i>	<i>Nee</i>
<i>Werknemer B</i>	<i>Nee</i>	<i>Ja</i>	<i>Ja</i>	<i>Nee</i>
<i>Werknemer C</i>	<i>Ja</i>	<i>Ja</i>	<i>Nee</i>	<i>Nee</i>
	<i>Ja</i>	<i>Ja</i>	<i>Ja</i>	<i>???</i>

Pettit laat zien dat het afhankelijk is van de beslisprocedure wat de groep uiteindelijk als groep zal besluiten. Op basis van een conclusie-gebaseerde procedure komt eerst elk individu tot een totaalafweging om wel of niet tot betaling van de extra kosten over te gaan, waarbij alle drie de vraagstukken bevestigend beantwoord moeten zijn wil de werknemer bereid zijn de extra kosten te betalen. Op basis van de meerderheid van die conclusies wordt het uiteindelijke besluit genomen. Dat leidt tot een duidelijk *nee* (verticale kolom aan de rechterzijde). Maar als een vraagstuk-gebaseerde procedure wordt gehanteerd, die uitgaat van de totaalantwoorden op de drie vraagstukken afzonderlijk, waarbij een meerderheid van de antwoorden van de werknemers leidt tot een antwoord namens de groep, dan leidt dat tot een duidelijk *ja* (horizontale rij aan de onderkant). En met deze laatste procedure wordt duidelijk dat het goed mogelijk is dat elk groepslid tegen de betaling van de extra kosten is, maar het collectieve besluit van de groep toch is om de extra kosten te betalen. Dit resultaat noemt Pettit de *'doctrinal paradox'*.

Volgens Pettit wordt hiermee de rede 'gecollectiviseerd'. Het is een *'social integrate'*: de representatie van een integratie van individuen (en niet enkel een optelling), omdat het gedeelde doelen heeft en de rede collectiviseert in het vormen van oordelen en intenties. En een *social integrate* moet gezien worden als een

intentioneel subject omdat “*it is going to display all the functional marks of an intentional subject and that there is no reason to discount those marks as mere appearances.*” (Pettit, 2002, p. 182)

Pettit ziet hierbij de *social integrate* als onderscheidend van haar individuele leden, niet zozeer metafysisch, maar “*in the sense of being a centre for the formation of attitudes that are capable of being quit discontinuous from the attitudes of the members.*” (Pettit, 2002, p. 183) Volgens Pettit zou het onjuist zijn te denken dat de individuen eerst hun intenties vormen waarna deze verworpen tot de intenties van het collectief. In plaats daarvan is het het collectief dat het besluit neemt, op een manier die niet direct samenhangt met de attitudes van de individuele leden. En pas nadat het collectieve besluit is genomen vormen de afzonderlijke leden hun eigen intenties ten opzichte daarvan, om deze om te zetten in actie.

Pettit laat hiermee zien dat het mogelijk is dat een doelgerichte organisatie als collectief tot een besluit kan komen, zich intenties kan vormen, die anders is dan die van haar leden. De vraag is dan of dit voldoende is om de autonomie van de organisatie te beargumenteren.

Rönnegard bijvoorbeeld denkt van niet. Volgens hem is de bepleitte autonomie van Pettit slechts procedureel – beslissingen van de organisatie zijn het resultaat van een procedure – en kan de organisatie op geen enkele manier zich bewust zijn van de keuze die het gemaakt heeft. Het zijn in die zin geen ‘second-order intentions’. Het gegeven dat de te maken keuze is verschoven van de leden naar de procedure, betekent niet dat het *actorschap* is verschoven van de leden naar de procedure: “*All the morally relevant choices are done by devising the procedure itself and by the members accepting to abide by the result of the voting procedure.*” (Rönnegard, 2015, p. 54) Volgens Rönnegard kan een beslisprocedure simpelweg niet moreel verantwoordelijk zijn. We schrijven verantwoordelijkheid toe aan een actor met een intentionele actie, en niet aan de procedure die leidt tot die actie. Bovendien is de keuze om überhaupt die procedure te hanteren het resultaat van intentionele acties van de individuele leden. Een procedure kan wellicht de acties van individuele leden verklaren, zoals de notie van de ‘gecollectiviseerde rede’ kan verklaren waarom het collectief een actie intendeerde die de individuele leden niet verlangden, maar is daartoe niet moreel verantwoordelijk.

4.5. *De integere organisatie als morele actor*

En zo lijken we in een cirkel terecht te komen: als we de organisatie als morele actor willen denken, zo wordt verondersteld, dan dient de organisatie aan bepaalde condities van moreel actorschap te voldoen. Deze condities worden ontleend aan de minimaal noodzakelijke condities van individueel moreel actorschap. Het vermogen tot het hebben van intenties is zo’n conditie. Probleem met intenties is dat ze worden beschouwd als mentale toestanden, en omdat de organisatie als organisatie niet over mentale toestanden kan beschikken, kan de organisatie dus niet over het vermogen beschikken om intenties te hebben. Daarom is gepoogd het gebruik van intenties als mentale toestanden te omzeilen, door organisatie-intentionaliteit te denken in termen van gedeelde intenties. Gedeelde intenties kunnen wel worden toegeschreven aan de organisatie, gebruikmakend van de CID-Structure. Maar, zo wordt nu duidelijk, een gedeelde intentie is niet per se één van de noodzakelijk geachte condities om van moreel actorschap te kunnen spreken, en daar ging het nu juist allemaal om. Als het probleem bij het toeschrijven van gedeelde intenties aan de organisatie is dat het daarmee de organisatie nog ontbreekt aan het vermogen om *als organisatie* intenties of redenen voor acties reflectief te bekrachtigen, dan is het dus niet zo dat de gedeelde intenties een noodzakelijke conditie voor moreel actorschap zijn, maar is moreel actorschap een noodzakelijke conditie voor het *als organisatie* bezitten van gedeelde intenties.

Is er een uitweg? Ik denk het wel, en ik denk dat die uitweg uit de combinatie van twee delen bestaat, beide gebaseerd op het voorstel voor organisatie-integriteit op grond van Korsgaards theorie van zelf-constitutie.

Het eerste deel van de uitweg betreft het probleem van de mentale toestand en de aanname dat de organisatie *als organisatie* niet over mentale toestanden kan beschikken. Aan de hand van Korsgaard wordt duidelijk dat het *niet nodig is* om het vermogen van een actor tot het bezitten van een mentale toestand te beargumenteren, wil die actor een morele actor zijn. Wat *wel* nodig is volgens Korsgaard, is het vermogen om te *delibereren* over een te verrichten actie, conform principes die de actor constitueren tot doeltreffende en autonome eenheid aan wie de actie kan worden toegeschreven. En dat *delibereren* kan gezien worden als

een mentale activiteit, zoals bij Korsgaard wellicht het geval is wanneer ze spreekt over de individuele morele actor, maar daarmee is niet gezegd dat het *enkel en alleen maar* een mentale activiteit kan zijn. Het is niet voor niets dat Korsgaard voortdurend terugkeert bij Plato's analogie van de stad, onder andere namelijk om te laten zien dat het vermogen om te delibereren is wat maakt of een actor tot eenheid geconstitueerd kan worden. Zonder het vermogen te delibereren conform de imperatieven, is autonoom en doeltreffend actorschap, moreel actorschap, niet mogelijk.

En dat vermogen tot delibereren is in de eerste plaats het vermogen tot *interactie*. Korsgaard geeft aan dat “[t]hinking, after all, is just talking to yourself. And since meanings are public, talking is just thinking in the company of others.” (Korsgaard, 2009, p. 197) Daarmee probeert ze duidelijk te maken dat als we willen dat redenen voor acties normatief zijn, dat alleen maar mogelijk is wanneer het *publieke* redenen zijn, analoog aan het argument van Wittgenstein voor het noodzakelijk publieke karakter van de betekenis van taal. Op dezelfde manier als de betekenis die we aan woorden geven niet normatief kan zijn als het niet publiek is (in de zin dat je het niet verkeerd kunt doen), kunnen redenen voor acties niet normatief zijn als ze niet publiek zijn. Dus het delibereren over redenen voor acties is eerst en vooral een *publieke* aangelegenheid, interactie, en niet voorbehouden aan een individuele mentale capaciteit.

Willen we een actor als morele actor kunnen denken, dat dient die actor kortom het vermogen te bezitten om te delibereren over de te verrichten acties, conform de imperatieven. Aangezien de normativiteit van die te verrichten acties alleen maar kan bestaan wanneer ze om publieke redenen worden verricht, kan het niet zo zijn dat daartoe enkel en alleen een individuele mentale toestand noodzakelijk is. Delibereren is interacteren. En hiermee kan met Korsgaard dus de noodzaak (of beter: het probleem) om intenties *als mentale toestand* toe te schrijven aan een actor omzeild worden.

Het tweede deel van de uitweg betreft de noodzaak van het vermogen om intenties en keuzes voor acties reflectief te bekrachtigen. Immers, ook al hoeft er geen sprake te zijn van een mentale toestand om te kunnen delibereren, dan kan dit nog steeds de vraag oproepen door wat of wie op basis van de deliberatie de uiteindelijke keuze voor een te verrichten actie tot stand komt. Blijven we hier niet nog steeds de noodzaak voor actorschap houden, die op reflectieve wijze keuzes voor acties kan bekrachtigen? Nee. Het hele punt van Korsgaard is juist dat dat actorschap *geconstitueerd* wordt door het verrichten van de actie die op basis van deliberatie conform de imperatieven is verkozen. Stellen dat een actorschap noodzakelijk is wil er sprake zijn van het vermogen tot het hebben van (gedeelde) intenties, of tot het delibereren, of tot kiezen en verrichten van acties, is *veronderstellen* dat er eerst een actor als een eenheid moet zijn, die zich op de een of andere manier een bepaalde intentie vormt of afweging maakt, op basis waarvan die actor al dan niet handelt. Maar met Korsgaard hebben we juist gezien dat het de actie zelf is die die eenheid constitueert. En dit gaat op voor zowel individuen als voor organisaties. Als binnen de organisatie de deliberatie over de te verrichten acties aan bepaalde formele, d.w.z. constitutieve kenmerken voldoet, dan constitueert zij daarmee zichzelf tot autonome actor van die acties.

Het is kortom niet de actor die aan bepaalde kenmerken moet voldoen wil hij autonoom in staat zijn bepaalde acties te verrichten; het is de manier van delibereren over de acties die aan bepaalde kenmerken moet voldoen wil het in staat zijn de actor te constitueren tot een doeltreffende en autonome actor, en dus tot een morele actor.

En zoals we in het voorstel gezien hebben betekent dat, naast de eis dat de CID-Structure aan de constitutieve kenmerken moet voldoen, dat daartoe de verschillende delen binnen organisatie zich moeten identificeren met de CID-Structure, in de zin dat ze de constitutieve waarde van die structuur, van die constitutie, *als zodanig* begrijpen, en van daaruit conform die structuur *willen* handelen, ook al is dat soms niet overeenstemming met de eigen afzonderlijke doelen of belangen. Er dient kortom een normatieve kracht vanuit de structuur te gaan die haar grond heeft in de zelf-constitutie van de organisatie als eenheid. Daarmee kan gezegd worden dat *in* het door de verschillende delen van de organisatie op een reflectieve manier bekrachtigen van de constitutie, en dus conform die constitutie willen handelen, de organisatie als eenheid tot de keuze van acties komt, waar zij als autonome actor achter kan staan, omdat het *de werking van haar constitutie als geheel* is die tot de acties is gekomen.

5. De organisatie als een specifiek soort morele actor

5.1. *Inleiding*

De organisatie, als normatief idee, is niet hetzelfde als een persoon. Het is een door personen gecreëerd instrument dat op de een of andere manier van nut is voor de samenleving. Een organisatie heeft geen fysiek lichaam, is niet onderworpen aan de natuurlijke of biologische noodzaak te overleven zoals dieren en mensen dat wel zijn, maar is in die zin een kunstmatige entiteit.

En toch, binnen de context van het voorstel zoals ik dat ik formuleer, blijkt dat, als het gaat om actie en moreel actorschap, de verschillen niet zo groot zijn. Om dat te laten zien is het zinvol nog even stil te staan bij de wijze waarop Korsgaard uitlegt hoe mensen zich van dieren onderscheiden in relatie tot actie en actorschap. Dat is wat ik hieronder zal doen.

5.2. *De mens als actor onderscheidt zich van het dier in het kiezen van principes*

Wat maakt het onderscheid tussen acties van dieren en van mensen? Volgens Korsgaard (en Kant) is het startpunt van een actie een aansporing, een motivationeel geladen representatie van een object. In dat geval beantwoordt een waargenomen object aan iets in jou of aan de conditie waarin je je bevindt. Iets in jou zorgt ervoor dat het object aantrekkelijk is, welkom, interessant, of wat dan ook. Bij dieren werken aansporingen causaal, hoewel ze niet direct de bewegingen van een dier veroorzaken; het is nog geen actie. Als een aansporing direct de bewegingen van een dier zou veroorzaken, zou het niet het dier als geheel, maar *iets* in het dier zijn dat de beweging bepaalde, en dat valt niet binnen het concept van een actie. Volgens Kant zijn daar *principes* voor nodig, die bepalen – of beschrijven – wat de *reacties* van een actor op een aansporing zijn, begeleid door zijn representatie van de wereld. Dergelijke principes zijn de wetten van zijn eigen causaliteit. Ze zijn van hemzelf omdat ze ontstaan uit de vorm of *telos* van de actor, en ze zijn de wetten van zijn causaliteit omdat ze duiden wat te doen voor welk doel. Elke actie moet derhalve resultaat zijn van een aansporing en een principe. Bij dieren zijn dergelijke principes de instincten, bij mensen zijn dat rationele principes, daar de mens handelt op basis van redenen, en rationele principes bepalen wat iemand als reden voor een actie beschouwt. Dieren die instinctieve acties verrichten laten zich niet zozeer leiden door een aansporing, maar door het instinct als principe dat zegt dat als het honger heeft, het moet gaan jagen, als het bang is, het moet wegrennen etc.. Een instinctieve actie is autonoom in de zin dat de bewegingen van het dier niet gestuurd worden door externe oorzaken, maar door de wetten van zijn eigen causaliteit, zijn eigen natuur. Maar een dier kiest niet de wetten van zijn natuur. En wij mensen wel; het is aan te ons te kiezen wat te doen voor welk doel; wij kiezen onze eigen maximes, de inhoud van onze principes. En de categorische en hypothetische imperatieven zeggen hoe we dat moeten doen, hoe we de maximes moeten construeren: het is omdat we de wetten van onze eigen causaliteit moeten kiezen dat we gebonden zijn aan de categorische en hypothetische imperatieven. Door te delibereren in overeenstemming met de imperatieven constitueren we onszelf tot een doeltreffende en autonome eenheid, tot een morele actor, die met die acties zichzelf tot *oorzaak* van een doel bepaalt, en *zichzelf* tot oorzaak van een doel.

5.3. *De organisatie als actor onderscheidt zich van de mens in het inbedden in de samenleving*

Als het gaat om actie en moreel actorschap zoals gehanteerd in het voorstel, springt eerst en vooral een duidelijke overeenkomst tussen de organisatie en de persoon in het oog.

Elke actie is het resultaat van een aansporing en een principe, en die principes dienen constitutieve principes te zijn wil met de actie de morele actor geconstitueerd worden. Net als bij mensen zijn die constitutieve principes voor de organisatie rationele principes, daar de organisatie, evenals de mens, handelt op basis van redenen, en niet op basis van instincten als principes.

Ook de organisatie kiest de inhoud van haar principes, en we hebben gezien dat het ook voor de organisatie de imperatieven zijn die haar vertellen hoe ze dat moet doen, wil de organisatie überhaupt als actor gedacht kunnen worden die op autonome en doeltreffende wijze acties kan verrichten. Voor zowel

mensen als voor de organisatie is het kortom *de manier van delibereren* over de acties die aan bepaalde kenmerken moet voldoen wil het in staat zijn de actor te constitueren tot een doeltreffende en autonome actor, en dus tot een morele actor. In die zin is er geen verschil tussen het individu als morele actor en de organisatie als morele actor. In beide gevallen is er niet eerst sprake van een moreel actorschap voorafgaand aan het handelen, maar wordt het moreel actorschap in het handelen conform bepaalde principes geconstitueerd.

Een verschil is dat wij mensen als rationele wezens moeten handelen, omdat dat onze *noodzakelijke toestand* als mens is. En omdat we redenen nodig hebben om te handelen, en daarmee bepaalde principes waarmee we ons identificeren (omdat we anders geen redenen voor handelen zouden hebben), hebben we een *praktische identiteit* nodig. Het is dus door het onszelf te waarderen als dragers van praktische identiteiten (ook al kunnen die contingent zijn) dat we tegelijkertijd onszelf als *rationele wezens* waarderen, omdat we daarmee een reden omarmen die ontstaat uit onze rationele natuur, namelijk onze noodzaak om redenen te hebben om te kunnen handelen. Dus met het onszelf tot specifiek persoon constitueren, committeren we ons dus aan onze rol als rationele actor. En het is vanuit dat noodzakelijke commitment dat de hypothetische en categorische imperatief de twee principes blijken te zijn op basis waarvan we acties moeten kiezen, willen de acties ons tot een actor als eenheid constitueren.

Maar de organisatie is een *kunstmatige* entiteit, gecreëerd door mensen om instrumentele redenen. Dus *als* er al zoiets is als een noodzakelijke toestand van de organisatie, dan moet dat besloten liggen in het waarom en waartoe ervan, in de reden om überhaupt de organisatie te creëren. Waar de mens genoodzaakt is zich een (willekeurige) praktische identiteit aan te meten wil hij überhaupt redenen voor acties kunnen hebben, is voor de organisatie daarentegen die praktische identiteit deels gegeven: de *vorm* ervan moet zodanig zijn dat het betrekking heeft op het instrumentele karakter van de organisatie, op het van nut zijn, op het zijn van een in de samenleving ingebed instrument. Zonder een dergelijke praktische identiteit is er van de organisatie helemaal geen sprake. En dat betekent dat, zoals we gezien hebben, naast (of beter: geïntegreerd in) de imperatieven als constitutieve principes, er een derde principe geldt; de organisatie is genoodzaakt de acties namens de organisatie te waarderen in het licht van die praktische identiteit.

De organisatie kan daarom pas een morele actor zijn wanneer zowel de hypothetische en categorische imperatieven, als het principe van de inbedding in de samenleving worden gevolgd in de manier van delibereren. Dát is het essentiële verschil tussen het moreel actorschap van de mens en het moreel actorschap van de organisatie.

6. Conclusie

De organisatie kan gedacht worden als morele actor. Dat is wat ik heb willen laten zien met deze thesis. Maar, zo bleek, het is wel een specifiek soort morele actor.

Aan de hand van de zelf-constitutiethorie van Korsgaard heb ik geprobeerd een voorstel te formuleren voor het denken van de integere organisatie. Dat voorstel kwam neer op de claim dat, als de functie van een actie de zelf-constitutie van de actor is, dan is de functie van een *actie namens de organisatie*, de constitutie van de *organisatie als actor* van die actie. Acties namens de organisatie hebben op dezelfde wijze als bij personen en dieren de functie om de actor te constitueren, in het licht van zijn functie als actor. En waar de functie van persoon-zijn is het zijn van een persoon, is de functie van het zijn van een organisatie het zijn van een in de samenleving ingebed instrument dat daar iets aan bijdraagt. Dit betekent dat acties namens de organisatie de organisatie constitueren tot een in de samenleving ingebed instrument dat iets aan de samenleving bijdraagt.

Volgens Korsgaard kan er alleen maar sprake zijn van een actie die aan een actor als auteur ervan kan worden toegeschreven, als die actie de actor *doeltreffend* en *autonoom* maakt. Dit betekent dat acties namens de organisatie pas acties namens de organisatie zijn, als de acties de organisatie constitueren tot een doeltreffende en autonome organisatie. Acties dienen daarom deliberatief, conform bepaalde constitutionele principes, gekozen te worden, zodanig dat de organisatie als eenheid daar achter kan staan. Wanneer de organisatie haar causaliteit bepaalt door zich als eenheid te scharen achter een actie, *moet* zij opereren als een geheel, als iets dat boven haar delen uitstijgt. En om dit te kunnen moet zij zichzelf zodanig inrichten dat ze *kan* delibereren over de te verrichten acties zodanig dat dit tot eenheid *leidt*. En daartoe dient de constitutie van de organisatie, die zegt *hoe* die deliberatie plaats moet vinden, hoe de te verrichten acties gekozen moeten worden. De constitutie van de organisatie levert als het ware de normatieve principes op basis waarvan te verrichten acties gekozen kunnen worden. En dat moeten de hypothetische en categorische imperatief zijn.

Omdat de functie of praktische identiteit van de organisatie betrekking heeft op de wijze waarop de organisatie is ingebed in de samenleving, dient de organisatie zodanig teleologisch georganiseerd te zijn dat de organisatie die functie kan vervullen. En dat veronderstelt dat het zich als organisatie inbedden in de samenleving niet zozeer een materiële keuze is, maar een (derde) normatief principe waaraan acties zich dienen te conformeren, wil er *überhaupt* sprake zijn van een organisatie.

Bovendien volgt vanuit dat normatieve principe van het inbedden in de samenleving dat de organisatie *interacteert* met de samenleving, en dus dat redenen voor acties namens de organisatie *publieke* redenen moeten zijn, in de zin dat er een normatieve kracht vanuit moet gaan die door alle rationele wezens gedeeld kan worden, inclusief door de organisatie zelf op een ander moment of in een andere situatie: de maximes voor de acties dienen universeel gewild te kunnen worden.

De organisatie *is* ook daadwerkelijk in staat de normatieve principes te integreren in een constitutie. Dat wordt gemanifesteerd met de interne besluitvormingsstructuur, de CID-Structure, die als het ware de constitutie van de organisatie representeert, en waarin een bepaalde vorm van *redelijkheid* is geborgd op basis waarvan voorstellen tot acties kunnen worden getoetst op hun waarde als actie-ter-realiseren-van-een-bepaald-doel. Die redelijkheid binnen de structuur krijgt vooral een plaats in enerzijds een identificatie met en begrip van de functie die organisatierepresentanten binnen de organisatie vervullen, het waarom en waartoe daarvan, in relatie tot de functie van de organisatie als geheel: de ingebedde functie-rationaliteit. En anderzijds in de mate waarin een zelfcorrigerende of reflectieve werking binnen de structuur is geborgd, zodanig dat de interne communicatie, de hantering van conflicten en dialogen over meningsverschillen, in feite de deliberatie gestalte geven.

De waarde toetsen van een actie-ter-realiseren-van-een-doel betekent de maxime ervan toetsen aan de hypothetische en categorische imperatief, én aan het normatieve principe dat uit de praktische identiteit van de organisatie volgt, de inbedding in de samenleving, zodanig dat de acties namens de organisatie de organisatie tot een doeltreffende en autonome eenheid constitueren, die als een integer geheel bijdragen aan de samenleving levert. En daarmee goed is in het zijn van een organisatie.

Dit voorstel draagt op een tweeledige manier bij aan het debat over de organisatie als morele actor.

De eerste betreft het probleem van de mentale toestand en de aanname dat de organisatie *als organisatie* niet over mentale toestanden kan beschikken. We hebben gezien dat het *niet nodig is* om over het vermogen tot het bezitten van een mentale toestand te beschikken, wil die actor een morele actor zijn. Wat nodig is, is het vermogen om te *delibereren* over een te verrichten actie, conform principes die de actor constitueren tot doeltreffende en autonome eenheid aan wie de actie kan worden toegeschreven. En over dat vermogen beschikt de organisatie, wat gerepresenteerd wordt door de CID-Structure van de organisatie.

De tweede betreft de noodzaak van het vermogen om intenties en keuzes voor acties *reflectief te bekrachtigen*, en de veronderstelde noodzaak daartoe voor een actorschap, die op reflectieve wijze keuzes voor acties *kan bekrachtigen*. Het hele punt van het voorstel was juist dat dat actorschap *geconstitueerd* wordt door het verrichten van de actie die op basis van deliberatie conform de imperatieven is verkozen. Het is het verrichten van de actie zelf die de eenheid constitueert. Als binnen de organisatie de deliberatie over de te verrichten acties aan bepaalde formele, d.w.z. constitutieve kenmerken voldoet, dan constitueert zij daarmee zichzelf tot autonome actor van die acties.

Het is kortom niet de organisatie als actor die aan bepaalde kenmerken moet voldoen wil zij autonoom in staat zijn bepaalde acties te verrichten; het is de manier van *delibereren* over de acties namens de organisatie die aan bepaalde kenmerken moet voldoen wil het in staat zijn de organisatie te constitueren tot een doeltreffende en autonome actor, en dus tot een morele actor.

Maar wel tot een *specifiek soort* morele actor, die zich op een essentieel punt onderscheidt van het moreel actorschap van de mens. De organisatie dient als organisatie noodzakelijkerwijs niet alleen de imperatieven, maar ook het normatieve principe van de inbedding in de samenleving te hanteren in het waarderen van te verrichten acties. Pas dan kan er sprake zijn van de integere organisatie.

Bibliografie

- Ancona, D., Kochan, T., Scully, M., Maanen, J. V., & Westney, D. E. (1996). *Managing for the Future: Organizational Behavior and Processes* (Vol. Module 9). Cincinnati: South Western College Publishing.
- Aristoteles. (2009). *Ethica Nicomachea*. (C. Pannier, & J. Verhaeghe, Vert.) Groningen: Historische Uitgeverij.
- Arnold, D. G. (2006). Corporate Moral Agency. *Midwest Studies in Philosophy*(30), 279-291.
- Bratman, M. E. (1993). Shared Intention. *Ethics*(104), 97-113.
- Bratman, M. E. (2008). Intention, Belief, Practical, Theoretical. In S. Robertson (Red.), *Spheres of Reason. New Essays in the Philosophy of Normativity* (pp. 29-61). Oxford: Oxford University Press.
- Bratman, M. E. (2014). *Shared Agency. A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Chandler, A. D. (1962). *Strategy and Structure: Chapters in the History of the American Industrial Enterprise*. Cambridge: MIT Press.
- Collier, J. (1998). Theorising the Ethical Organization. *Business Ethics Quarterly*(8), 621-654.
- Donaldson, T. (1982). *Corporations and Morality*. Englewood Cliffs: Prentice Hall.
- Dubbink, W., & Smith, J. (2011, April). A Political Account of Corporate Moral Responsibility. *Ethic Theory and Moral Practice*, 14(2), 223-246.
- Frankfurt, H. G. (1969). Alternative possibilities and Moral Responsibility. *Journal of Philosophy*, 66, 828-839.
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*(68), 5-20.
- French, P. (1979, July). The Corporation as a Moral Person. *American Philosophical Quarterly*, 16(3), 207-215.
- French, P. (1984). *Collective and Corporate Responsibility*. New York: Columbia University Press.
- French, P. (1995). *Corporate Ethics*. New York: Harcourt Brace.
- French, P. (1996). Integrity, Intentions, and Corporations. *American Business Law Journal*(34), 141-156.
- Friedman, M. (1970, September 13). The Social Responsibility of Business is to Increase its Profits. *the New York Times Magazine*, p. SM17.
- Giddens, A. (1979). *Central Problems in Social Theory*. Berkeley: University of California Press.
- Hume, D. (1739/2014). *A Treatise of Human Nature*. New Delhi: General Press.
- Kant, I. (1785/2008). *Fundering voor de metafysica van de zeden*. (T. Mertens, Vert.) Amsterdam: Boom Uitgevers.
- Kant, I. (1793/2010). *De Religie Binnen de Grenzen van de Rede*. (G. v. Eekert, W. v. Herck, & W. Lemmens, Vert.) Amsterdam: Uitgeverij Boom.
- Kant, I. (1996). *The Metaphysics of Morals*. (M. Gregor, Vert.) Cambridge: Cambridge University Press.
- Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, C. M. (2009). *Self-constitution. Agency, Identity and Integrity*. Oxford: Oxford University Press.
- Ladd, J. (1970). Morality and the Ideal of Rationality in Formal Organizations. *The Monist*(54), 488-516.
- Leavitt, H. J. (1965). Applied Organizational Change in Industry: Structural, Technological and Humanistic Approaches. In J. D. March, *Handbook of Organizations* (pp. 1144-70). Chicago: Rand McNally.
- May, L. (1987). *the Morality of Groups. Collective Responsibility, Group-bases Harm, and Corporate Rights*. Notre Dame: University of Notre Dame Press.
- Pettit, P. (2001). *A Theory of Freedom. From the Psychology to the Politics of Agency*. Oxford: Oxford University Press.
- Pettit, P. (2002). Groups With Minds of Their Own. In F. Schmitt (Red.), *Socializing Metaphysics* (pp. 167-193). Lanham: Rowman & Littlefield.

- Phillips, M. J. (1992). Corporate Moral Personhood and Three Conceptions Of The Corporation. *Business Ethics Quarterly*(2), 435-459.
- Plato. (2010). *De Ideale Staat; Politeia*. (G. Koolschijn, Vert.) Amsterdam: Athenaeum - Polak & Van Gennep.
- Rønnegard, D. (2015). *The Fallacy of Corporate Moral Agency*. Dordrecht: Springer Science+Business Media.
- Scott, W. R. (1981/2003). *Organizations. Rational, Natural and Open Systems* (5th ed.). Upper Saddle River: Prentice Hall.
- Velasquez, M. G. (1983). Why Corporations are not Morally Responsible for Anything they do. *Business and Professional Ethics Journal*, 2(3), 1-18.
- Werhane, P. (1985). *Persons, Rights & Corporations*. Englewood Cliffs: Prentice Hall.
- Werhane, P. (1989). Corporate and Individual Moral Responsibility: A Reply to Jan Garrett. *Journal of Business Ethics*(8), 821-822.