

Dit is een pre-print versie van het volgende artikel:

Brey, P. (2000). 'Ethische aspecten van virtual reality,' *Filosofie en Praktijk*, 20/1: 31-44.

Ethische aspecten van virtualiteit

1. Inleiding

Zoals iedere belangrijke nieuwe technologie geeft ook virtual reality technologie aanleiding tot nieuwe ethische dilemma's. Tot nu toe zijn er echter slechts enkele studies verschenen over de ethische aspecten van virtual reality (Cranford, 1996; Whalley, 1995; Thurmel, 1994; Beardon, 1992). Zoals te zien is in deze studies hebben sommige ethische vraagstukken betrekking op specifieke toepassingsgebieden van virtual reality (VR) technologie. Zo kunnen er ethische vragen worden gesteld bij militaire en therapeutische toepassingen van VR, en bij de gevaren van overgebruik van VR-systemen. VR technologie geeft echter ook aanleiding tot ethische vragen die specifieke toepassingen ervan overstijgen, en die betrekking hebben op de technologie in zijn algemeenheid. In dit paper zal ik twee zulke ethische vraagstukken bespreken, welke betrekking hebben op de *representatieve* aspecten en op de *interactieve* of gedrags-aspecten van VR-systemen. Representatieve aspecten van VR-applicaties hebben betrekking op eigenschappen van zulke applicaties die definiëren op welke wijze objecten, toestanden en gebeurtenissen worden afgebeeld of gesimuleerd. Interactieve aspecten hebben betrekking op de handelingen en gedragsvormen die worden mogelijk gemaakt of worden uitgevoerd in VR omgevingen.

De structuur van het paper is als volgt. In de volgende sectie zal een definitie van VR technologie worden gegeven en verdedigd, zullen verscheidene soorten van virtual reality worden onderscheiden die relevant zijn voor de te verrichten ethische analyse, en zullen verschillende toepassingsgebieden van VR-systemen kort worden besproken. Sectie 3 bevat een ethische analyse en discussie van gedrag en interactiviteit in virtual reality. Sectie 4 bevat hetzelfde voor representatie in VR-systemen. In een concluderende sectie worden de belangrijkste ethische implicaties van deze twee discussies nog eens op een rij gezet.

2. VIRTUAL REALITY SYSTEMEN EN TOEPASSINGSGEBIEDEN

De term 'virtual reality' heeft geen standaard betekenis, zelfs niet binnen de virtual reality industrie. Er zijn zeer ruime betekenissen van deze term, volgens welke elk visueel medium, inclusief televisie of zelfs grafische afbeeldingen als voorbeelden van virtual reality gelden. Meer gebruikelijk zijn zeer strikte opvattingen van de term, waarin VR systemen gedefinieerd worden als computersystemen die met behulp van head-mounted displays (op het hoofd gedragen beeldschermen) en datagloves en datasuits (handschoenen en pakken die de stand van het lichaam registreren en druk op het lichaam kunnen uitoefenen) een het gehele veld van de waarneming bestrijkende, interactieve omgeving simuleren. Het wordt momenteel erkend in de virtual reality industrie dat deze opvatting van virtual reality te beperkend is, en dat er vormen van VR zijn die geen head-mounted displays en datagloves en geen totale immersie van de zintuigen vereisen. In *projection virtual reality*, bijvoorbeeld, worden driedimensionale virtuele modellen in een kamer geprojecteerd, die met behulp van speciale brillen vanuit verschillende hoeken kunnen worden waargenomen. In *desktop virtual reality* wordt een virtuele omgeving afgebeeld op een

computerscherm, waarbij dieptewerking wordt verkregen door het dragen van speciale stereoscopische brillen. De gebruikers interageren met de gerepresenteerde omgeving met een dataglove of, vaker nog, met een muis.

In dit paper wordt virtual reality nog iets ruimer gedefinieerd als *een driedimensionale, interactieve, computer-gegenereerde omgeving waarin een eerste-persoons perspectief is opgenomen*.

De totale immersie van de zintuigen wordt in deze definitie niet als een essentiële eigenschap van virtual reality systemen opgevat; deze vereiste wordt immers in projection en desktop virtual reality niet gehonoreerd. Maar ook interactie door middel van datagloves en datasuits is geen noodzakelijk kenmerk van VR, aangezien interactie ook plaats kan vinden via een muis of joystick. Stereo-visie is eveneens niet noodzakelijk, aangezien dieptewerking in virtuele omgevingen ook zonder stereoscopie gesuggereerd kan worden.

Wat wel essentiële eigenschappen van virtual reality zijn, volgens deze definitie, zijn interactiviteit, het gebruik van driedimensionale afbeeldingen, en een eerste-persoons perspectief. *Interactiviteit* is aanwezig wanneer een afgebeelde omgeving manipulatie en navigatie toestaat. Manipulatie heeft betrekking op de directe modificatie van aspecten van de omgeving, bijvoorbeeld door erop te klikken met een muis of door ze te grijpen met behulp van datagloves. Navigatie impliceert de mogelijkheid om de locatie te veranderen vanwaaruit aspecten van de omgeving worden waargenomen en gemanipuleerd, inclusief het kunnen bekijken van voorwerpen vanuit verschillende invalshoeken. Een dergelijke interactiviteit vereist een *driedimensionale grafische weergave*. Een eerste-persoons perspectief, tenslotte, impliceert het kunnen waarnemen van en interageren met een omgeving vanuit een specifieke locatie. Een *eerste-persoons perspectief* suggereert een zekere mate van onderdompeling in een wereld, in plaats van de ervaring van een wereld als een object dat (ten dele) van buitenaf kan worden gemanipuleerd.

De mate van gelijkenis tussen een virtuele omgeving en een echte omgeving hangt af van de mate waarin de perceptuele en interactieve eigenschappen van het VR-systeem of toepassing zijn ontworpen. Zeer realistische VR is VR waarin de gebruiker een gevoel van totale onderdompeling ervaart, zoals in een echte omgeving, omdat de perceptuele en interactieve eigenschappen ervan een rijkheid hebben die vergelijkbaar is met die van een echte omgeving. Deze rijkheid wordt bepaald door de mate waarin het VR-systeem is afgestemd op de *sensomotorische vermogens* van het menselijk lichaam. Het gaat dan om perceptuele vermogens zoals stereo-visie, de waarneming van driedimensionaal gesitueerd geluid, en tactiele waarneming, en motorische vaardigheden zoals grijpen, het draaien van het hoofd, en lopen. In realistische VR zullen de visuele en auditieve waarneming bovendien meestal afgesloten zijn van de werkelijke omgeving. In het algemeen geldt dat een realistische, immersieve virtuele omgeving op succesvolle wijze alledaagse perceptuele en motorische interactie met een omgeving simuleert.

Momenteel het meest aanwezig in de maatschappij zijn vormen van VR met een relatief lage graad van realisme en immersie, waarin driedimensionale omgevingen worden geprojecteerd op een beeldscherm en bestuurd met een muis, toetsenbord of joystick. Immersievere vormen van VR, die meestal gebruik maken van een head-mounted display, zijn nog steeds relatief zeldzaam. Toch worden ze in professionele contexten steeds vaker gebruikt. Zo worden dergelijke systemen tegenwoordig bijvoorbeeld gebruikt voor prototyping (...) en het ergonomisch testen van producten en machines in het industrieel ontwerpen, voor het ontwerpen en het communiceren van ontwerpen in de architectuur en stadsplanning, voor medische en chirurgische simulaties, voor visualisatie en simulatie in de natuurwetenschappen, voor data-visualisatie in de financiële sector, en voor simulaties van gevechtshandelingen in militaire trainingen. Het gebruik van dergelijke professionele systemen is echter nog beperkt, en immersieve VR heeft daarnaast ook nog geen doorbraak gemaakt op de consumentenmarkt.

Een echte doorbraak voor immersieve VR zou echter wel kunnen plaatsvinden in de nabije toekomst, aangezien de verwachting is dat toenames in de verwerkingssnelheid van audiovisuele informatie en kostenafnamen dergelijke systemen binnen bereik brengen van de gewone consument en ze ook aantrekkelijker maakt voor zakelijk en professioneel gebruik. Experts op dit gebied wijzen op belangwekkende potentiële toepassingsgebieden, zoals onderwijs, training, communicatie (van ontwerpen en ideeën), coöperatief werk, amusement, en therapeutische toepassingen, en verwachten dat VR-systemen belangrijke hulpmiddelen zullen worden op deze gebieden in de komende tien tot twintig jaar (b.v., Briggs, 1996; Valente and Bardini, 1995).

In dit essay zal de aandacht uitgaan naar ethische aspecten van representatie en interactie in zowel weinig immersieve als zeer immersieve VR-applicaties. Er zal echter geen aandacht worden geschonken aan ethische vraagstukken die worden opgeworpen bij het ontwerp en gebruik van multi-user (in een netwerk geplaatste, door meerdere gebruikers tegelijkertijd gehanteerde) VR systemen, zodat de discussie beperkt blijft tot ethische aspecten van VR-systemen met een enkele gebruiker ('single-user VR'). Daarnaast gaat de grootste interesse uit naar de ethische aspecten van zeer immersieve VR, vanwege het grotere realiteitsgehalte ervan, waardoor ethische vragen met betrekking tot representatie en interactie in deze systemen het geprononceerdst zijn. Een aantal van de te behandelen voorbeelden zijn echter toch afkomstig van laag-immersieve VR-toepassingen, omdat er nog geen immersieve systemen bestaan die de zelfde morele vragen oproepen.

3. GEDRAG EN MORALITEIT IN VIRTUELE OMGEVINGEN

In virtuele omgevingen kunnen handelingen worden verricht die moreel verwerpelijk zouden zijn in de fysieke wereld. Een VR-toepassing kan in principe elke voorstelbare handeling mogelijk maken. Er zijn dus ook VR-toepassingen mogelijk waarin de meest immorele handelingen zijn toegestaan en worden afgebeeld, zoals moord, marteling, verkrachting, roof, en andere misdaden. Tevens mogelijk in VR zijn rollenspelen waarin uitvoerige scenario's worden uitgespeeld, zoals terroristische missies of de planning en uitvoering van seriemoorden. Momenteel zijn er vrijwel geen zeer immersieve VR-toepassingen waarin zeer immorele handelingen mogelijk worden gemaakt, behoudens enkele medische en militaire toepassingen. Zoals beargumenteerd in sectie 2 kan het echter worden verwacht dat zeer immersieve VR-systemen prominent zullen gaan fungeren in onderwijs, trainingen, therapie, en amusement, en zullen zulke systemen steeds vaker ook worden gebruikt voor het modelleren van sociale situaties, waarin gebruikers interageren met ander virtuele of echte mensen. In zulke toepassingen is de mogelijkheid van onethisch gedrag veel duidelijker aanwezig dan in de meeste huidige applicaties.

Met dit mogelijke toekomstige scenario in het achterhoofd zal ik nu kort stilstaan bij een toepassingsgebied waarbinnen de mogelijkheid van onethisch gedrag in virtuele omgevingen reeds heeft geleid tot controverse, namelijk in computerspellen. Veel hedendaagse computerspellen maken gebruik van realistische, driedimensionale afbeeldingen en incorporeren ook vaak een eerste-persoons perspectief. Interacties worden ook steeds realistischer in zulke spellen, door het gebruik van speciale joysticks, autosturen, en de mogelijkheid tot het spelen over netwerken ('multi-player'). Dergelijke spellen gelden dus als VR-applicaties zoals gedefinieerd in sectie 2. Sterker nog, het is het enige type virtual reality dat op grote schaal is doorgedrongen op de consumentenmarkt, en de ervoor ontwikkelde, relatief goedkope technologie biedt een goede basis voor toekomstige meer geavanceerde VR-systemen voor zowel persoonlijk als professioneel gebruik. In wat volgt zal ik enkele van de handelingen en scenario's analyseren die in huidige driedimensionale computerspellen mogelijk zijn, tesamen met de maatschappelijke reacties die zulke spellen hebben uitgelokt. Vervolgens zal ik overgaan tot een algemene discussie van de

ethische aspecten van gedrag in virtuele omgevingen en de morele verantwoordelijkheid van de ontwerpers ervan.¹

- In het computerspel *Postal*, uitgebracht door Take 2 Interactive, personifieert de speler een massamoordenaar die actief is in een realistische weergegeven stad, met het doel om de politie te ontwijken en zoveel mogelijk onschuldige passanten te vermoorden. Onder meer is het mogelijk om kinderen op een school om te brengen, een fanfare in brand te zetten met napalm, en een protestdemonstratie tegen geweld in computerspellen op te blazen, handelingen die worden beloond met extra punten. Het spel bevat realistische geluidsfragmenten zoals de stemmen van vrouwelijke slachtoffers die om genade smeken, een executie-knop waarmee gewonden kunnen worden afgemaakt, en een zelfmoordoptie, waarbij de moordenaar uitroept 'I regret nothing.' In Australië en verscheidene andere landen is na overheidsensuur een afgezwakte versie uitgebracht.

- In het computerspel *Grand Theft Auto*, uitgebracht door BMG, neemt de speler de rol aan van een crimineel die begint met kleine klusjes voor misdaadbazen, en steeds hoger klimt op de ladder van de georganiseerde misdaad. Het spel bevat inbraak, autodiefstal, en snelheidsraces met de politie, en kent extra punten toe voor handelingen als het neerschieten van politieagenten en het opblazen van een bus vol nonnen. In Engeland hebben conservatieve politici aangedrongen op een verbod op het spel.²

- Het verkoopsucces *Carmageddon*, uitgebracht door Interplay, is een zeer gewelddadig racespel waarin het doel is om andere auto's total-loss rijden en zoveel mogelijk voetgangers en dieren omver te rijden. Deze handelingen worden grafisch afgebeeld en beloond met extra speltijd. In Engeland werd een gecensureerde versie uitgebracht waarin het bloed werd vergangen door een groene substantie en de voetgangers door zombies. In Duitsland resulteerde censuur in de substitutie van menselijke slachtoffers door robots. De recent uitgebrachte opvolger *Carmageddon II* is weer een stuk realistischer dan de eerdere versie, met name in de realistische afbeelding van voetgangers en de wijzen waarop deze kunnen worden verminkt en gedood. Deze versie heeft zelfs tot acties geleid in landen waar de eerste versie dat niet deed. Zo werd in Nederland tegen dit spel geprotesteerd door kamerleden en de vereniging Veilig Verkeer Nederland, en werd besloten tot een verkoopverbod aan jongeren.

- In het spel *The Wild Nines*, gepubliceerd door Interplay, is marteling een thema. De speler neemt de rol aan van Wex, een rebellenleider in een fantasiewereld. Het spel laat toe dat de speler vijanden martelt voor informatie, en vertoont deze martelingen op grafische wijze.

- Het spel *Gender Wars*, uitgebracht door SCI, simuleert een toekomstige maatschappij waarin mannen en vrouwen in oorlog zijn met elkaar. De speler leidt een squadron van een sekse die militaire missies onderneemt tegen vijandelijke soldaten of burgers van de andere sekse, bijvoorbeeld het kidnappen of vermoorden van een wetenschapper. Deze missies hebben namen als 'Castration' en 'Die, Bastards!'.³

- Het spel *Virtual Surgeon: Open Heart*, uitgebracht door ISM Interactive, is een zogenaamd 'edutainment' produkt, waarin de speler leert en zich amuseert door het uitvoeren van een open hart-operatie, die zeer gedetailleerd wordt gesimuleerd. Dit spel bewijst dat ook spellen die geen geweld, sex of criminaliteit bevatten controversieel kunnen zijn. In Nederland werd er

¹ Niet alle besproken spellen incorporeren een eerste-persoons perspectief. Deze zouden echter kunnen worden aangepast zodat een dergelijk perspectief aanwezig is.

² Het anti-establishment thema in het spel was blijkbaar een belangrijkere reden om op zo'n verbod aan te dringen dan het geweld dat het spel bevat. Ter vergelijking, het vroegere computerspel *Terminator 2*, gebaseerd op de gelijknamige film, werd bestreden door Amerikaanse politiebonden omdat het de speler plaatste tegenover politieagenten, die konden worden neergeschoten.

³ In een verboden advertentie voor het spel was een afbeelding opgenomen van een man bovenop een berg vrouwenlichamen met het onderschrift 'How many women did you kill today?'

geprotesteerd tegen het spel door hartpatiënten en hun familie, vanwege de speelse behandeling van een emotioneel geladen onderwerp.

Het publieke debat omtrent dergelijke zeer gewelddadige of anderszins moreel controversiële computerspellen vertoont gelijkenis met het debat omtrent geweld en moraliteit in andere media, zoals televisie, film, en literatuur. Sleutelthema's in de publieke discussie zijn of dergelijke spellen moeten worden gecensureerd, welke geestelijke en morele schade ze gebruikers kunnen berokkenen, met name kinderen, en of ze individuen kunnen aanzetten tot het onethisch handelen in de werkelijke wereld. In de standaard pro-censuur positie wordt gesteld dat dergelijke spellen immoreel zijn, dat ze de morele ontwikkeling negatief beïnvloeden, dat ze tot immoreel of asociaal gedrag leiden in de werkelijke wereld, en dat de staat op basis hiervan het recht heeft om ze te censureren. In de standaard anti-censuur positie wordt het libertijnse standpunt verdedigd dat aangezien immorele handelingen in een virtuele omgeving geen schade berokkenen aan anderen, de beslissing tot zulke handelingen een persoonlijke is, en dat de staat daarom niet het recht heeft om censuur op te leggen. Het wordt dan vaak toegevoegd dat er geen bewijs is dat dergelijke spellen leiden tot immoreel gedrag in de werkelijke wereld, en soms ook dat zulke spellen positieve effecten kunnen hebben doordat ze gebruikers toestaan om fantasieën of opgekropte frustraties uit te leven die ze anders misschien in de werkelijke wereld zouden uitleven.

In de rest van deze sectie zal ik bijzondere eigenschappen van gedrag in virtuele omgevingen belichten die voor dit morele debat relevant zijn, maar die tevens implicaties hebben voor de ethische reflectie op interactiviteit in niet voor amusement ontworpen virtuele omgevingen. Ten eerste moet een belangrijk verschil tussen VR-applicaties en andere media worden aangewezen, dat in het publieke debat onvoldoende wordt herkend. Dit is het verschil dat waar de meeste media, zoals televisie en lectuur, passieve media zijn die slechts worden ervaren door hun gebruikers, VR-toepassingen vereisen van de gebruiker dat deze actief gedrag vertoont. In VR media is de gebruiker geen spectator maar een actor. Het is daarom niet alleen de ervaren inhoud van het medium waarbij vragen kunnen worden gesteld (b.v. de afbeelding van geweld of sex), ook de gedragskeuzen waarvoor de gebruiker wordt gesteld kunnen een onderwerp van morele reflectie worden.

Hoewel in sommige media, zoals bordspelen, de gebruiker ook actor is, geldt voor VR-applicaties het verschil dat zij streven naar realistische simulatie van een omgeving. De gelijkenis tussen handelen in de echte wereld en handelen in het medium is het grootste bij zeer immersieve VR-systemen. Het verschil tussen VR en andere media is dat VR de gebruiker in staat stelt om een eerste-persoons ervaring te verkrijgen van hoe het is en hoe het voelt om bepaalde immorele handelingen uit te voeren of om criminele rollen aan te nemen. Verder is deze ervaring, zoals Cranford (1996) heeft benadrukt, meestal onvergezd door een dreiging van (echte of virtueel weergegeven) straf. Daarnaast stellen VR-applicaties de gebruiker in staat om de perceptuele en motorische vaardigheden en de denkwijzen aan te leren die gepaard gaan met dergelijke handelingen en rollen. Kortom, VR media hebben een aparte morele status in vergelijking met andere media, omdat ze (realistische) handelingen mogelijk maken, waarvan sommigen moreel controversieel zijn.

Welke argumenten kunnen echter überhaupt aangevoerd worden voor de morele verwerpelijkheid van gewelddadige of anderszins controversiële handelingen in virtuele omgevingen, die immers niet direkt aan derden schade lijken te berokkenen? Ik zal deze vraag beantwoorden door te onderzoeken hoe zij zou kunnen beantwoord vanuit de twee meest

invloedrijke benaderingen in de hedendaagse ethiek, namelijk de Kantiaanse plichtenethiek en het consequentialisme.

De *Kantiaanse plichtenethiek*, ten eerste, houdt het fundamentele principe hoog dat mensen een plicht hebben om andere mensen met respect te behandelen, dat wil zeggen, om ze te behandelen als doel en niet als middel, ofwel om hen te behandelen zoals men zelf ook verwacht door anderen behandeld te worden. Uit dit principe volgen geen directe morele dilemma's voor gedrag in virtuele omgevingen. De virtuele personen die in virtuele omgevingen mishandeld en vermoord kunnen worden zijn namelijk geen echte personen, maar slechts simulaties. Mensen hebben dus geen intrinsieke plicht om virtuele personen met respect te behandelen. Wat wel mogelijk beargumenteerd zou kunnen worden is dat de plicht om echte personen met respect te behandelen een plicht impliceert om ook virtuele personen met respect te behandelen. Maar met welke argumenten zou een dergelijke bewering ondersteund kunnen worden? Voor zover ik kan zien kunnen hiervoor twee argumenten verschaft worden, die ik zal omschrijven als het *morele ontwikkelings-argument* en het *psychologische schade-argument*.

Het eerste genoemde argument kan worden afgeleid uit opmerkingen van Kant zelf. Kant heeft beweerd, zoals ook andere denkers als Augustinus en Descartes, dat hoewel dieren geen intrinsieke waarde hebben en daarom geen respect vereisen, men toch dieren niet wreed mag behandelen, omdat dit ertoe kan leiden dat men ook mensen wreed gaat behandelen. Dit zou gebeuren omdat volgens Kant de gevoelens waarop een beroep wordt gedaan in de behandeling van dieren dezelfde zijn als de gevoelens die gepaard gaan met de behandeling van dieren, vanwege de grote gelijkennis tussen de handelingen. Volgens Kant zou wreedheid jegens dieren dus resulteren in wrede omgang met mensen, en zachtaardige omgang met dieren in respectvolle omgang met mensen (1963, p. 240-241). Het zal duidelijk zijn dat als deze redenering klopt voor dieren, hij ook zal gelden voor virtuele personen, waarbij de gelijkennis met echte mensen nog groter is. Anderszijds moet geconstateerd worden dat de redenering van Kant deels steunt op een psychologische hypothese, namelijk dat het handelen en het gevoelsleven van mensen zodanig zijn ingericht dat wreed of immoreel gedrag gepraktiseerd in één domein vaak leidt tot wreed of immoreel gedrag in gelijkende domeinen. Dit is echter een moeilijk bewijsbare stelling, zodat het argument betwistbaar is.

Het tweede argument, het psychologische schade-argument, houdt in dat derden psychologische schade kunnen oplopen door de kennis dat afbeeldingen van henzelf of van andere mensen of dingen die zij waarderen niet met respect worden behandeld. Uitgangspunt is hierbij dat mensen zich identificeren met afbeeldingen van zichzelf of personen of zaken die zij waarderen. Als zulke afbeeldingen niet met respect worden behandeld, dan voelen mensen zichzelf daardoor geweld aangedaan. Zo hebben de hartpatiënten die zich identificeren met de generische hartpatiënt in het spel *Virtual Surgeon* blijkbaar de ervaring dat het onserieuze karakter van het programma niet verenigbaar is met respect voor henzelf. De virtuele operaties in het spel zijn immers niet bedoeld om medische studenten te helpen levens te redden, maar voor een sensatiebelust publiek om zich te amuseren. Anders gezegd, in het spel fungeren hartpatiënten als middel, en niet als doel.

Tegen het psychologische schade-argument kan echter ingebracht worden dat de mate van identificatie met afbeeldingen sterk verschilt tussen personen, en dat zij die zich hiermee sterk identificeren dit mogelijk te veel doen, en zouden moeten leren inzien dat de meeste mensen respectloze handelingen in de fantasiewerelden van spellen kunnen combineren met een respectvolle behandeling van mensen en dingen in de werkelijke wereld. De vraag is nu wie zich moet conformeren aan de ander: de onrespectvolle spelers, of de zich zo sterk identificerende partijen. Deze controverse verwijst naar een fundamentele controverse in Westerse liberale samenlevingen. In de meeste huidige Westerse maatschappijen geldt als dominant principe dat individuele handelingen moeten worden toegestaan zolang ze anderen geen schade berokkenen.

Schade verwijst hier meestal naar lichamelijk letsel, eigendomsschade, of aantasting van de sociale positie. Psychische schade hoort ook in dit rijtje thuis, maar wordt controversieel zodra het om zulke subjectieve zaken gaat als de eigen levensbeschouwing en identiteit. Bij inperking van de vrijheid van handelen op grond van toegedane psychische schade aan derden moet een zorgvuldige afweging worden gemaakt tussen het recht op vrijheid van handelen en het recht om niet gekrenkt of beledigd te worden. Een dergelijke afweging kan niet a priori gemaakt worden, maar vereist een analyse van de argumenten die voor beide partijen gemaakt kunnen worden.

Cruciaal voor het bij elkaar brengen van de belangen van deze twee partijen is de aanwezigheid van een wederzijds begrip van elkaars motieven en overtuigingen. Wil bijvoorbeeld een spel als *Virtual Surgeon* ooit acceptabel kunnen worden voor hartpatiënten, dan zal er tenminste bij hen de overtuiging moeten bestaan dat de spelers ervan desondanks een minimum aan respect voor hen hebben, op basis van inlevingsvermogen in en waardering voor hun ervaringen en maatschappelijke situatie. Wat de gekrenkte partij mogelijk nodig heeft is een minimaal vertrouwen dat het verlangen tot immoreel of onrespectvol handelen in virtuele omgevingen niet voorkomt uit een gebrek aan respect voor de echte equivalenten van de virtuele wezens of dingen die worden geschaad of ontheiligd in virtuele omgevingen. Overigens geldt dit niet alleen voor de gebruikers van de applicatie, maar tevens voor het ontwerp ervan en de wijze waarop het wordt geadverteerd.

Deze twee argumenten, het morele ontwikkelings-argument en het psychologische schade-argument, zijn met wat aanpassingen ook de twee meest plausibele argumenten die vanuit het *consequentialisme* tegen immoreel gedrag in VR kunnen worden ingebracht. Consequentialistische theorieën stellen meestal dat die handelingen dienen te worden uitgevoerd die het meeste goed en het minste kwaad tot gevolg hebben voor iedereen die gevolgen ondervindt van de handeling. Immorele handelingen zijn dan meestal handelingen die anderen onnodig schaden. Dit argument is even suggestief als zijn Kantiaanse tegenhanger, maar eveneens controversieel vanwege de empirische bewijslast. Het morele ontwikkelings-argument, gereconstrueerd binnen het *consequentialisme*, is dan dat immoreel gedrag in VR leidt tot handelingen in de werkelijke wereld die daarin schadelijke gevolgen hebben. Het nadruk ligt hier dus op schadelijke gevolgen, en niet op de afwezigheid van respect. De consequentialistische versie van het psychologische schade-argument is dat schijnbaar immoreel gedrag in VR vaak ook werkelijk immoreel is omdat het door de gebruiker ervaren plezier of nut niet opweegt tegen de psychologische schade ervaren door hen die door deze handelingen worden gekrenkt. Net als bij zijn Kantiaanse tegenhanger geldt hier dat deze afweging per geval verschillend zal zijn en door middel van onderhandelingen en dialoog beslist zal moeten worden.

Tenslotte wil ik uit bovenstaande discussie enkele lessen trekken uit deze discussie voor het *ontwerpen* van VR-applicaties. De manier waarop een VR-applicatie ontworpen is kan grote invloed hebben op het gedrag van de gebruikers ervan en de manier waarop zij hun eigen gedragingen leren waarderen. Deze invloed van het ontwerp op de gebruiker komt op verschillende manieren tot stand. Ten eerste bepaalt het ontwerp welke handelingen überhaupt mogelijk zijn. Ten tweede bepaalt het ontwerp hoe handelingen en hun gevolgen worden afgebeeld, en of bepaalde gevolgen überhaupt worden afgebeeld. Keuzes hierin van ontwerpers bepalen bijvoorbeeld of moorden wel of niet grafisch worden afgebeeld, of welke ecologische gevolgen in beeld worden gebracht wanneer de gebruiker een rivier vervuult. Ten derde kunnen ontwerpers gedrag stimuleren of afremmen door het expliciet te suggereren of ontraden door middel van in de applicatie opgenomen tekst of symbolen, of door dergelijke handelingen te belonen of bestraffen, bijvoorbeeld door het bijtellen of aftrekken van punten.

Omdat handelingen in virtuele omgevingen moreel omstreken kunnen zijn, en omdat ontwerpers van VR-applicaties mede vorm geven aan deze handelingen en de manier waarop zij ervaren worden, hebben ontwerpers van VR-applicaties een verantwoordelijkheid om moreel te reflecteren op hun ontwerpkeuzen. Specifiek dienen ontwerpers zich af te vragen welke ogenschijnlijk immorele handelswijzen zij in hun applicaties mogelijk maken, hoe deze handelingen en hun gevolgen worden afgebeeld, en op welke wijze deze handelingen door het ontwerp worden gestimuleerd of afgeremd, en beloond of juist bestraft. Het lijkt mij onhoudbaar om te stellen dat de verantwoordelijkheid uiteindelijk alleen bij de gebruiker ligt, omdat het ontwerp mede bepalend is voor de keuzemogelijkheden en keuze-ervaringen van de gebruiker. Met name bij toepassingen die mede voor gebruik door kinderen bedoeld zijn, die nog in een morele ontwikkelingsfase verkeren, dient hier bij stilgestaan te worden. Anderszijds is er het gevaar van paternalisme, wanneer immorele handelingen systematisch onmogelijk worden gemaakt of bestraft of veroordeeld. Het onmogelijk maken van immorele handelingen heeft met name het gevolg dat het moreel handelen in virtuele omgevingen niet meer gebaseerd is op een vrije keuze van de gebruiker. Gebruikers zijn dan geen ‘free agents,’ die vrijelijk kiezen voor een morele handelswijze, en kunnen op deze wijze ook niet door ervaring de grenzen van hun eigen moreel handelen leren kennen.

Momenteel is het morele debat omtrent het ontwerp van de interactieve eigenschappen van VR-applicaties grotendeels beperkt tot realistische computerspellen, maar zoals beweerd in sectie 2 kan er ook grootstalige toepassing verwacht worden in onderwijs, training, en therapie. Applicaties binnen deze gebieden zullen expliciet ontworpen zijn om mensen kennis en vaardigheden aan te leren, en om hun gedrag en opvattingen te beïnvloeden. In dergelijke applicaties zijn ontwerpkeuzen aangaande de interactieve mogelijkheden van de applicaties nog een stuk belangrijker dan in amusementstoepassingen. Zij rusten immers de gebruiker toe met kennis, vaardigheden en inzichten die zeer waarschijnlijk ook in de echte wereld zullen worden gebruikt. Een militair simulatieprogramma dat vereist dat de gebruikers opponenten doden in plaats van hen op een andere manier onschadelijk te maken leert de gebruiker om bepaalde keuzen te maken, zonder hem of haar de vrijheid te geven om alternatieven te verkennen. Omdat op leren gerichte VR-applicaties sterk bepalend kunnen zijn voor de manier waarop later in de werkelijke wereld met situaties wordt omgegaan impliceren zij een extra verantwoordelijkheid voor de ontwerper om te reflecteren op de vormgeving van de interactieve eigenschappen van de applicatie.

4. MISREPRESENTATIE EN BEVOORDEELDE REPRESENTATIE IN VIRTUAL REALITY APPLICATIES

Zoals betoogd in de vorige sectie maken ontwerpers van VR-applicaties soms morele keuzen in de vormgeving van de interactieve mogelijkheden van de applicatie. Zij maken echter ook keuzen omtrent de wijze waarop in VR-applicaties objecten en gebeurtenissen worden afgebeeld, ofwel de *representatieve* eigenschappen van VR-applicaties. In deze sectie zal beargumenteerd worden dat de wijze waarop in virtual reality omgevingen worden gerepresenteerd eveneens ethische vraagstukken oproept, net als de wijze waarop in virtual reality handelingen wordt mogelijk gemaakt en uitgevoerd. Specifiek zal beargumenteerd worden dat keuzen in de representatie van objecten, gebeurtenissen, personen en plaatsen in VR moreel controversieel kunnen zijn doordat zij de werkelijkheid onjuist voorstellen (*misrepresentatie*), of doordat zij individuele of maatschappelijke vooroordelen uitdragen (*bevoordeelde representatie*).

Misrepresentatie treedt op wanneer een VR-applicatie een onjuiste voorstelling van de werkelijkheid geeft. Misrepresentatie kan alleen optreden bij een VR-applicatie waarvan het een doelstelling is van de applicatie om bepaalde aspecten van de werkelijkheid correct of realistisch weer te geven. VR-applicaties verschillen sterk in de werkelijkheidsclaims die zij maken, d.w.z. de impliciete of expliciete beloften aangaande de waarheidsgetrouwheid van (eigenschappen van) de virtuele omgeving. Zo kan verwacht worden van een VR-simulatie van een open hart-operatie die gebruikt wordt in de training van chirurgen dat deze zeer waarheidsgetrouw is, terwijl een edutainment produkt als Virtual Surgeon minder waarheidsgetrouw mag zijn (tenzij er een expliciete belofte van waarheidsgetrouwheid wordt gedaan) en een open hart-operatie in een fantasiespel helemaal niet realistisch hoeft te worden weergegeven. De specifieke waarheidsclaims die bij een VR-applicatie worden gemaakt bepalen normatieve standaarden voor waarheidsgetrouwheid: standaarden die bepalen welke eigenschappen van de gepresenteerde werkelijkheid ook in de simulatie aanwezig moeten zijn, met hoeveel detail deze eigenschappen moeten worden weergegeven, en welke idealiseringen hierbij gehanteerd mogen worden. Misrepresentatie treedt op als aan deze standaarden onvoldoende gevolg wordt gegeven.

Het gevolg van misrepresentatie in VR-applicaties zal veelal zijn dat de gebruikers van de applicatie verkeerde veronderstellingen over de werkelijkheid opdoen. Dit hoeft geen grote negatieve gevolgen te hebben, maar in sommige gevallen kunnen deze zeker optreden, namelijk wanneer de voorstelling van zaken in de VR-applicatie wordt gebruikt als basis voor beslissingen die een grote impact hebben. Dit geldt bijvoorbeeld bij VR-simulaties van chirurgie, of virtuele modellen van gebouwen, bruggen en machines zoals die worden gebruikt in de architectuur en het industrieel ontwerpen. Onjuistheden in de simulatie kunnen hier leiden tot levensbedreigende situaties.

Er kan alleen van misrepresentatie in VR-applicaties sprake zijn als er duidelijke criteria voor waarheidsgetrouwheid zijn waarover overeenstemming bestaat. Veel modelleerkeuzen in het ontwerp van virtuele omgevingen zijn echter subjectiever, omdat zij onderbepaald zijn door de empirische werkelijkheid en ten dele rusten op *pragmatische criteria* die zelf weer gebaseerd zijn op de *praktische waarde* van modelleerkeuzen. In veel modelleerkeuzen in het ontwerp van VR-applicaties speelt waarheidsgetrouwheid een beperkte rol, en wordt rekening gehouden met praktische zaken zoals de aantrekkelijkheid van het ontwerp, de vereisten en beperkingen van de programmatuur waarmee wordt gemodelleerd en de systemen waarop de applicatie later wordt gebruikt, en de hoeveelheid detail die voor praktische redenen gewenst is. Bij deze praktische afwegingen kunnen de waarden en belangen van zowel de ontwerpers als de personen en groepen waaraan zij in het ontwerp tegemoet proberen te komen een belangrijke rol spelen. Zo zijn er veel representatieve aspecten van VR-simulaties die afhankelijk van pragmatische keuzen:

- *Het wel of niet opnemen van objecten en eigenschappen in een simulatie.* in een VR-simulatie van een bestaand gebouw kan besloten worden om graffiti of scheuren in de muren niet af te beelden, terwijl fresco's wel gedetailleerd worden weergegeven. Een VR-simulatie van een bomaanval kan militaire structuren in een gebied weergeven en verzuimen om nabijgelegen civiele structuren weer te geven. In een VR-simulatie van bloedvaten in de hersenen als ondersteuning bij hersenoperaties kunnen kleinere vaten worden weggelaten.

- *Keuzen met betrekking tot de gedetailleerdheid en de mate van realisme waarmee objecten en eigenschappen in een simulatie worden afgebeeld.* Als de keuze is gemaakt om een object of eigenschap in een simulatie wordt opgenomen, moet er vervolgens worden gekozen met welk detail het wordt weergegeven. Mensen en dieren in VR-simulaties kunnen bijvoorbeeld minder of meer realistisch worden afgebeeld, en met een heel beperkt of juist heel uitvoerig gedragsrepertoire

worden toegerust. In een VR-representatie van oorlogshandelingen kunnen wonden worden weergegeven als vage plekken, of juist in grafisch detail worden weergegeven.

- *De selectieve misrepresentatie of geïdealiseerde representatie van objecten of eigenschappen om praktische redenen, inclusief weergave door middel van iconen of symbolen.*

Soms worden objecten of eigenschappen opzettelijk onjuist weergegeven, om een bepaald effect te bereiken of om de simulatie inzichtelijker te maken.

For example, for the communication of design, an idealized model of an industrial complex that simplifies its structure may be more useful than a model that faithfully reproduces every curve and connector piece. In a simulation of surgery, organs and tissues may be rendered in unrealistic colors to facilitate recognition. In a simulation of combat, it may be convenient to remove targets that have been killed or destroyed from the simulation because they no longer serve a purpose in it. Generally, also, objects, texts and symbols in VR environments that are important for the purpose of the application may be foregrounded or highlighted so as to render them more visible to the user.

- *The use of stereotypes in the representation of people, things, and events.* Stereotypical representations of generic situations may not be factually incorrect, in that they might exist in reality as they are represented, but may rest on stereotypes that are not representative of reality at large. In particular, the representation of individuals and their behaviors and cultural artifacts may be based on ethnic or gender stereotypes. The representation of other phenomena like animals and natural environments may draw on stereotypes as well.

- *Built-in assumptions about cause-and-effect relationships and implicit narratives.* VR simulations include principles laws that the modeled environment or objects in it obey to, including a number of causal laws that define cause-and-effect relationships. Some of these laws may be difficult to verify and their inclusion is therefore dependent on personal conviction. For example, a simulation program like SimCity, which simulates city development, contains all kinds of assumptions about causal relationships, for example assumptions about the relation between poverty and crime, that are controversial and difficult to prove. Relatedly, VR simulations may by their sequencing of events contain implicit narratives or story lines. For example, a combat simulation may be structured such that any disobedience of orders ultimately results in the shooting of captivation of its participants.

- *Choices in the visualization of abstract domains.* Data visualization and the simulation of abstract domains require even more pragmatic choices and conventions than the simulation of concrete, observable domains. For example, it involves choices on what data to use, what categories to group data in, what basic modeling parameters to use, and what symbolisms (e.g., color, light intensity) to employ in the representation of abstract features.

Any VR simulation necessarily entails some or all of these pragmatic modeling choices.

Consequently, any VR simulation is to a significant extent value-laden. This value-ladenness does not in itself make VR applications morally controversial. Value choices in VR simulations are only morally controversial, I claim, when these choices disadvantage certain individuals and groups that can justifiably be claimed to have a stake in the simulation because they hold certain values or have certain interests that may be compromised or may fail to be promoted by the way the simulation is set up. Particularly problematical are VR applications used for learning or training purposes that promote stereotypes, obscure consequences of actions, and suggest causal relations that are in fact politically controversial. When a VR application favors certain values or interests over others due to its choices in representation, it may be said that the model makes use of *biased representations*.

I hence distinguish two types of representational failures or shortcomings in VR simulations: *misrepresentation*, in which some aspects of the simulation are clearly mistaken according to

accepted standards of accuracy, and *biased representation*, in which the values or interests of some stakeholders of the simulation are insufficiently accounted for. The principal moral importance of these representational failures is that they may induce false or biased beliefs in users that may ultimately have undesirable practical consequences. A secondary issue is that representational biases may make VR applications less useful or agreeable to users whose values or interests are disregarded in the application.

What moral responsibilities do VR modelers have regarding the possibility of misrepresentation and biased representation in their simulations? In a discussion of mathematical models, which may also be extended to VR models, Richard Mason claims: 'At a minimum, a model builder is obliged to do at least three things: (1) to represent reality to clients adequately, (2) to understand and to incorporate the clients' values into the model in an effective way, and (3) to ensure that actions the client takes based on the model have the desired effect.' (1994: 184). I agree with this assessment, with the addition that from a moral point of view, it is not just the clients' values and desires that are relevant in the assessment of a model, but the values and interests of other stakeholder of the model as well. By dropping Mason's third principle, which I think is implicit in his first two principles, his principles can then be formulated as follows: A model builder is obligated to (1) avoid misrepresentation in his models, and (2) to avoid biased representation in his models, by recognizing stakeholders of the model and by taking into account their values and interests.

The responsibility to avoid misrepresentation entails the responsibility to take proper precautions to ensure that modeling mistakes do not occur, especially when the stakes are high (e.g., in professional simulations of surgery, or design models of collapsible structures like bridges and buildings). It also entails the responsibility to inform users if such mistakes do occur and are difficult to correct. And, finally, it entails the responsibility not to participate in intentional deception of users (e.g., embellishment, dramatization, virtual censorship). These responsibility can be derived from general responsibilities of engineers, as found in professional codes, to hold paramount the safety, health and welfare of the public, to inform clients and other stakeholders of engineering mistakes, and to act in an objective and truthful way.

The responsibility to avoid biased representation can be derived from the general responsibility of engineers (and other professionals) to use their knowledge and skill for the enhancement of human welfare. Certainly, it is in the interest of human welfare that VR applications do not contain biases that disregard the values and interests of some of their stakeholders. Developers of VR applications should therefore take care to recognize how biases may enter into their designs and take steps to avoid such biases. This will require the development of a methodology for the recognition and avoidance of such biases. Although a developed methodology does not currently exist, Friedman and Nissenbaum (1997) is a good starting point, as it analyzes various kinds of biases that may emerge at different stages in the design of computer systems.

It may be noted, finally, that many of the above remarks on misrepresentation and biased representation also apply to other representational media. Paintings can be inaccurate, graphs can be misleading, documentaries may be biased, etc. An important difference between such media and VR, however, is that the representations of VR media present themselves as elements in full-blown, three-dimensional, interactive environments. VR simulations of objects may approach the perceptual complexity and interactive richness of everyday physical objects, and may for this reason more easily generate belief in their veracity and objectivity than other sorts of representations. This effect may be strongest for VR simulations of abstract domains. Financial data, for example, may be harder to disbelieve when one can hold it in one's hand or navigate through it than when it is displayed as numbers on a sheet of paper. However, whether VR media

indeed have this effect of more easily inducing belief in their veracity than other media requires further empirical study.

5. CONCLUSION

The representational and interactive features of VR applications have been argued to raise significant ethical questions. VR applications, especially those aimed at realistic simulation, may involve misrepresentation and biased representation, and developers have been argued to have the responsibility to take proper precautions to avoid misrepresentation or inform users about its occurrence, and to chart the stakeholders of their applications and ensure that their values and interests are accounted for in design choices regarding representational format and content. Behavior in virtual environments also raises ethical questions. These questions have, however, proven to be more difficult to resolve. A more extended discussion is needed of the ethical aspects of immoral behavior in VR environments, along with more empirical research on the implications for real-world behavior of behavior in virtual worlds. Designers were claimed to have the responsibility to reflect on their own standpoint on this issue, and to reflect on the moral aspects of the way in which behavioral options and the consequences of actions are structured and represented in VR applications. This has been argued to be especially important in VR systems used in education, training and therapy, and applications for use by children and adolescents.

Literatuur

Beardon, C. (1992) The Ethics of Virtual Reality. *Intelligent Tutoring Media*, 3, 23-28.

Briggs, J. (1996) The Promise of Virtual Reality, *Futurist*, 50, 13-18.

Cranford, M. (1996) The Social Trajectory of Virtual Reality: Substantive Ethics in a World Without Constraints, *Technology in Society*, 18, 79-92.

Friedman, B. en Nissenbaum, H. (1997) Bias in Computer Systems, in *Human Values and the Design of Computer Technology* (red. B. Friedman), Cambridge University Press, Cambridge.

Kant, I. (1963) Duties to Animals and Spirits, in *Lectures on Ethics* (trans. L. Infield), Methuen & Company, London.

Mason, M. (1994) Morality and Models, in *Ethics in Modeling* (red. W. Wallace), Elsevier Science, Oxford.

Thurmel, S. (1994) Ethische Aspecte der Virtuellen Realität, in *Analyomen 1. Proceedings of the 1st Conference on Perspectives in Analytical Philosophy* (red. G. Meggle and U. Wessels), De Gruyter, Berlijn.

Valente, T. & Bardini, T. (1995) Virtual Diffusion or an Uncertain Reality: Networks, Policy, and Models for the Diffusion of VR Technology, in *Communication in the Age of Virtual Reality* (red. F. Biocca and M. Levy), Lawrence Erlbaum, Hillsdale, NJ.

Whalley, L. (1995) Ethical Issues in the Application of Virtual Reality to Medicine. *Computers in Biology and Medicine*, 25, 107-114.