

# Op zoek naar morele illusies

**Abstract – Looking for Moral Illusions.** Perceptual illusions and cognitive distortions are proof that our spontaneous intuitive judgements are not always trustworthy. In this article I will use the paradigmatic example of a visual illusion developed by Müller-Lyer in order to develop a method for identifying illusions, without thereby referring to the existence of an objective world of primary qualities or factual properties. This method will also enable us to identify moral illusions – persistent moral intuitions that are at odds with a coherent system of backing intuitive judgements and ethical principles.



Perceptuele illusies zijn perceptuele ervaringen die inconsistente oordelen genereren. Een paradigmatisch voorbeeld is de Müller-Lyer optische illusie waar twee lijnsegmenten verschillende lengtes lijken te hebben (Müller-Lyer 1889). Onze oordelen en beslissingen zijn kwetsbaar voor cognitieve illusies, vertekeningen en vooroordelen (Pohl 2004). Daarom kunnen we verwachten dat er ook morele illusies bestaan: hardnekkige maar onjuiste intuïtieve morele oordelen.

Het idee van morele illusies is niet nieuw. Sir David Ross (1930) vergeleek onze morele overtuigingen of intuïties met zintuiglijke waarnemingen: de morele intuïties zijn de basisgegevens van de ethiek, net als de zintuiglijke waarnemingen de basisgegevens zijn van de natuurwetenschappen. Maar hij merkte op: “Net zoals sommige van de laatste wellicht als illusie moeten worden afgewezen, zo ook sommige van de eerste; maar zoals die laatsten worden afgewezen als ze in strijd zijn met andere, nauwkeuriger zintuiglijke waarnemingen, zo worden de eersten afgewezen als ze in conflict zijn met andere overtuigingen die beter de test der reflectie doorstaan” (Ross 1930, 41).

- 
1. Stijn Bruers (1981) is doctor in de wetenschappen (theoretische natuurkunde en ecologie, KU Leuven), doctor in de moraalwetenschappen (dierenethiek, UGent) en projectverantwoordelijke ecologische voetafdruk bij Ecolife vzw. Als activist zet hij zich reeds vele jaren in voor mens, dier en natuur. Zo is hij als vrijwilliger betrokken bij verschillende organisaties die werken op het vlak van ecologie, mensenrechten, dierenrechten, sociale rechtvaardigheid, vrede en natuurbescherming. Stijn was de eerste Belg die met Sea Shepherd meevoerde in Antarctica om te strijden tegen de illegale Japanse walvisjacht. Verder geeft Stijn regelmatig voordrachten en schrijft hij artikelen en boeken over milieufilosofie en ethiek. Website: <https://stijnbruers.wordpress.com/>

Zeventig jaar later, wanneer wetenschappers en filosofen meer en meer geïnteresseerd werden in de neurobiologie van de moraal, verscheen de idee van morele illusies en de analogie met optische illusies opnieuw op het toneel. Zo verklaarde bijvoorbeeld Sam Harris dat het verschil in morele afkeuring tussen martelen van een vermoedelijke terrorist (om de locatie te vinden van een bom die op het punt staat honderden personen te doden) en onbedoelde nevenschade (*collateral damage*) in een oorlog een morele illusie is: “Ik geloof dat we hier stuiten op een ethische illusie – analoog aan de perceptuele illusies die zodanig interessant zijn voor wetenschappers die de visuele banen in de hersenen willen bestuderen. De volle maan aan de horizon is niet groter dan de volle maan als die recht boven ons staat, maar ze lijkt wel groter, om redenen die nog onbekend zijn voor neurowetenschappers” (Harris 2004, 198).

In dit artikel probeer ik een stap verder te gaan dan de loutere suggestie dat morele illusies zouden kunnen bestaan. Twee cruciale vragen moeten worden opgelost:

1. wat bedoelen we met een onjuist moreel oordeel? en
2. hoe gaan we dergelijke morele illusies opsporen, als ze bestaan?

In dit artikel beargumenteer ik in welke richting we kunnen beginnen zoeken naar antwoorden op deze twee vragen. Ik zal dit doen door te wijzen op interessante analogieën tussen de Müller-Lyer optische illusie en twee hypothetische morele illusies: het tramdilemma (Thomson 1985; Kamm 2007) en groepsdiscriminatie.

Ik maak doelbewust gebruik van de analogie van de optische illusie, om ons meteen te confronteren met de eerste vraag van hierboven: de lengtes in de Müller-Lyer figuur zijn ‘primaire kwaliteiten’, d.w.z. feitelijke, objectieve eigenschappen van de wereld. Een correspondentietheorie van de waarheid zegt nu dat ons oordeel over de lengtes ofwel juist is ofwel fout, afhankelijk van of de lengtes in de figuur objectief (‘in werkelijkheid’) gelijk zijn of niet.

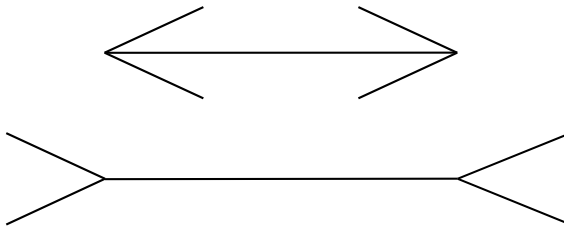
Echter, in de ethiek geloven morele niet-cognitivisten niet dat morele oordelen ware of onware feiten over een objectieve morele wereld uitdrukken. Een eenvoudige correspondentietheorie van de waarheid is daarom niet van toepassing als het gaat om morele oordelen. Dat is de reden waarom ik in de volgende paragraaf de onjuistheid van ons oordeel in de Müller-Lyer illusie aantoon aan de hand van een coherentistische benadering van de waarheid, waarbij ik elke verwijzing naar objectieve morele feiten vermijd. Ik zal alleen gebruik maken van onze intuïties als uitgangspunten. Met deze intuïties construeer ik een samenhangend, coherent systeem in reflectief evenwicht (Rawls 1971). Het toevoegen van enkele achtergrondtheorieën uit de psychologie versterkt de coherentie tot een breed reflectief evenwicht (Daniels 1979). Met andere woorden: ik stel een methode voor om

optische illusies op te sporen die we kunnen gebruiken, zelfs als we in een geconstrueerde *virtual reality* zouden leven waar de lijnstukken in de Müller-Lyer figuur niets meer zijn dan subjectieve, virtuele indrukken.

## De Müller-Lyer optische illusie

Als paradigmatisch voorbeeld van een optische illusie kunnen we de Müller-Lyer illusie nemen (Müller-Lyer 1889). Deze figuur bestaat uit twee horizontale lijnstukken, met naar binnen en naar buiten wijzende pijlpunten op de uiteinden (zie figuur). Het lijnstuk met de naar buiten wijzende pijlpunten lijkt kleiner dan het andere lijnstuk. Met andere woorden: onze perceptuele intuïtie is van oordeel dat de onderste horizontale lijn langer is dan de bovenste.

Figuur 1: de Muller-Lyer optische illusie



Maar hoe weten we eigenlijk dat deze intuïtie een illusie is, zonder te verwijzen naar een objectieve realiteit van de figuur? We mogen niet veronderstellen dat we directe kennis van deze objectieve werkelijkheid kunnen krijgen, want we moeten een methode hebben die ook geschikt is om in de ethiek toegepast te worden. Om tegemoet te komen aan de positie van morele niet-cognitivisten (en ons dus niet te beperken tot een naturalistische meta-ethische positie) is het beter niet te verwijzen naar een objectieve realiteit. De methode mag dus niet verwijzen naar een wereld van morele feiten.

De benadering die ik voorstel is gebaseerd op coherentisme en is een eenvoudig voorbeeld van hoe de methode van reflectief evenwicht werkt. Het vertrekpunt zijn intuïties in plaats van objectieve feiten. Die intuïties zijn als het ware input data voor de constructie van een coherent systeem. Eerst moeten we al onze meest evidente intuïties over meetkunde expliciet maken. Als intuïties onderling inconsistent zijn, houden we de sterkste en elkaar onderling ondersteunende intuïties over. Deze resterende intuïties worden vertaald in algemene principes. De hardnekkige intuïties die niet compatibel zijn met dit sterk samenhangend systeem van principes, zijn dan illusies.

In de Müller-Lyer figuur hebben we drie intuïties. Ten eerste is er de *perceptie van ongelijke lengten* van de horizontale lijnstukken.

Een tweede, sterke intuïtie zegt dat de lengte van een meetlat niet verandert wanneer ze wordt verschoven. Dit lijkt vanzelfsprekend, maar het is onmogelijk om met andere argumenten dit idee te bewijzen<sup>2</sup>, dus daarom is dit een fundamentele optische (of geometrische) intuïtie. Deze intuïtie kan worden uitgedrukt in het zeer belangrijke geometrische principe van de *translatie-invariantie*. Dit is een universeel principe dat geldt voor *alle* meetlatten en *alle* verschuivingen in *alle* richtingen. Dus het moet ook gelden voor de Müller-Lyer figuur. Wij aanvaarden dit principe als vanzelfsprekend, hoewel het mogelijk is om hypothetische werelden of complexe geometrische systemen te bedenken waar translatie-invariantie niet geldig is (wiskundigen hebben reeds vele contra-intuïtieve geometrische systemen geconstrueerd: bv. projectieve, gekromde en zelfs niet-commutatieve geometrieën).

Een derde intuïtie zegt dat de lengte van een lijnstuk niet afhangt van de aanwezigheid van andere lijnen rondom. Deze intuïtie kan worden uitgedrukt in het zeer belangrijke principe van *context onafhankelijkheid*. Ook dit is een universeel principe: de pijlpunten in de Müller-Lyer figuur zijn de ‘context’, en de bewering is dat lengtes *altijd* onafhankelijk zijn van *eender welke* andere geometrische objecten rondom. Dit principe van context onafhankelijkheid is ook gerelateerd aan een intuïtieve afkeer van *willekeur* en *kunstmatigheid* (of ingewikkeldheid).

Willekeur heeft twee aspecten: een verticale en een horizontale. *Verticale willekeur* zegt dat het arbitrair is om te beweren dat lengtes van lijnstukken worden beïnvloed door ‘vierbenige’ figuren (twee pijlpunten met elk twee benen) in plaats van ‘N-benige’ figuren of lijnstukken met andere objecten dan pijlpunten (bijv. lijnstukken met cirkels in plaats van pijlpunten). *Horizontale willekeur* zegt dat het willekeurig is dat binnen de klasse van vierbenige figuren, de lengte van een lijnstuk afneemt wanneer de pijlpunten naar buiten wijzen in plaats van naar binnen. Vergelijk deze dubbele willekeur met een klerkast met daarin verticaal gerangschikte lades. Verticale willekeur is de willekeur om bv. de derde lade met broeken te selecteren. Horizontale willekeur zegt dat in deze derde lade, het arbitrair is om de bruine broek in plaats van de blauwe nemen.

Kunstmatigheid zegt dat de invloed van de hoeken van de pijlpunten op de lengte van een lijnstuk een geometrische regel genereert die veel te ingewikkeld en vergezocht is. Deze kunstmatigheid introduceert een vaagheid: wat gebeurt er als we geleidelijk de hoeken van de pijlpunten openen? Hoe zou dat de lengte van de

2. Met ‘verder argument’ bedoel ik een argument op basis van een ander fundamenteel principe of intuïtie. Natuurlijk moet de intuïtie van translatie-invariantie in een specifieke situatie stroken met gelijkaardige intuïties in andere specifieke situaties, maar deze vorm van samenhang is niet wat ik bedoel met een ‘verder argument’.

horizontale lijnstukken beïnvloeden? Dergelijke mysterieuze invloed lijkt kunstmatig.

De twee principes van de translatie-invariantie en contextonafhankelijkheid impliceren dat we twee instrumenten kunnen gebruiken: een meetlat en iets om de pijlpunten af te dekken of uit te wissen. Met deze instrumenten kunnen we duidelijk zien dat beide horizontale lijnstukken een gelijke lengte hebben. Onze twee principes zijn samenhangend in de zin dat ze tot hetzelfde oordeel komen. Maar ze zijn wel in tegenspraak met het eerste intuïtieve oordeel over de verschillende lengtes van de horizontale lijnen.

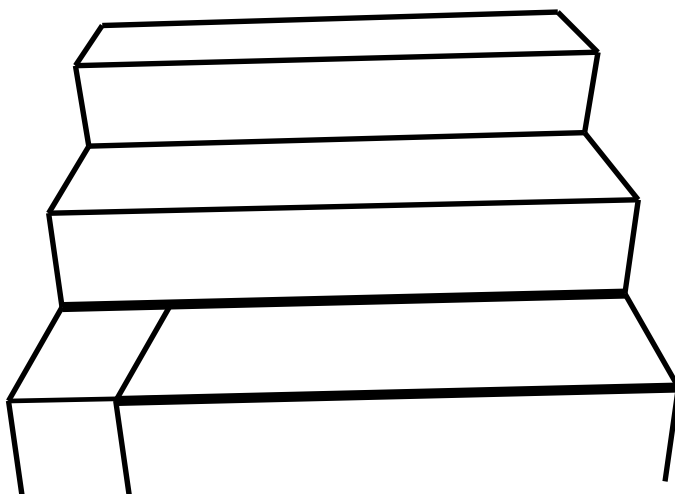
We hebben nu twee opties. 1) We kunnen twee van onze sterkste en coherente intuïties (translatie-invariantie en contextonafhankelijkheid) opgeven en proberen een consistent geometrisch systeem te construeren zonder die twee principes, om onze intuïtie dat de horizontale lijnstukken verschillende lengtes hebben te redden. Zoals wiskundigen een aantal zeer exotische geometrische systemen kunnen construeren, is deze strategie niet per se onmogelijk. Maar iedereen is het erover eens dat een dergelijke procedure om een nieuwe geometrische systeem uit te vinden heel moeilijk zou zijn, en het resulterende geometrische systeem zal zeer kunstmatig lijken. 2) Dus een tweede optie is om te erkennen dat onze intuïtie over de lengte van de horizontale lijnstukken verkeerd is. De intuïtie verdwijnt niet maar ze heeft een hardnekkige cognitieve ondoordringbaarheid (Pylyshyn 1999), wat ze tot een echte illusie maakt.

De meeste personen geven automatisch de voorkeur aan de laatste optie, omdat de combinatie van de intuïties van de translatie-invariantie en contextonafhankelijkheid zeer sterk is, en we ze niet zo gemakkelijk wegwerpen. Translatie-invariantie en contextonafhankelijkheid genereren een samenhangend systeem in smal reflectief evenwicht (Daniels 1979). Aan dit coherente systeem kunnen we wat achtergrondtheorieën toevoegen om te komen tot een breed reflectief evenwicht. Twee achtergrondtheorieën hebben betrekking op de psychologie en antropologie.

Het toevoegen van kennis over het onderliggende psychologische mechanisme genereert meer bewijs dat we te maken hebben met een optische illusie. We weten dat de Müller-Lyer illusie wordt gecreëerd door onze hersenen om een 2D-beeld op het netvlies te vertalen naar een 3D-visie. Twee objecten van gelijke fysieke lengte kunnen verschillende beelden op ons netvlies projecteren, als het ene object verder weg is dan het andere. Onze hersenen corrigeren voor dit verschil in grootte. Het mechanisme is dus perspectiefaanpassing: we leven in een 3D-wereld en daarom zijn onze hersenen getraind om 3D-aanpassingen te maken van 2D-beelden. Soms werkt het mechanisme misleidend, zoals bij het kijken naar de 2D Müller-Lyer figuur.

Wat er eigenlijk gebeurt, kan worden omschreven als een psychologische heuristiek (Kahneman & Shane 2002). Heuristieken zijn intuïtieve, efficiënte vuistregels die worden toegepast wanneer we geconfronteerd worden met complexe problemen. Ze werken door een proces van attribuutsubstitutie: een doelattribuut dat complex is voor onze hersenen (een vraag of probleem dat moeilijk is te beantwoorden, berekenen of begrijpen) wordt (onbewust) gesubstitueerd door een heuristisch attribuut dat gemakkelijker is te berekenen. Kahneman (2003) stelde dat sommige optische illusies worden gegenereerd door heuristische attribuutsubstituties. De volgende afbeelding van een trap kan dat verduidelijken.

Figuur 2: de Muller-Lyer optische illusie in een trap



In 3D weten we dat elke trede dezelfde afmetingen heeft, zodat elke horizontale lijn dezelfde lengte heeft in een 3D-wereld. Maar in 2D hebben de lijnen duidelijk verschillende lengtes. Nu zien we deze afbeelding in 2D, en de taak is het bepalen van de lengtes van de twee dikke horizontale lijnen van de onderste trede. Deze lengtes vormen het doelattribuut. Maar wat doen onze hersenen? We nemen vaak de trap, zodat onze hersenen geoefend zijn in het bepalen van afmetingen van objecten in 3D. Afmetingen in 3D zijn dus voor onze hersenen gemakkelijk te bepalen, en onze hersenen gebruiken ze onbewust als heuristische attributen om afmetingen in 2D-afbeeldingen te bepalen. Zo worden onze oordelen van lengtes in 2D-afbeeldingen vervormd: de lengtes van de twee dikke lijnen zijn in feite gelijk. Maar in een 3D trap is de onderste dikke lijn veel korter, dus dat is waarom ze korter lijkt.<sup>3</sup>

3. Voor een gedetailleerde uitleg van deze illusie, gebaseerd op het mechanisme van tijdsachterstand, zie Changizi e.a. (2008).

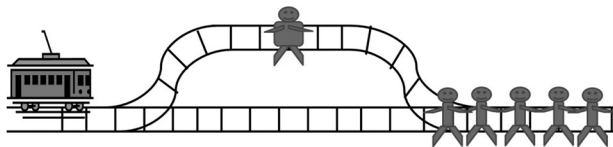
Een tweede achtergrondtheorie verwijst naar de antropologie. Interessant is dat de Müller-Lyer illusie niet aangeboren is. Antropologische studies hebben aangetoond dat de illusie cultureel afhankelijk is (Segall 1963, Ahluwalia 1978). In het bijzonder zijn sommige inheemse volkeren (die niet in een omgeving wonen met rechte randen van huizen, tafels en trappen) minder vatbaar voor deze optische illusie. Dus als zij andere oordelen hebben over de lengtes van de lijnstukken, wie heeft er dan gelijk? Dit is een extra aanwijzing of bewijs dat we inderdaad te maken hebben met een optische illusie.

Twee principes die samenhangen met elkaar, aangevuld met achtergrondtheorieën over het psychologisch mechanisme en de culturele relativiteit, vormen samen heel wat bewijs voor de bewering dat de Müller-Lyer figuur een optische illusie is. Het enige tegenbewijs is dat de illusie niet gewoon verdwijnt nadat we de lengtes gemeten hebben en erover hebben nagedacht. Foutieve intuïties kunnen erg hardnekkig (cognitief ondoordringbaar) zijn.

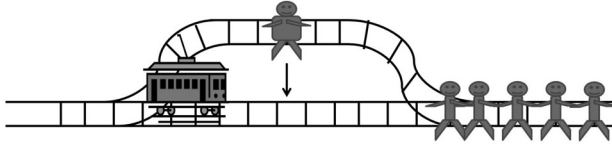
We hebben gezien hoe we een illusie kunnen opsporen in het geval van de Müller-Lyer figuur, zelfs als we geen directe toegang tot een objectieve werkelijkheid zouden hebben. We kunnen een aantal basisoordelen, uitgangspunten of intuïties gebruiken en vertalen naar principes zoals translatie-invariantie en contextonafhankelijkheid. De onderliggende intuïties achter deze principes hebben geen verdere rechtvaardiging. Dit zijn principes die wij moeten accepteren.

## Voorbeeld 1: het tramdilemma

Neem het bekende tramdilemma (Thomson 1985): een tram raast in volle vaart naar vijf personen op het hoofdspoor. Je kunt die vijf personen redden door aan een hendel te draaien waardoor de tram naar het kort zijspoor gestuurd wordt. Dit korte zijspoor buigt terug naar het hoofdspoor, maar op dit zijspoor is een zware persoon die zwaar genoeg is om de tram te blokkeren. Indien die zware persoon er niet was geweest, zou de tram terug naar het hoofdspoor rijden en alsnog de vijf personen doden. Samengevat: niets doen resulteert in de dood van vijf personen, handelen (de hendel overhalen) resulteert in de dood van een zware persoon.



Er is een gelijkaardig dilemma: stel dat je zo lang aarzelde, dat de tram reeds de eerste splitsing passeerde en op weg is naar de vijf personen. Het plan om de wissel over te halen zal niet meer werken. Maar het zijspoor loopt over een brug boven het hoofdspoor. Dus je kunt de tram nog wel stoppen door op de zware persoon van de brug te duwen om de tram op het hoofdspoor te blokkeren. Zou je de zware persoon duwen om de vijf anderen te redden?



Volgens psychologische studies (Hauser e.a. 2008) antwoorden ongeveer 50% van de ondervraagden dat men de wissel in het eerste dilemma mag overhalen, terwijl slechts 10% aangeeft dat men de zware persoon van de brug mag duwen in het tweede dilemma. Een tweede studie (Waldmann & Dieterich 2007) gaf vergelijkbare resultaten. Respondenten konden de toelaatbaarheid van de handeling beoordelen op een schaal van 1 (zeker niet toegestaan) tot 6 (zeker toegestaan). Het dilemma met de wissel kreeg een gemiddelde beoordeling van 4,8, dus de actie (de wissel overhalen) is eerder toegestaan. Maar het tweede brugdilemma kende een gemiddelde score van slechts 3,7, relevant lager dan het eerste dilemma.

In beide dilemma's wordt de zware persoon gebruikt als louter middel, als een menselijk schild of tramblokkeerder, om de vijf personen op het hoofdspoor te redden. Zonder de zware persoon op het zijspoor zou men de vijf anderen niet kunnen redden. Het enige verschil tussen de twee dilemma's is de positie van de tram: vóór of achter de eerste splitsing. Dit beïnvloedt de 'plaats van interventie' (Waldmann & Dieterich 2007) in de causale keten. In het brugdilemma wordt het slachtoffer (de zware persoon) naar het gevaar (de tram) gestuurd, terwijl in het eerste dilemma het gevaar naar het slachtoffer wordt gestuurd. De 'plaats van interventie' in het brugdilemma ligt dus bij de persoon op het zijspoor. Het sturen van het slachtoffer naar de tram betekent dat het traject van het slachtoffer direct wordt beïnvloed. Anderzijds, de wissel overhalen in het eerste dilemma betekent dat de interventie plaatsgrijpt bij de dreiging. Het sturen van de tram naar het zijspoor betekent dat het traject van het gevaar direct wordt beïnvloed (en het traject van het slachtoffer wordt indirect beïnvloed, omdat de dreiging uiteindelijk het slachtoffer zal raken).

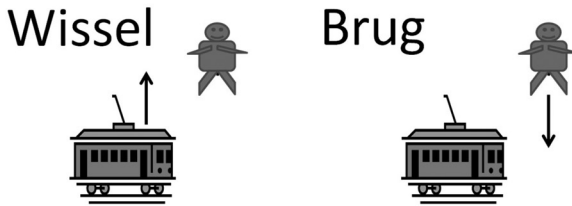


## Morele illusies in het tramdilemma

Is de morele intuïtie in de tramdilemma's een illusie? Om te laten zien dat het een morele illusie (zoals werd geïnsinueerd door Unger 1996, 101), presenteer ik (net als bij de optische illusie) twee argumenten gebaseerd op sterke morele intuïties, en een extra argument gebaseerd op kennis van het onderliggend psychologisch mechanisme.

Het eerste argument is gebaseerd op een vorm van translatie-invariantie. We moeten kijken naar wat constant blijft als we van het wisseldilemma overgaan naar het brugdilemma. In de Müller-Lyer optische illusie hebben we gezien dat de lengte van een meetlat een intrinsieke eigenschap is die constant blijft wanneer we de meetlat verschuiven van het bovenste naar het onderste lijnstuk in de figuur. De lengte van een meetlat verandert niet. In de tramdilemma's blijft er ook een intrinsieke en moreel relevante eigenschap constant als we van het eerste naar het tweede dilemma verschuiven: de morele status van een individu is een intrinsieke en moreel relevante eigenschap die niet verandert als het eerste dilemma overgaat in het tweede (dus als de tram voorbij de eerste splitsing rijdt). De morele status is onafhankelijk van de situatie. Dit lijkt vanzelfsprekend, maar is in feite een sterke intuïtie die wij accepteren. Nu kan deze morele status bestaan uit verschillende dingen, zoals het recht om te leven en het recht om niet te worden gebruikt als louter middel. Dit zijn rechten die personen hebben, ongeacht de situatie. Belangrijker nog: we hebben een sterke intuïtie dat de morele status (de rechten) van een individu niet verandert wanneer de plaats van interventie in de causale keten verandert. De morele status van een slachtoffer is niet hoger wanneer de plaats van interventie bij het slachtoffer ligt in plaats van bij het gevaar.

Het tweede argument is gebaseerd op contextonafhankelijkheid. In de Müller-Lyer illusie kunnen we de irrelevante eigenschappen, zoals de pijlpunten, afdekken of uitgommen om de gelijkheid van de lengtes aan te tonen. Kijken we naar de twee tramdilemma's, wat is dan de irrelevante context die we moeten uitwissen? Het enige relevante aspect is de botsing tussen de tram en het slachtoffer. Waar deze botsing plaatsvindt, is niet relevant. Dus laten we gewoon de tramsporen en alle andere dingen in de omgeving uitwissen. Wat overblijft is een geheel lege ruimte met daarin enkel de tram en het slachtoffer (de zware persoon). In het wisseldilemma wordt de tram omhoog gestuurd richting de zware persoon. In het brugdilemma wordt de zware persoon naar beneden gestuurd richting de tram.



In de Müller-Lyer illusie hebben we een relatie gezien tussen contextonafhankelijkheid en een afkeer voor willekeur. In het tramdilemma, is er ook een vorm van (horizontale) willekeur: het is arbitrair om te zeggen dat het sturen van A naar B moreel erger is dan het sturen van B naar A. Omdat er geen absoluut referentiepunt is (we hebben alle irrelevante context uitgewist), zegt het principe van relativiteit dat beide situaties in feite gelijk zijn. Volgens Unger (1996, 101) is dit verschil tussen het sturen van een slachtoffer naar een gevaar versus een gevaar naar een slachtoffer sturen een vorm van foutieve 'protofysica', wat in strijd is met de relativiteitstheorie in de natuurkunde. Dit toont dat het verschil tussen de twee situaties louter door de irrelevante context wordt bepaald. Maar natuurlijk is dit argument ook gebaseerd op een (sterke) intuïtie over wat de context is en wat moreel (ir)relevant is. De positie van de tramsporen is niet moreel relevant. De enige relevante aspecten zijn de relatieve posities van het gevaar en het slachtoffer, omdat die de botsing bepalen.

Een derde aanvullend argument is gebaseerd op een achterliggend psychologisch mechanisme. Dit mechanisme wordt duidelijk uitgelegd door Waldmann en Dieterich (2007). Hun begrip 'interventiebijziendheid' geeft al aan dat we te maken hebben met iets in ons moreel oordeelvermogen dat niet goed functioneert. Wanneer de plaats van interventie bij het gevaar ligt, zoals in het wisseldilemma, richt onze aandacht zich tot het gevaar, en alle personen (op het hoofdspoor en het zijspoor) gelden dan als achtergrond. Dus vanuit het perspectief van het gevaar (de tram) zijn alle personen gelijk en is het gemakkelijker om consequentialistische berekeningen te maken en het aantal te redden levens te tellen. Maar als de plaats van interventie bij het slachtoffer ligt (de zware persoon op het zijspoor), dan is ook onze aandacht bij dat slachtoffer. De andere vijf personen op het hoofdspoor zijn nu deel van de achtergrond. En de interventiebijziendheid geeft aan dat het moeilijk is om die personen op de achtergrond volledig in rekening te brengen. Als gevolg van die bijziendheid zien we hun morele status niet zo duidelijk, in vergelijking met de status van het slachtoffer op het zijspoor. Dit vertekent onze consequentialistische afwegingen. In het brugdilemma hebben we de neiging ons te richten op het lot van het slachtoffer, en verwaarlozen we de dood van de personen op het hoofdspoor. Deze focus op het slachtoffer (de plaats van interventie), resulteert in een verwaarlozing van andere

personen op de achtergrond. Deze vijf mensen op de achtergrond lijken te ontbreken (ze zijn ver weg in de causale keten).

Kortom, de algemene hypothese is dat mensen de neiging hebben zich te richten op de causale paden van bedreigingen of slachtoffers als doelwit van een interventie en dat ze andere causale processen negeren die zich buiten deze focus, in de achtergrond, bevinden. (...) We beweren niet dat in geval van interventiebijziendheid mensen volledig blind zijn voor de slachtoffers op de achtergrond (dat wil zeggen, de dood van de vijf personen). We zeggen eerder dat als gevolg van een aandachtsfocus over de gevolgen van interventies, mensen die de moraliteit van opties evalueren slachtoffers in de achtergrond een lager gewicht toekennen dan slachtoffers in het aandachtsgebied. (Waldmann & Dieterich 2007, 249)

Interventiebijziendheid kan ook worden beschouwd als een generalisatie van situaties met een dichtbij en persoonlijk contact (Greene 2008). Mensen oordelen negatief over het duwen van de zware persoon van de brug, omdat dat een vorm van dichtbij en persoonlijk geweld is. Maar stel dat in het brugdilemma men niet de zware persoon persoonlijk hoefde te duwen, maar dat men op een knop kon drukken die de brug laat omkantelen en zo de zware persoon voor de tram werpt. Er is niet langer een persoonlijk contact, maar nog steeds is er een zekere directheid.

Met de bovenstaande argumenten kunnen we besluiten dat het verschil tussen het sturen van een gevaar naar een slachtoffer en het versturen van een slachtoffer naar een gevaar een morele illusie is. De morele irrelevantie van dit verschil werd al aangehaald door Fischer (1992) en Fischer & Ravizza (1994), maar nu hebben we een duidelijker beeld en kunnen we spreken van een morele illusie.

Het eerste tramdilemma wordt vaak aangehaald als een tegenvoorbeeld van een deontologisch louter-middel principe dat zegt dat iedereen een recht heeft om niet te worden gebruikt als louter middel (zie bijvoorbeeld Singer 2005; Scanlon 2008). Vele personen oordelen dat men in het wisseldilemma de zware persoon mag opofferen om de tram de blokkeren, hoewel diens deontologische recht daarbij dus wel geschonden wordt: het lichaam van de zware persoon wordt tegen diens wil in gebruikt als menselijk schild. Aanhangers van een consequentialistische ethiek verwijzen naar dit intuïtieve oordeel in het tramdilemma om te betogen dat het deontologische recht een morele illusie is. Maar uit het bovenstaande onderzoek leren we iets interessants: de toelaatbaarheid in het tramdilemma is misschien zelf een morele illusie. En indien dit het geval is, dan is het deontologische recht misschien nog wel geldig. Dus de methode om morele illusies op te sporen zou het deontologische recht kunnen redden, en het tramdilemma kan

dan niet meer worden gebruikt om een deontologische ethiek te bekritisieren. Of het deontologische recht om niet te worden gebruikt als louter middel ook een morele illusie is, valt nog verder te onderzoeken.

## Voorbeeld 2: discriminatie als morele illusie

Laten we discriminatie definiëren als het veroorzaken van een nadeel bij een individu (of groep) ten opzichte van een ander individu (of groep), gebaseerd op een waardegeleden onderscheid tussen individuen (of groepen), waarbij dat onderscheid verwijst naar eigenschappen van de individuen (of groepen) die niet moreel relevant zijn in de betreffende situatie. Met andere woorden: een persoon A discrimineert B ten opzichte van C, als A gelooft (en handelt volgens dat geloof) dat B een lagere morele status heeft dan C (dus dat C volgens A meer rechten, voordelen of kansen moet hebben dan B), waar dit statusverschil gebaseerd is op eigenschappen van B en C die niet moreel relevant zijn of geen aanvaardbaar motief vormen voor de beslissingen en de handelingen van A.

Met morele status verwijs ik naar een intrinsieke, relatie onafhankelijke morele status, wat betekent dat relationele aspecten (nog) geen discriminatie vormen. Beschouw het brandende-huis-dilemma: ik moet kiezen tussen het redden van mijn kind of jouw kind. Ik geef de voorkeur aan mijn kind, want ik voel een sterkere verbinding of relatie met mijn kind dan met jouw kind. Mijn voorkeur (mijn partijdigheid) is geen discriminatie indien ik jouw keuze om jouw kind te redden zou tolereren. Maar als ik denk dat mijn kind een hogere morele status heeft onafhankelijk van mijn relatie met mijn kind, dan zou ik jouw keuze om iemand met een lagere morele status te redden niet tolereren, en dan wordt het wel discriminatie.

In dit artikel richt ik me op groepsdiscriminatie, wat betekent dat de moreel irrelevante eigenschappen verwijzen naar groepen. In de bovenstaande definitie behoort B tot de in-groep van A, terwijl C behoort tot de out-groep.

De Müller-Lyer optische illusie is een zeer verhelderende analogie van de morele illusie van groepsdiscriminatie. De twee horizontale lijnen in de figuur kunnen worden geïnterpreteerd als de morele status (of intrinsieke waarde) van respectievelijk een lid van de in-groep en een lid van de out-groep. De lengte van het lijnstuk stelt de morele status of waarde voor. We hebben dan de intuïtie dat de morele status van een lid van de out-groep lager is dan de morele status van een lid van de in-groep, net zoals we het spontane oordeel hebben dat het ene lijnstuk korter is dan het andere. Een oordeel dat de twee horizontale lijnen in het Müller-Lyer illusie ongelijk zijn, is vergelijkbaar met het morele oordeel van discriminatie.

De kleine lijnen (pijlpunten) stellen de moreel irrelevante eigenschappen voor: de dingen die iemands morele status niet zouden mogen beïnvloeden. Zo beïn-

vloedt de huidskleur het intuïtieve oordeel van racisten dat donkerkleurige mensen een lagere morele status hebben. Net zoals dergelijke moreel irrelevante criteria zijn de pijlpunten in de Müller-Lyerfiguur geometrisch irrelevant. Ze bepalen niet de lengtes van de twee horizontale lijnstukken.

Het breed reflectief evenwicht in de studie van de Müller-Lyer illusie betrof vier elementen: twee principes gebaseerd op sterke intuïties (contextonafhankelijkheid en translatie-invariantie), en twee achtergrondtheorieën uit de psychologie en antropologie. Dezelfde vier elementen treffen we aan in de studie van discriminatie.

## Contextonafhankelijkheid

Eén eigenschap wordt gedeeld door bijna alle optische illusies: de invloed van de context. In de Müller-Lyer figuur vormen de kleine pijlpunten de context. Het geometrische systeem dat we verkiezen, vereist een contextonafhankelijkheid, omdat de invloed van de context irrelevant en arbitrair is. Toch kan de context onze percepties en oordelen beïnvloeden. Dit kan ook het geval zijn met morele illusies zoals discriminatie. We kunnen daarom verwachten dat morele illusies kunnen worden herkend aan hun contextafhankelijkheid, en in het bijzonder aan de karakteristieke kenmerken van de context: willekeur en kunstmatigheid (ingewikkeldheid).

Het maken van een groepsonderscheid is geen triviale zaak. Net als bij de Müller-Lyer illusie is keuze van de in-groep willekeurig en vergezocht in zijn kunstmatigheid.

### *Verticale willekeur*

Ten eerste stuiten we vaak op een hiërarchie van groepen. Neem als voorbeeld religieuze discriminatie. Er bestaat een verticale hiërarchie van classificaties. Neem bijvoorbeeld een Benedictijn. Die behoort tot de groep van alle gelovigen, maar ook tot de Abrahamieten, de christenen, de katholieken, de rooms-katholieken en tot slot de Benedictijnen. De groep van gelovigen is het hoogste niveau in de hiërarchie. Deze groep kan op het tweede niveau onderverdeeld worden in groepen zoals de Abrahamieten en de Dharmische gelovigen. Op het derde niveau worden de Abrahamieten onderverdeeld in onder andere joden, christenen, moslims en rastafari's.

Als iemand christenen tegen moslims wil discrimineren, moet hij/zij aangeven waarom het derde niveau in deze hiërarchie zo belangrijk is. Als protestanten vechten tegen katholieken, moeten ze kunnen argumenteren waarom ze discrimineren op dat niveau in de verticale hiërarchie.

Speciësisme (Singer 1975) heeft een soortgelijk probleem van willekeur. Kijken we naar de biologische classificatie (taxonomie), dan behoor ik tot het koninkrijk der dieren, de stam van chordaten en gewervelde dieren, de klasse van de zoogdieren, de orde van de primaten, de infraorde van smalneusapen, de familie van mensapen, het geslacht *Homo*, de soort *Homo sapiens*, de ondersoort *Homo sapiens sapiens* en de etnische populatie van Europeanen. Er zijn verschillende genetische affiniteiten. Bij speciësisme is er verticale willekeur omdat het arbitrair is om te wijzen naar het niveau van de soorten in deze lijst van biologische categorieën. Ik ben net zo goed een smalneusaap als een mens.

### *Horizontale willekeur*

Na het selecteren van een willekeurig niveau in de verticale hiërarchie gaat discriminatie nog gepaard met een willekeurige selectie van de in-groep op dit niveau. Voorbeelden van deze horizontale willekeur zijn: het selecteren van de groep van islamitische Abrahamieten in plaats van andere Abrahamieten, of het selecteren van de groep van menselijke dieren (*Homo sapiens*) in plaats van andere dieren.<sup>4</sup>

### *Kunstmatigheid (ingewikkeldheid)*

Het geven van een definitie van de in-groep is vaak erg moeilijk. Neem zelfs een ogenschijnlijk simpele vraag als het verschil in sekse: hoe kunnen we de groep van mannen definiëren? Dit lijkt misschien simpel, maar hoe zit het met de verschillende soorten van transseksuelen en interseksuelen (die genitale ambiguïteit of gemengde chromosomale genotypes hebben)? Dergelijke vage grenzen verstoren het schijnbare vanzelfsprekende onderscheid tussen mannen en vrouwen, wat het afbakenen van een in-groep in een seksistische ideologie bemoeilijkt. En zelfs als we een complex geheel van eigenschappen zouden kunnen geven om de groep van mannen te definiëren, dan nog zou die morele relevantie erg vergezocht worden.

Hetzelfde geldt voor speciësisme. Eén van de vele definities van biologische soorten verwijst naar de mogelijkheid om vruchtbare nakomelingen te krijgen met andere leden. Maar waarom zou deze mogelijkheid relevant zijn bij het toekennen van rechten? Het is te vergezocht om te zeggen dat een wezen een hogere morele status heeft als zijn naaste familieleden (bv. ouders)<sup>5</sup> vruchtbare nakomelingen hadden kunnen krijgen met andere wezens. Het is oneerlijk dat een individu rechten krijgt omdat zijn ouders in staat zijn om iets te doen met

- 
4. Merk op dat de verwijzing naar intelligentie, rationaliteit of moreel redeneervermogen geen goede rechtvaardiging zijn voor het selecteren van de soort *Homo sapiens*, want sommige *Homo sapiens* (bv. mentaal gehandicapten) zijn even rationeel als sommige niet-menselijke wezens.
  5. Ik verwijs naar zijn naaste familieleden, omdat het individu zelf onvruchtbaar zou kunnen zijn.

anderen. Het is onrechtvaardig om een principe te hanteren waarbij niet-menselijke dieren gewoon de pech hebben om de verkeerde ouders te hebben. Rechten mogen niet afhankelijk zijn van toevallige empirische feiten.

Hoewel men zou kunnen denken dat mensen duidelijk gescheiden zijn van andere dieren, is er een potentieel vage grens net als bij interseksuelen en gemengde rassen. Ten eerste is het niet onwaarschijnlijk dat een mens-chimpansee hybride geboren kan worden, net zoals paard-ezel, leeuw-luipaard, kameel-lama, dolfijnen-orca en schapen-geiten hybriden kunnen bestaan. Ook een mens-dier chimeer (een individu dat bestaat uit genetisch verschillende cellen afkomstig van menselijke en niet-menselijke zygoten) is biologisch gezien niet onmogelijk. Chimeren kunnen variëren van 100% menselijk tot 100% niet-menselijk. Waar is dan de lijn te trekken van de mensheid? Ook de biologische mogelijkheid van genetische modificatie vervaagt de grens tussen soorten.

Natuurlijk bestaan mens-dier hybriden, chimeren en genetisch gemodificeerde mensachtige wezens (nog) niet, maar andere tussenvormen tussen mensen en niet-mensen hebben wel degelijk ooit bestaan. Een moderne *Homo sapiens* zou vruchtbare nakomelingen kunnen krijgen met een voorouder, die voorouder met een oudere voorouder, en zo gaan we verder omhoog langs een tak in de fylogenetische stamboom totdat we een gemeenschappelijke voorouder tegenkomen van een *Homo sapiens* met een andere soort. Dan dalen we de tak af naar die andere soort. Er is een continue keten van populaties die onze soort verbindt met een willekeurige andere soort. Er is geen punt in deze keten waar een individu plotseling ophoudt een *Homo sapiens* te zijn. Het vreemde is dat een vermeende hogere morele status van A (een *Homo sapiens*) vergeleken met D (een individu van een andere soort) sterk afhankelijk is van het toevallige feit dat de tussenvormen B en C zijn dood.

Al deze overwegingen vervagen de lijn tussen mensen en niet-menselijke dieren. De wetenschap zal nooit in staat zijn om criteria voor te stellen om te bepalen of wezens zoals voorouders, hybriden, chimeren en genetisch gemanipuleerde wezens 'mensen' zouden moeten heten, net zoals wetenschappers niet in staat zijn om te bepalen of zandkorrels een 'hoop' moeten worden genoemd. Speciësisen worden uitgedaagd door de vraag naar de morele status van grensvallen zoals hybriden, chimeren, voorouders en genetisch gemodificeerde mensachtige wezens, net als racisten en seksisten worden geplaagd door vage grenzen tussen rassen en seksen.

### *Het uitwissen van de context*

In de Müller-Lyer illusie werd de irrelevante context (de pijlpunten) gekenmerkt door willekeur, kunstmatigheid en vaagheid. Zoals een groepsonderscheid die-

zelfde eigenschappen heeft, kunnen we zeggen dat dit onderscheid onderdeel is van een irrelevante context. Als we de irrelevante context van groepen uitwissen, dan resulteert dat in een moreel individualisme zoals gedefinieerd door McMahan (2005) en Rachels (1990): de morele status van een individu is uniek bepaald door zijn/haar eigen specifieke kenmerken. Vandaar dat zijn/haar groepslidmaatschappen irrelevant zijn.

## Translatie-invariantie

In de Müller-Lyer illusie konden we een hulpmiddel gebruiken om van de ene positie naar de andere te verschuiven. Dit hulpmiddel, een meetlat, veranderde niet van lengte bij verschuivingen. Daarom is er translatie-invariantie. De analogie van de translatie-invariantie in de ethiek is onpartijdigheid. Empathie, evenals gedachte-experimenten zoals de sluier van onwetendheid (Rawls 1971) zijn mentale hulpmiddelen om van positie te verschuiven, om jezelf in de positie van iemand anders te plaatsen en te meten hoe sterk zijn of haar voorkeuren zijn, hoe belangrijk dingen zijn voor die andere persoon. Ethici moeten medeleven ontwikkelen als een morele deugd, net zoals meetkundigen de juistheid van de meetlaten moeten waarderen.

Net zoals lengte een inherente eigenschap is van een lijnstuk, zo is welzijn een waardevolle eigenschap van voelend wezens. Deze eigenschap is onafhankelijk van context zoals het hebben van een blanke huidskleur, het hebben van een mannelijke geslachtsorgaan, het hebben van menselijke genen, het hebben van tien vingers of het geloven in een specifieke religie.

Rawls (1971) beperkte het gedachte-experiment van de sluier van onwetendheid tot de rationele wezens (die een concept van het goede en een gevoel van rechtvaardigheid hebben). Maar dit gedachte-experiment kan nog onpartijdiger worden gemaakt door het toe te passen op alle entiteiten in het heelal, zoals voorgesteld werd door bijvoorbeeld Rowlands (1997). Stel je voor dat je een entiteit in het universum zou kunnen zijn, maar je weet niet wie of wat je zou kunnen zijn. Je zou een niet-bewust ding zonder welzijn kunnen zijn, of een bewust wezen. Hoe zou je willen worden behandeld? Als je een niet-bewust ding zou zijn, zou deze vraag niet relevant zijn voor jou, omdat niets jouw welzijn zou kunnen beïnvloeden (je hebt namelijk geen welzijn). Je zou niets ervaren en geen enkele voorkeur hebben voor iets. Je zou niet in staat zijn om iets erg te vinden of om iets te willen. Dit impliceert dus dat voor de voelende wezens, die wel een welzijn en een wil hebben, de behandeling wel relevant wordt. Want voor dat voelend wezen is diens welzijn belangrijk. Stel dat je moest kiezen tussen het behouden van je mentale vermogens of het behouden van je menselijk lichaam. Als je liever je mentale vermogens behoudt maar dan met een niet-menselijk lichaam, dan waar-



deur je het hebben van mentale vermogens boven het behoren tot een biologische categorie zoals de soort.

## Psychologische achtergrondtheorieën: cognitieve vertekeningen

### *In-groep-out-groep vertekening (bias)*

Psychologen hebben de mechanismen achter optische illusies zoals de Müller-Lyer illusie bestudeerd (Purves & Lotto 2002). Ook naar de psychologie van discriminatie werd veel onderzoek gedaan, met aandacht voor bijvoorbeeld stereotyperingen en vooroordelen (Whitley & Kite 2010). Groepsdiscriminatie is gebaseerd op een cognitieve vertekening: in-groep-out-groepsvertekening (Tajfel 1981; Whitley & Kite 2010) verwijst naar een patroon van het bevoordelen van leden van de in-groep boven leden van de out-groep. Deze bias bevat elementen zoals out-groep homogeniteit (Quattrone & Jones 1980; Rubin & Badaea 2012), het systematisch onderschatten van de verschillen tussen outgroepsleden. Net zoals bij racisme en seksisme hebben we ook in het geval van speciësisme een hoeveelheid psychologische kennis over vooroordelen, stereotyperingen en de invloed van taal en woorden (zie bijvoorbeeld Plous 2003).

### *Essentialisme*

Groepsdiscriminatie gebruikt ook een andere cognitieve vertekening: essentialisme. Volgens Gelman (2003) is een essentialistische overtuiging het geloof dat er voor elke natuurlijke categorie een (niet-waarneembare) essentiële eigenschap, substantie of kwaliteit bestaat die alle leden van die categorie met elkaar gemeen hebben. Die essentie zorgt ervoor dat alle leden er hetzelfde uitzien en zich hetzelfde gedragen. Het betekent dat er kenmerken zijn die alle elementen van een specifieke verzameling (bijvoorbeeld een soort of een ras) bezitten en elementen van andere verzamelingen niet bezitten. Alle elementen van deze verzameling kunnen dan nauwkeurig worden beschreven en gedefinieerd door deze kenmerken. De specifieke verzameling heeft dus een unieke definitie.

Verschillende studies toonden aan dat mensen een snelle (maar verkeerde) neiging hebben om entiteiten te categoriseren in termen van geëssentialiseerde groepen (Gil-White 2001). Volgens Gelman (2003) en vele andere psychologen beschrijven kinderen en volwassenen spontaan biologische entiteiten in essentialistische termen. Mensen (van verschillende culturen en achtergronden) denken automatisch dat biologische categorieën of rassen onzichtbare essenties hebben (Bloom 2010). Het is alsof onze hersenen zijn getraind in essentialistisch denken om groepen te categoriseren.

Vele filosofen die nadenken over het dierenvraagstuk zijn essentialisten door te verwijzen naar de mensheid (de menselijke soort) als een 'aard', het hebben van een 'substantiële natuur'.<sup>6</sup> Ook onze taal weerspiegelt essentialistisch denken dat mensen van dieren onderscheidt. Om de vreemdheid van dit automatische mensdenken te zien, kunnen we systematisch het woord 'mens' vervangen in bijvoorbeeld 'aap' of 'mensaap' en 'soort' in 'infraorde' of 'familie'. Wij zijn net zo goed apen als dat we mensen zijn, dus blijven de argumenten van essentialistische denkers even geldig als we van categorie veranderen. Bijvoorbeeld: de meeste (maar niet alle) mensen hebben rationaliteit, maar ook de meerderheid van de mensapen hebben rationaliteit. Ook de definitie van discriminatie wordt vaak automatisch beperkt tot de mensen (men definieert discriminatie vaak als: "een nadelige behandeling van een groep van mensen...").

Essentialisme in de biologie vindt zijn wortels ondermeer in oude filosofische stromingen (bijv. Platonisme) en in religies. Maar sinds Darwin is er een wetenschappelijke consensus die zegt dat er niets bijzonders is aan een soort. Het is gewoon een willekeurige abstracte indeling met zijn beperkingen en moeilijkheden (Rachels 1990). Ook racisten denken dat rassen of etnische groepen geëssentialiseerde natuurlijke groepen vormen, ook al is het nu bekend dat er geen essentie gekoppeld is aan een etnische groep of ras. Het feit dat een racist, seksist of speciësisist de neiging heeft om essenties toe te schrijven aan rassen, geslachten of soorten, betekent niet dat deze groepen daadwerkelijk zulke essenties hebben.

De biologie spreekt ons intuïtieve essentialistische denken tegen. Onze intuïties zijn niet altijd in lijn met de wetenschap. Het zien van een essentie in een willekeurige en abstracte (moeilijk te definiëren) verzameling van wezens is als het zien van een essentie in lijnstukken met naar buiten wijzende pijlpunten. Alleen eenvoudige rechte lijnstukken hebben een belangrijke essentie, want ze zijn primitieve geometrische objecten en kunnen op een unieke, eenvoudige manier worden gedefinieerd (bijvoorbeeld als de kortste afstand tussen twee punten). We construeren een meetkundige systeem met lijnen en niet met speciale complexe figuren met pijlpunten. Zo ook moeten we een ethiek construeren dat gebruik maakt van moreel relevante criteria (bijvoorbeeld welzijn), in plaats van met zeer complexe categorieën zoals rassen, geslachten en soorten.

## *Heuristieken*

Heuristieken (intuïtieve vuistregels) zijn een ander psychologisch mechanisme dat cognitieve vertekeningen kan genereren. De Müller-Lyerfiguur wordt gegenereerd door een heuristiek van attribuutsubstitutie, waar onze hersenen een moei-

6. Zie bijvoorbeeld Chappell 2011; Cohen 2001; zie ook McMahan 2005, en Tanner, 2006, voor meer referenties en een uitgebreide kritiek op het 'argument van soorten'.

lijk probleem oplossen door het probleem te vereenvoudigen (het doelattribuut te vervangen door een eenvoudiger heuristisch attribuut). Zo ook bestaan er morele heuristieken (Sunstein 2005). Groepsdiscriminatie zoals speciësisme kan worden worden opgevat als een morele heuristiek (Bruers 2013). Kort gezegd wordt in de speciësismeheuristiek een doelattribuut (het hebben van een mentale capaciteit zoals bv. rationeel denkvermogen) vervangen door een eenvoudiger heuristisch attribuut (het hebben van een menselijk uiterlijk). Dit heuristisch attribuut is voor ons vaak makkelijker op te sporen, omdat we goede patroonherkenningvaardigheden hebben die ons in staat stellen om snel te zien of iets een mens is.

## **Psychologische achtergrondtheorieën: cognitieve ondoordringbaarheid**

Een kenmerk van optische illusies is de cognitieve ondoordringbaarheid (Pylyshyn 1999): zelfs na het meten van de lengtes in de Müller-Lyerfiguur lijken ze nog steeds verschillend. De vraag is of schadelijke discriminatie een zelfde soort cognitieve ondoordringbaarheid heeft: weerspiegelen onze spontane, intuïtieve oordelen over de morele status van verschillende individuen nog een zekere discriminatie nadat we inzicht hebben verworven in bovenstaande argumenten van contextonafhankelijkheid en translatie-invariantie?

Misschien wel het beste wetenschappelijke bewijs dat discriminatie tot op zekere hoogte cognitief ondoordringbaar is, komt uit studies over impliciete associaties (Greenwald e.a. 1995; Devine 2001). De Impliciete Associatie Test (IAT) is een experiment om spontane impliciete attitudes die mensen hebben tegenover bijvoorbeeld rassen te meten. De reactiesnelheid wordt gemeten wanneer proefpersonen zo snel mogelijk paren van concepten met elkaar moeten associëren. Die concepten kunnen verwijzen naar rassen (bv. gezichten van zwarten en blanken) en waarden (bv. positieve en negatieve woorden zoals ‘vreugde’ en ‘pijn’). Volgens IAT-studies hebben een heleboel mensen impliciete racistische attitudes, zelfs als die mensen een expliciete antiracistische houding hebben: ze kunnen expliciet zeggen dat ze tegen racisme zijn en dat ze geen waarde hechten aan het rassens onderscheid, hoewel ze kortere reactietijden blijken te hebben wanneer ze in de IAT zwarte mensen moeten associëren met negatieve waarden. Impliciete (onbewuste) vooroordelen en stereotypen kunnen dit verschil tussen expliciete en impliciete attitudes verklaren (Devine 2001).

Wat deze IAT-studies aantonen is een soort van cognitieve ondoordringbaarheid: zelfs als een racist alles leert over foutieve racistische vooroordelen en stereotypering en zelfs als hij of zij erkent dat racisme immoreel is (dat ras arbitrair, kunstmatig en niet moreel relevant is), verdwijnt zijn/haar impliciete negatieve houding ten opzichte van andere rassen niet zomaar.

## De antropologische achtergrondtheorieën: culturele relativiteit

We hebben gezien dat niet iedereen even gevoelig is voor de Müller-Lyerillusie (sommige inheemse mensen zijn niet vatbaar voor de illusie als ze opgroeien in een omgeving zonder rechte lijnen van tafels en trappen). Hoewel de in-groep-out-groep waardeordenen intuïtief opduiken, tonen verschillende studies (Kurzban e.a. 2001; Cosmides e.a. 2003) aan dat de keuze van de in-groep (bijvoorbeeld op basis van ras) niet aangeboren is, maar cultureel afhankelijk is en kan worden beïnvloed of gewijzigd. Net zoals mensen die opgroeien in een omgeving met huizen, trappen en tafels vaak rechte randen zien en daardoor optische intuïties over lengtes van lijnstukken in zich opnemen, zo nemen mensen die opgroeien in een racistische samenleving de racistische ideologie in zich op en beginnen hun groepsonderscheid als natuurlijk te ervaren. De aanleg voor een dergelijk assimilatieproces, bv. de neiging tot het maken van een groepsonderscheid, is misschien wel natuurlijk (aangeboren), maar het resultaat is dat niet: het intuïtieve oordeel wie er dan precies tot de eigen groep behoort, is niet universeel en niet aangeboren.

## Conclusie

In de afgelopen jaren hebben verschillende bevindingen in verschillende onderzoeksgebieden bewijsmateriaal aangebracht voor het bestaan van morele illusies. Ten eerste begonnen een aantal filosofen onze foutieve morele oordelen te vergelijken met optische illusies (Harris, 2004). Ten tweede is er de studie van 'morele heuristieken' (Sunstein 2005) en de ontdekking dat attribuutsubstitutie ook het mechanisme achter vele optische illusies is (Kahneman 2003). Ten derde is er de studie van 'grammaticale illusies' (Phillips et al., 2010.), in combinatie met de taalkundige analogie tussen ons morele denkvermogen en ons taalvermogen (Mikhail 2007; Hauser et al. 2008). Tot slot is er het (neuro)psychologisch onderzoek naar bv. het tramprobleem (Greene 2008), met de hypothese van de interventiebijzandheid in morele intuïties (Waldmann et al. 2007). Deze vier recente ontwikkelingen geven aan dat de tijd rijp is om morele illusies grondig te onderzoeken.

De analogie tussen optische en morele illusies kan ons helpen om morele illusies beter op te sporen en te begrijpen. We hebben een strategie om optische illusies te detecteren met behulp van een aantal samenhangende principes op basis van sterke intuïties (translatie-invariantie en contextonafhankelijkheid), versterkt door achtergrondtheorieën zoals psychologische mechanismen. Een soortgelijke coherentistische strategie bleek te werken om morele illusies op te sporen in het tramdilemma en groepsdiscriminatie. We hebben gezien dat het tramdilemma

een morele illusie genereert en dat het gedachte-experiment daarom niet kan worden gebruikt als een tegenvoorbeeld van een deontologische ethiek die bestaat uit het recht om niet gebruikt te worden als louter middel. Ook discriminatie (racisme, seksisme en speciësisme) is gebaseerd op een morele illusie: het spontane oordeel van een verschil in morele status tussen groepen is een hardnekkige intuïtie die niet compatibel is met een verzameling van sterkere en samenhangende morele intuïties die uitgedrukt kunnen worden in principes van translatie-invariantie en contextonafhankelijkheid (anti-willekeur). Dit resultaat wordt verder ondersteund door psychologische en antropologische achtergrondtheorieën. Het bestuderen van groepsdiscriminatie als een morele illusie zou ons kunnen helpen om discriminerende ideologieën te begrijpen en te bekritisieren.

## Bibliografie

- Ahluwalia, A. 1978. "An Intra-cultural Investigation of Susceptibility to 'Perspective' and 'Pon-perspective' Spatial Illusions." *British Journal of Psychology*, 69, 2: 233-241.
- Bruers, S. 2013. Speciesism as a Moral Heuristic. *Philosophia*, 41, 2: 489-501.
- Changizi M.A., A. Hsieh, R. Nijhawan, R. Kanai & S. Shimojo. 2008. "Perceiving the Present and a Systematization of Illusions." *Cognitive Science*, 32, 3: 459-503.
- Chappell, T. 2011. "On the Very Idea of Criteria for Personhood." *Southern Journal of Philosophy*, 49, 1: 1-27.
- Cohen, C & T. Regan. 2001. *The Animal Rights Debate*. Lanham, MD: Rowman & Littlefield.
- Cosmides, L., J. Tooby, R. & Kurzban. 2003. "Perceptions of Race." *Trends in Cognitive Sciences*, 7, 4: 173-179.
- Daniels, N. 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy*, 76, 5: 256-82.
- Devine, P.G. 2001. "Implicit Prejudice and Stereotyping: How Automatic Are They?" *Journal of Personality and Social Psychology*, 81, 5: 757-759.
- Fischer, J.M. 1992. "Thoughts on the Trolley Problem." In *Ethics: Problems and Principles*, J.M. Fischer & M. Ravizza, red., 308-317. New York: Holt, Rinehart & Winston.
- Fischer, J. M. & M. Ravizza. 1994. "Ducking Harm and Sacrificing Others." *Journal of Social Philosophy*, 25, 3: 135-145.
- Gelman, S. 2003. *The Essential Child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
- Gil-White, F.J. 2001. "Are Ethnic Groups Biological 'Species' to the Human Brain? Essentialism in our cognition of some social categories." *Current Anthropology*, 42, 4: 515-555.

- Greene, J.D. 2008. "The Secret Joke of Kant's Soul." In *Moral Psychology, Vol. 3: The Neuroscience of Morality*, W. Sinnott-Armstrong, red., 359-372. Cambridge, MA: MIT Press.
- Greenwald, A.G. & M.R. Banaji. 1995. "Implicit Social Cognition: Attitudes, self-esteem, and stereotypes." *Psychological Review*, 102: 4-27.
- Harris S. 2004. *The End of Faith*. New York & London: W.W. Norton.
- Hauser M., L. Young & F. Cushman. 2008. "Reviving Rawls' Linguistic Analogy: Operative principles and the causal structure of moral actions." In *Moral Psychology and Biology*. W. Sinnott-Armstrong, red., 107-143. New York: Oxford University Press.
- Kahneman D. & F. Shane. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." In *Heuristics and Biases: The Psychology of Intuitive Judgment*. Thomas Gilovich, Dale Griffin & Daniel Kahneman, red., 49-81. Cambridge: Cambridge University Press.
- Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review*, 93, 5: 1449-1475.
- Kamm, F.M. 2007. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.
- Kurzban, R., J. Tooby & L. Cosmides. 2001. "Can Race Be Erased? Coalitional computation and social categorization." *Proceedings of the National Academy of Science*, 98, 26:15387-15393.
- McMahan, J. 2005. "Our Fellow Creatures." *The Journal of Ethics*, 9: 353-380.
- Mikhail, J. 2007. "Universal Moral Grammar: Theory, Evidence and the Future." *Trends in Cognitive Sciences*, 11, 4: 143-152.
- Müller-Lyer, F.C. 1889. "Optische Urteilstäuschungen." *Archiv für Physiologie Suppl.* 263-270.
- Phillips C., W. Wagers & E.F. Lau, E.F. 2010. "Grammatical Illusions and Selective Fallibility in Real-Time Language Comprehension." In *Experiments at the Interfaces*, J. Runner, red., Syntax & Semantics, vol. 37. Emerald Publications.
- Plous, S. 2003. "Is There Such a Thing as Prejudice Towards Animals?" In *Understanding Prejudice and Discrimination*, S. Plous, red., 509-528. McGraw-Hill, New York.
- Pohl, R. 2004. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgment and Memory*. Hove & New York: Psychology Press.
- Purves, D. & B. Lotto. 2002. *Why We See What We Do: An Empirical Theory of Vision*, Sunderland, MA: Sinauer Associates.
- Pylyshyn, Z. 1999. "Is Vision Continuous With Cognition? The case for cognitive impenetrability of visual perception." *Behavioral Brain Science*, 22: 341-65.
- Quattrone, G.A. & E.E. Jones. 1980. "The Perception of Variability Within In-groups and Out-groups: Implications for the law of small numbers." *Journal of Personality and Social Psychology*, 38, 1: 141-152.
- Rachels, J. 1990. *Created From Animals. The Moral Implications of Darwinism*. Oxford University Press.

- Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Ross, W.D. 1930, 2002. *The Right and the Good*. (With an introduction and bibliography by Philip Stratton-Lake). Oxford: Oxford University Press.
- Rowlands, M. 1997. "Contractarianism and Animal Rights." *Journal of Applied Philosophy*, 14, 3: 235-247.
- Rubin, M. & C. Badea. 2012. "They're all the same!...But for several different reasons: A review of the multicausal nature of perceived group variability." *Current Directions in Psychological Science*, 21: 367-372.
- Segall, M.H., D.T. Campbell & M.J. Herskovits. 1963. "Cultural Differences in the Perception of Geometric Illusions." *Science*, New Series, 139, 3556:769-771.
- Scanlon, T.M. 2008. *Moral Dimensions*. Harvard: Harvard University Press.
- Singer, P. 1975, 1990. *Animal Liberation: A new ethics for our treatment of animals*. 2nd ed. New York: New York Review of Books.
- Singer, P. 2005. "Ethics and Intuitions." *The Journal of Ethics*, 9: 331-352.
- Sunstein, C. 2005. "Moral Heuristics." *Behavioral and Brain Sciences*, 28: 531-573.
- Tajfel, H. 1981. *Human Groups and Social Categories*. Cambridge: Cambridge University Press.
- Tanner, J. 2006. "Marginal Humans: the argument from kinds and the similarity argument." *Facta Universitas*, 5, 1: 47-63.
- Thompson, J.J. 1985. "The Trolley Problem." *The Yale Law Journal*, 94: 1395-1415.
- Unger, P. 1996. *Living High and Letting Die*. Oxford: Oxford University Press.
- Waldmann M. & J. Dieterich. 2007. "Throwing a Bomb on a Person Versus Throwing a Person on a Bomb. Intervention myopia in moral intuitions." *Psychological Science*, 18, nr.3: 247-253.
- Whitley, B.E. & M.E. 2010. *The Psychology of Prejudice and Discrimination*. (2e). Belmont, CA: Thomson-Wadsworth.