



Nationaal Comité
advies dierproevenbeleid

Afwegingskader voor het Prioriteren van Dierproeven voor Vervanging

Zienswijze van het Nationaal Comité advies dierproevenbeleid (NCad)
bij het rapport Prioriteren van Dierproeven voor Vervanging



Voor proefdieren van nu en innovaties van morgen

Nationaal Comité advies
dierproevenbeleid

Het NCad en haar werkwijze

Het Nationaal Comité advies dierproevenbeleid (NCad) is een onafhankelijk adviesorgaan dat het welzijn van proefdieren beschermt. Dit doet het comité door gevraagd en ongevraagd advies uit te brengen, innovatie en kennisontwikkeling te stimuleren en stakeholders aan elkaar te verbinden. Hiermee realiseert het NCad zichtbare verbeteringen voor het Vervangen, Verminderen en Verfijnen (3V's) van dierproeven en proefdiervrije innovatie.







Leden NCad

Boven van links naar rechts: Henk Smid (voorzitter), Reineke Hameleers, Coenraad Hendriksen.
Onder van links naar rechts: Wim de Leeuw, Jan-Bas Prins, Roger Adan.

Omslagfoto

pexels.com

Inhoud

Inleiding	4		
 1. Prioriteren van Dierproeven voor Vervanging - samenvatting	6		
Transleerbaarheid	7		
Relevantie van het doel	8		
Ongerief	9		
Integriteitsaantasting	9		
Concluderend	11		
 2. Zienswijze NCad	12		
Reflectie van het NCad op de criteria van Rademaker	13		
Schade-batenanalyse en de criteria van Rademaker	14		
Prioriteren in de ontwikkeling van NAMs	16		
Conclusies en aanbevelingen	16		
 Bijlage 1 Auteurs	18		
NCad-werkgroepleden	18		
Externe deskundige	18		
 Bijlage 2 Lijst van gebruikte afkortingen	19		
 Bijlage 3 Rapport 'Prioriteren van dierproeven voor vervanging' van Corné Rademaker	20		
Samenvatting	20		
Inleiding	21		
Methodologie	22		
Ongeschikte criteria ter prioritering vervanging dierproeven	23		
Mate van proportionaliteit	23		
Ontwikkelingsstatus van alternatieven voor proefdieronderzoek	23		
Aantallen dieren of procedures	23		
Gebruikte diersoort	23		
		Geschikte criteria ter prioritering vervanging dierproeven	24
		Transleerbaarheid van het onderzoek	24
		Diermodellen en modeldieren	24
		Klassieke validiteitscriteria van diermodellen	25
		Nieuwere criteria van validiteit van diermodellen	26
		Toepassingsbereik van validiteitscriteria	28
		Transleerbaarheid als criterium	28
		Tools om de validiteit van diermodellen te bepalen	30
		Is translatie überhaupt mogelijk mogelijk?	33
		Samenvatting	33
		Relevantie van het doel van een dierproef	35
		Typen en varianten van proefdieren onderzoek	35
		Beoordeling van verschillen typen en varianten van proefdierenonderzoek	35
		Samenvatting	38
		Ongerief van proefdieren	39
		Samenvatting	40
		Integriteitsaantasting van de proefdieren	41
		Achtergrond van het begrip integriteit	41
		Receptie van Rutgers en Heeger's definitie van integriteit	42
		Empirische bruikbaarheid van het begrip integriteit	45
		Operationalisering van integriteit	47
		Samenvatting	51
		Beoordelen middels een kader van criteria	52
		Conclusie	52
		Aanbevelingen	53
		Bibliografie	55
		Appendix A	67
		 Bijlage 4 Teksten over fundamenteel onderzoek uit het Werkdocument over projectevaluatie en beoordeling achteraf	68

Inleiding

Het is nu nog niet mogelijk om op grote schaal het gebruik van dieren in veiligheids- en toxiciteitstesten, of in het fundamenteel en translationeel onderzoek te vervangen door dierproefvrije methoden. Wel worden steeds meer zogenaamde New Approach Methodologies (NAMs)¹ ontwikkeld.

Het Nationaal Comité advies dierproevenbeleid (NCad) stelde zich in 2021 de vraag of het mogelijk is te prioriteren welke dierproeven als eerste zouden moeten worden vervangen door NAMs. Hierbij werd onder andere gedacht aan diermodellen met een beperkte translationele waarde waar translatie van onderzoeksresultaten wèl van belang is, diermodellen die ernstig ongerief met zich meebrengen en diermodellen met een weinig relevant of ethisch discutabel doel. Het NCad wilde echter niet met lijstjes komen, maar wel met criteria aan de hand waarvan diermodellen geprioriteerd kunnen worden.

De Europese Richtlijn 2010/63/EU (overweging 12)² gaat ervan uit dat dieren niet louter een instrumentele en/of economische waarde vertegenwoordigen. Dieren moeten daarom altijd behandeld worden als wezens met gevoel. Het gebruik ervan in dierproeven moet worden beperkt tot gebieden die uiteindelijk van nut kunnen zijn voor de gezondheid van mensen, dieren of voor het milieu. In de Wet dieren wordt de *intrinsieke waarde* van het dier expliciet als uitgangspunt genomen.³ Intrinsieke waarde vormt ook de basis van de Wet op de dierproeven (Wod) waarin de bescherming van het proefdier voorop staat.⁴

¹ New Approach Methodologies (NAMs) zijn technologieën en aanpakken die in potentie dezelfde of betere veiligheids- en risk assessment informatie opleveren zonder daarbij dierproeven te doen. Hieronder vallen *in silico*, *in chemica*, *in vitro* en *ex-vivo* aanpakken, maar ook het gebruik van alternatieve diersoorten die minder waarnemingsvermogen (sentience) hebben of waarvan er minder nodig zijn. Inmiddels wordt het begrip ook gebruikt in het biomedisch onderzoek.

² Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. European Commission. Official Journal of the European Union 2010 Vol. L276 Issue L276 Pages 33-79

³ Aanhef Wet Dieren: Alzo Wij in overweging genomen hebben dat het wenselijk is ter uitvoering van Europese verplichtingen en in het belang van de gezondheid en het welzijn van dieren en in dat van de volksgezondheid, regels te stellen betreffende dieren, in het bijzonder door de mens gehouden dieren, onder erkenning van de intrinsieke waarde van het dier (...).
<https://wetten.overheid.nl/BWBR0030250/2022-12-22>

⁴ Aanhef Wod: Alzo Wij in overweging genomen hebben, dat het wenselijk is met het oog op de bescherming van het dier regelen te stellen met betrekking tot het verrichten van proeven op dieren.
<https://wetten.overheid.nl/BWBR0003081/2021-07-01>

In de EU-Richtlijn wordt het wenselijk en ook als einddoel genoemd om het gebruik van dieren in proeven te vervangen door andere methodes waarbij geen levende dieren worden gebruikt (overweging 10, 2010/63/EU, noot 2). Onderkend wordt dat dit op dit moment vaak nog niet mogelijk is. Bij het kiezen van een methode waarbij dieren gebruikt worden, moeten de beginselen van vervanging, vermindering en verfijning in de praktijk gebracht worden waarbij uitdrukkelijk op de eerste plaats staat het gebruik van *alternatieve methoden* als deze minstens even goed zijn voor het beantwoorden van de wetenschappelijke vraag.

Zodra een alternatieve methode is erkend voor regelgevend onderzoek, moet deze ook worden toegepast volgens de EU-Richtlijn. Dat geldt zowel voor alternatieve methoden waarbij *geen proefdieren* meer worden gebruikt als voor methoden waardoor het gebruik van dieren wordt *verminderd of verfijnd* (overweging 10, 2010/63/EU, zie noot 2). Voor dat laatste gelden de volgende richtlijnen (overweging 11, 2010/63/EU, zie noot 2):

- Kies die methode die de beste resultaten oplevert en waarschijnlijk het minste pijn, lijden en angst veroorzaakt.
- Gebruik het kleinste aantal dieren dat betrouwbare resultaten oplevert.
- Gebruik diersoorten die het minst gevoelig zijn voor pijn, lijden, angst of blijvende schade en het meest geschikt zijn voor extrapolatie van de resultaten naar de doelsoort.
- Vermijd zoveel mogelijk dood als eindpunt van een dierproef en vervang die door in een vroege fase vaststelbare humane eindpunten (of beter: humane interventiepunten⁵). Als de dood als wetenschappelijk eindpunt onvermijdelijk is, doe dan wat mogelijk is om het lijden te beperken.

Dit betekent dat als het niet mogelijk is een diermodel te vervangen door een proefdiervrij alternatief, het wenselijk is te bezien of de dierproef kan worden vervangen door een andere dierproef met minder impact op de dieren, of door minder dieren te gebruiken of door het ongerief voor de dieren zoveel mogelijk te beperken.

Bij het verlenen van een vergunning voor dierproeven moet een commissie - in Nederland is dat de Centrale Commissie Dierproeven die geadviseerd wordt door lokale Dierexperimentencommissies - de aanvraag op een aantal aspecten beoordelen, waaronder (artikel 18, 2010/63/EU, zie noot 2):

- De doelstellingen van het project en de voorspelde wetenschappelijke of maatschappelijke baten of educatieve waarde;
- De aandacht die is besteed aan vervanging, vermindering en verfijning;
- De inschatting van het verwachte maximale ongerief.

Vervolgens wordt in een *schade-batenanalyse* nagegaan of het lijden of ongerief gerechtvaardigd wordt door het verwachte voordeel voor mens, dier of milieu met inachtneming van ethische overwegingen (artikel 38, lid 2, 2010/63/EU, zie noot 2).

In het licht van bovenstaande heeft het NCad aan Corné Rademaker, ethicus en lid van meerdere dierexperimentencommissies, gevraagd een rapport te schrijven over prioriteren van dierproeven voor vervanging, dat wil zeggen over de criteria waarmee beoordeeld kan worden voor welke diermodellen met voorrang NAMs zouden moeten worden gezocht. Het rapport dat Rademaker heeft opgesteld is een uitgebreid literatuuronderzoek naar criteria op grond waarvan diermodellen geprioriteerd zouden kunnen worden. Het rapport *Prioriteren van Dierproeven voor Vervanging* is als bijlage 3 bijgevoegd.

Hieronder volgt eerst een uitgebreide samenvatting van het rapport van Rademaker. Daarna volgt de zienswijze van het NCad op hetgeen door Rademaker is aangereikt. Vervolgens wordt doorgedacht over wat de criteria die Rademaker heeft opgesteld, betekenen voor het ethisch afwegingskader. In de laatste paragraaf wordt ingegaan op hoe de criteria behulpzaam zijn bij het stellen van prioriteiten bij het vervangen van diermodellen. Geëindigd wordt met wat een eventueel vervolg kan zijn op deze studie.

⁵ W. O. Williams and P. Baneux. *Humane Intervention Points: Refining endpoint terminology to incorporate non-euthanasia intervention options to improve animal welfare and preserve experimental outcomes*. Lab Anim 2022 Vol. 56 Issue 5 Pages 482-489.
<https://doi.org/10.1177/00236772221090801>



1.

Prioriteren van Dierproeven voor Vervanging - samenvatting

In het rapport *Prioriteren van Dierproeven voor Vervanging* (bijlage 3) gaat Rademaker in op de overwegingen – criteria – waarop een dierproef beoordeeld kan worden en die helpen bij de afweging of deze al dan niet met prioriteit uitgefaseerd zouden moeten worden. Zijn gedegen literatuuronderzoek gaat in op de gronden voor het prioriteren. Hij bespreekt acht mogelijke criteria en komt tot de conclusie dat vier daarvan bruikbaar zijn voor het gestelde doel.

Rademaker laat vier criteria afgefallen:

- Mate van proportionaliteit. Proportionaliteit zegt immers iets over de geschiktheid, noodzakelijkheid en adequaatheid van een dierproef in de context van schadeberokkening, en verwijst daarbij naar *andere* criteria;
- De ontwikkelingsstatus van alternatieven zegt iets over laaghangend fruit, maar niet over waar op ethische gronden meer inspanning gewenst is. Dat neemt niet weg dat ‘nabije’ oplossingen aandacht verdienen⁶;
- Het aantal dieren (dierproeven) is van belang in relatie tot andere karakteristieken zoals het doel van de proef of het ongerief dat de dieren ondervinden, maar is minder geschikt als criterium op zich⁷;
- De gebruikte diersoort kan relevant zijn in relatie tot de validiteit van de dierproef en daarmee de transleerbaarheid en tot het vermogen tot ervaren van ongerief. Deze beide criteria worden in het rapport van Rademaker al meegenomen als zelfstandige criteria. Onder filosofen en ethici is er geen consensus over of sommige diersoorten principieel beschermwaardiger zijn dan andere. Echter, de EU-Richtlijn verbiedt het gebruik van een aantal primatensoorten (Artikel 8, 2010/63/EU zie noot 2) en stelt voor andere primatensoorten strengere regels dan voor de overige diersoorten⁸.

Vervolgens werkt Rademaker de vier volgens hem geschikte criteria nader uit, te weten de mate van transleerbaarheid, de relevantie van het doel van de proef, de mate van het ongerief en de mate van integriteitsaantasting.

⁶ Reactie NCad in hoofdstuk 3: Reflectie van het NCad op het rapport.

⁷ Reactie NCad in hoofdstuk 3: Reflectie van het NCad op het rapport.

⁸ Reactie NCad in hoofdstuk 3: Reflectie van het NCad op het rapport.



Transleerbaarheid

Transleerbaarheid (vertaalbaarheid) gaat over de vraag in hoeverre de data verkregen uit een dierexperiment vertaalbaar zijn naar de mens of het doeldier. Transleerbaarheid heeft met validiteit te maken: een diermodel is valide wanneer het een goed model vormt van de ziekte, aandoening of fenomeen waar het model voor staat en meer specifiek wanneer onderliggende mechanismen overeenkomen.

De transleerbaarheid of vertaalbaarheid van diermodellen is al geruime tijd een punt van aandacht in het biomedisch onderzoek. Uit een recente Systematic Review (Leenaars e.a.⁹) blijkt dat de voorspellende waarde van bevindingen in proefdieren naar fase I/II *clinical trials* enorm kan variëren.

In het rapport van Rademaker wordt uitgebreid ingegaan op verschillende aspecten van validiteit. In de literatuur worden tenminste drie aspecten van validiteit onderscheiden:

- Voorspellende validiteit zegt iets over de mogelijkheid om voorspellingen te doen met betrekking tot een menselijk of dierlijk fenomeen op basis van het presteren van een (dier)model.
- Fenotypische validiteit heeft betrekking op de mate van overeenkomst tussen het (dier)model en de onderzochte ziekte of aandoening voor een zo breed mogelijk scala aan symptomen.
- Constructvaliditeit zegt iets over de mate waarin de mechanismen die ten grondslag liggen aan de ziekte of aandoening overeenkomen tussen het (dier)model en het doel¹⁰.

Dan zijn er nog de interne validiteit (betrouwbaarheid en reproduceerbaarheid) en externe validiteit (brede toepasbaarheid; generalisatie) deze worden eerder bepaald door een combinatie van de eigenschappen van het model en de experimentele context waarin het model wordt gebruikt. Bij *reproduceerbaarheid* gaat het om de vraag of het model consistente en betrouwbare resultaten oplevert. Bij *generalisatie* draait het om de vraag of de resultaten behaald met het diermodel in de onderzochte populatie of de foklijn te extrapoleren zijn naar het doelorganisme.

⁹ Leenaars, Cathalijn H. C., Carien Kouwenaar, Frans R. Stafleu, André Bleich, Merel Ritskes-Hoitinga, Rob B. M. De Vries, and Franck L. B. Meijboom. *Animal to Human Translation: A Systematic Scoping Review of Reported Concordance Rates*. *Journal of Translational Medicine* 2019 Vol. 17 Issue 1 Pages 1–22. <https://doi.org/10.1186/s12967-019-1976-2>.

¹⁰ Bijvoorbeeld de mens als het diermodel wordt gebruikt om onderzoek te doen naar een humane ziekte.

Ook wordt onderscheid gemaakt tussen convergerende validiteit en onderscheidende validiteit. Convergerende validiteit geeft aan dat maten die hetzelfde meten aanzienlijk met elkaar horen te correleren. Onderscheidende validiteit is precies het tegenovergestelde.

De validiteitscriteria moeten worden toegepast op onderzoek waarbij transleerbaarheid een doel is, zoals alle regulatorisch en preklinisch translationeel onderzoek. Bij fundamenteel onderzoek gaat het met name om het ontrafelen van biologische mechanismen met als vervolgvraag in hoeverre het gevonden mechanisme meer of minder algemeen geldend is. De resultaten van fundamenteel onderzoek dragen bij aan de beantwoording van de validiteitsvraag en dus in hoeverre datgene wat onderwerp van studie is gemodelleerd kan worden door bijvoorbeeld een bepaald cel- of orgaansysteem of diermodel. Rademaker zegt dat een deel van wat onder fundamenteel onderzoek wordt geschaard, ook translationeel van aard is.

Rademaker stelt vervolgens de vraag op welke manier transleerbaarheid als criterium gehanteerd kan worden.

De transleerbaarheid van een diermodel hangt volgens Rademaker af van de mate waarin je met zekerheid kunt stellen dat een gevonden causale relatie niet uitgelegd kan worden door andere factoren dan de factor die het doel van het onderzoek was (de interne validiteit) en de brede toepasbaarheid van de resultaten (de externe validiteit). Beide validiteiten zeggen iets over de mate waarin het model en de manier waarop het in het experiment gebruikt is, de veronderstelde achterliggende mechanismen van een biologisch of etiologisch fenomeen in de mens of het doeldier simuleert.

Rademaker bespreekt ook een viertal hulpmiddelen (instrumenten) om de transleerbaarheid te bepalen. Meerdere onderzoekers hebben namelijk een scoresysteem ontwikkeld om een diermodel te kunnen evalueren op validiteit. Rademaker beveelt aan om in samenspraak met het veld een keuze te maken tussen deze scoresystemen.

Tenslotte stelt Rademaker de vraag of translatie van de ene diersoort naar de andere überhaupt mogelijk is. Er zijn voorbeelden van succes en van falen.

Evident is dat diersoorten en ook foklijnen binnen diersoorten van elkaar verschillen. Vraag is of dat ook betekent dat een dier alleen model kan staan voor dieren van dezelfde soort (of foklijn) en een diermodel hooguit een hypothese kan opleveren voor een andere



diersoort. In dat laatste geval weet je pas of een model transleerbaar is *nadat* de resultaten zijn vergeleken met resultaten in het doeldier of de mens.

Ook legt Rademaker de vraag op tafel of dit laatste betekent dat het ontwikkelen van nieuwe, nog niet beproefde, diermodellen daardoor ethisch meer betwistbaar is dan het gebruik van beproefde modellen. Zijn antwoord is dat het niet altijd nodig is dat het causale mechanisme exact hetzelfde is. Bovendien zijn ook hypothese-stellende modellen van waarde.

Dierproeven waarbij diermodellen gebruikt worden die een lage mate van transleerbaarheid hebben, komen als eerste in aanmerking om vervangen te worden, mede afhankelijk van de beoordeling op andere criteria, zoals de mate van relevantie van het doel van een dierproef.

Relevantie van het doel

Proefdieronderzoek wordt uitgevoerd met een doel. Het directe doel van een concrete dierproef, bijvoorbeeld het ontrafelen van een mechanisme, kan worden onderscheiden van het uiteindelijke doel dat verder in de toekomst ligt en waarvoor meerdere experimenten en onderzoeken nodig zijn.

Sommige van die uiteindelijke doelen staan ter discussie. Zo vragen velen zich af of routinematige veiligheidstesten van chemicaliën op dieren niet anders kunnen. Onderzoek gericht op de gezondheid van mens en dier scoort daarentegen hoog als doel. Ook binnen die categorieën zijn sommige doelen beter te rechtvaardigen dan andere, omdat sommige aandoeningen ook op een andere wijze kunnen worden aangepakt of dragelijk gemaakt. Medisch onderzoek is vaak gericht op genezing, terwijl preventie ook een optie is, denk bijvoorbeeld aan voedingsinterventies ter versterking van het immuunsysteem. Bij onderzoek ten behoeve van de veehouderij maakt het uit of dat gericht is op de gezondheid en het welzijn van de dieren of op verduurzaming van de veehouderij in plaats van op productieverhoging (Raad voor Dieraangelegenheden, 2018).

Een eenvoudige hiërarchie is niet aan te brengen. Rademaker geeft aan dat het behulpzaam is onderscheid te maken tussen basale, serieuze en perifere belangen (VanDeVeer¹¹):

- Basale belangen zijn specifieke condities of goederen die nodig voor mens en dier om “minimaal adequaat” functioneren mogelijk te maken. Rademaker past dit onderscheid toe op de relevantie van doelen van onderzoek. Een onderzoek met een basaal doel is bijv. gericht op de preventie, genezing en het draagbaar maken van levensbedreigende ziekten en ziekten die ernstig belemmerend zijn voor de kwaliteit van leven van mens en dier en het behoud van diersoorten en ecosystemen.
- Serieuze belangen zijn goederen of condities die van belang zijn voor het welzijn van mens en dier. Bij serieuze doelen gaat het om bijvoorbeeld onderzoek gericht op duurzame vormen van veehouderij en veiligheidstesten van voedingsmiddelen.
- Bij perifere belangen gaat het om goederen en condities die het leven aangenaam maken. Denk aan decoraties, een deel van de verzorgingsartikelen, maar ook aan bont. In de huidige moderne samenleving zouden bijv. productiedoelen in de veehouderij perifeer genoemd kunnen worden.

Als doelen van onderzoek kunnen al deze doelen legitiem zijn, maar op het moment dat proefdieren nodig zijn om het doel te bereiken, zou een hiërarchie moeten worden aangebracht.

¹¹ VanDeVeer, Donald. (1995) *Interspecific Justice*. In: *People, Penguins, and Plastic Trees: Basic Issues in Environmental Ethics*, edited by Christine Pierce and Donald VanDeVeer, 85–99. Belmont, CA: Wadsworth Publishing Company.



Ongerief

Onderzoek met dieren geldt pas als een dierproef als er sprake is van ongerief boven een ondergrens. De Europese Richtlijn 2010/63/EU (zie noot 2) maakt het volgende onderscheid en benoemt daarbij twee drempels:

- Onder de ongerief drempel; het ongerief is minder dan het ongerief veroorzaakt door het inbrengen van een injectienaald volgens goed diergeneeskundig gebruik. Dan is geen sprake van een dierproef;
- Terminaal; het dier wordt onder narcose gebracht; alle handelingen worden onder narcose uitgevoerd; het dier komt niet meer bij;
- Mild ongerief;
- Matig ongerief;
- Ernstig ongerief;
- Zeer ernstig ongerief, d.w.z. ernstige pijn, lijden of angst die waarschijnlijk lang zal duren en niet kan worden verzacht, is alleen bij hoge uitzondering en na expliciete toestemming van de minister toegestaan en valt buiten de normale procedures. In Nederland en in Europa zijn geen dierproeven gemeld met zeer ernstig ongerief.

Voor het vervangen van dierproeven op grond van ongerief zou volgens Rademakers de focus allereerst gericht moeten zijn op dierproeven met diermodellen die ernstig ongerief veroorzaken. In Nederland gaat het om ca. 1% van alle dierproeven¹² en in Europa om ca. 11%¹³. Binnen de EU (incl. Noorwegen) vormt *batch potency testing* van vaccins en andere biologische producten de belangrijkste bron van ernstig ongerief, op afstand gevolgd door onderzoek naar het zenuwstelsel en diagnostische testen.

Integriteitsaantasting

Begin jaren '90, rond de opkomst van genetisch gemodificeerde (GM) dieren - waaronder stier Herman - ontstond in Nederland discussie over genetische modificatie. Mensen ervaren blijkbaar moreel ongemak met GM-dieren ook als er geen sprake is van aantasting van de gezondheid of het welzijn. Daarom werd toen gezocht naar een begrip dat dit kon duiden. In eerste instantie werd gesproken over 'heelheid' en 'natuurlijkheid'. Later werd daar het begrip 'integriteit' voor gebruikt. Met dit begrip konden ook andere 'aantastingen' van dieren worden geduid die niet direct met welzijn of gezondheid te maken hebben.

¹² p. 25 <https://www.nvwa.nl/documenten/dier/dierenwelzijn/zo-doende/publicaties/zo-doende-2020-jaaroverzicht-dierproeven-en-proefdieren>

¹³ p. 32 https://ec.europa.eu/environment/chemicals/lab_animals/pdf/SWD2019_Part_A_and_B.pdf

Naast genetische modificaties kan dat ook spelen bij het doorfokken van dieren op bepaalde kenmerken of het verwijderen van snavels, hoorns of voortplantingsorganen.

Volgens de definitie van Rutgers en Heeger¹⁴ gaat integriteit over de heelheid van het dier, de soort-specifieke balans van het dier en het vermogen om zich onafhankelijk te handhaven in een voor de soort passende omgeving.

Als voorbeeld van de aantasting van de *heelheid* van het dier noemen Rutgers en Heeger het onthoornen van vee. Een voorbeeld dat Rutgers en Heeger geven met betrekking tot de *aantasting van de soort-specifieke balans* van het dier is het fokken op snelle spieraanzet in vleeskuikens, waardoor deze dieren al vóór de slachtleeftijd niet goed meer kunnen lopen en soms plotseling sterven. Bij deze vleeskuikens is de stofwisseling uit balans. Een voorbeeld van de aantasting van *het vermogen om zich onafhankelijk te handhaven in een voor die soort passende omgeving*, is het fokken van dikbilkoaien als het Belgisch Witblauw. Deze dieren hebben een buitengewone sterke spiergroei in combinatie met een smal geboortekanaal. Hierdoor kunnen de kalveren niet op een natuurlijke wijze geboren worden. Deze dikbilkoaien zijn bij de bevalling altijd aangewezen op een keizersnede.

Rademaker gaat vervolgens in op de ontvangst van de definitie van integriteit van Rutgers en Heeger door latere auteurs. Hij constateert dat twee interpretaties van *heelheid* gehanteerd worden, namelijk:

- Intervenieren in het dierlijk lichaam. Dan valt amputatie, bijv. de teenknip, binnen de definitie en valt een fokresultaat, bijv. naakte (vachtloze) muizen, erbuiten. Dit is een smalle opvatting van heelheid.
- Het moeten missen van iets wat de dieren van nature, als soort, fysiek hebben (bijv. een vacht, horens). Dit is de brede opvatting van integriteit. Hier valt ook een fokresultaat binnen. De lichamelijke configuratie van dieren voorziet het dier van allerlei vermogens. Soms worden spontaan dieren zonder vacht geboren en wordt hiermee vervolgens een foklijn gecreëerd (naakte katten; naakte muizen). Deze dieren zijn minder goed in staat om hun lichaamstemperatuur te handhaven en zijn aangewezen op een kunstmatige omgeving.

¹⁴ Rutgers, B. and Heeger, R. (1999) *Inherent worth and respect for animal integrity*. In: *Recognizing the intrinsic value of animals: beyond animal welfare*, Van Gorcum, Assen, The Netherlands.



De brede opvatting van heelheid betreft de capaciteiten of *vermogens* van dieren in de opvatting van integriteit. Dit is ook in lijn met de oorspronkelijk definitie van Rutgers en Heeger die immers spreken over het *vermogen* om zich onafhankelijk te handhaven in een voor de soort passende omgeving.

Rademaker constateert dat door de tijd heen deze focus op dierlijke vermogens sterker is geworden. Rademaker bekritiseert echter de sterke focus op onafhankelijkheid in de definitie van integriteit. Liever spreekt hij van *soort-specifieke vermogens* van dieren, zoals knaaggedrag soort-specifiek is voor knaagdieren. Het gaat hierbij om de mogelijkheid van dieren om soorteigen gedrag te vertonen.

Volgens Rademaker doet de brede opvatting van integriteit het meest recht aan het morele onbehagen van mensen met bijvoorbeeld naakte muizen of dikbilkoaien.

Vervolgens gaat Rademaker in op enkele vragen die door filosofen in het debat van de jaren '90 zijn opgeworpen en besproken.

Is er alleen sprake van een integriteitsaantasting wanneer de aantasting verricht wordt ten bate van menselijke gebruiksdoeleinden?

Oftewel doet de intentie ertoe als bijv. de staart van een hond wordt gecoupeerd? Dit zou erop kunnen wijzen dat integriteit geen eigenschap of biologisch kenmerk van het dier is (zoals welzijn) maar een duiding van het morele onbehagen van mensen bij bepaalde handelingen in een bepaalde context. Rademaker betoogt echter dat ook als een ingreep ten behoeve van het dier zelf is, dit wel degelijk een integriteitsaantasting kan zijn. De amputatie van een staart tast de heelheid van het dier aan, maar kan veterinaire noodzakelijk zijn. Welzijn is dan belangrijker. Rademaker concludeert dat integriteit net als het begrip welzijn een evaluatief (waarderend) begrip is en ook een empirische kant heeft die kan worden vastgesteld aan de hand van de gevolgen voor het dier.

Is integriteit een absoluut of een relatief begrip?

Integriteit lijkt een absoluut begrip te zijn: de integriteit van het dier wordt aangetast of niet. Rademaker werpt echter de vraag op of er ook *gradaties* kunnen zijn in integriteit (vergelijkbaar met welzijn). Verwijzend naar Brom¹⁵ stelt hij dat dit inderdaad het geval kan

zijn. Een dier dat in zijn integriteit is aangetast (bijv. staart gecoupeerd) kan op een later moment nog verder in zijn integriteit worden aangetast (ook nog oren gecoupeerd).

Gaat integriteit alleen over individuele dieren of ook over foklijnen of soorten?

Deze vraag speelt vooral bij genetische modificatie (GM). GM vindt plaats in de vroegembryonale fase en hoeft niet altijd gevolgen voor het fenotype te hebben. Vraag is dus wanneer gesproken kan worden van integriteitsaantasting. Vermoedelijk is dat niet het geval zolang er geen fenotypische verandering detecteerbaar is. Als veranderde genetische eigenschappen worden overgedragen op het nageslacht zijn die eigenschappen voor het nageslacht aangeboren. Is de integriteit van het nageslacht dan aangetast? Het hangt ervan af wat je als referentiepunt neemt, het wild-type dier, de foklijn (of het ras) of de waarneembare kenmerken of eigenschappen van het dier (het fenotype). Als de foklijn als referentie wordt genomen, dan tellen de veranderingen in de huidige generatie niet mee. Die worden dan voor lief, als vanzelfsprekend genomen. Dat sluit niet goed aan bij het morele ongemak dat mensen kunnen ervaren bij zo'n fokresultaat. In de beide andere gevallen kan wel sprake zijn van fok met integriteitsaantasting, naar analogie van fok met welzijnsaantasting.

Is integriteit bruikbaar als toetsingscriterium in de praktijk?

Om van integriteitsaantasting een toetsingscriterium te kunnen maken zou je de aantasting moeten kunnen inschatten, zoals je dat ook met welzijnsaantasting (ongerief) doet. Dat leidt tot de vraag of er gradaties zijn in integriteitsaantasting. Ingeval van de smalle definitie (lichamelijke heelheid) kan alleen onderscheid gemaakt worden naar het aantal aantastingen (bijv. een vasectomie én een teenknip). Ook bij de brede definitie speelt het aantal aantastingen een rol, maar daarnaast speelt daar ook de aantasting van soort-specifieke vermogens. De soort-specifieke vermogens kunnen in meer of mindere mate worden aangetast. Sowieso gaat het, net als bij welzijn, om het ontwikkelen van een intersubjectief gekwantificeerde inschatting.

Operationalisering van het brede begrip van integriteit is niet eenvoudig omdat dit veel kennis vraagt van de soort-specifieke vermogens van dieren en de aantasting daarvan. Omdat voor de inschatting van integriteitsaantasting nog geen handvatten bestaan, adviseert Rademaker om een werkgroep op te zetten die een voorstel ontwikkelt voor een classificatie van aantasting van integriteit.

¹⁵ Brom, Frans W.A. (1997). *Onherstelbaar verbeterd: Biotechnologie bij dieren als een moreel probleem*. Assen: Van Gorcum.



Dit zou bijvoorbeeld kunnen door een classificatie in termen van licht, matig en ernstig te koppelen aan de frequentie van ingrepen, de mate waarin een specifiek dierlijk vermogen aangetast wordt, en het aantal vermogens dat aangetast wordt.

Concluderend

De opdracht aan Rademaker was om criteria te ontwikkelen om prioriteiten te kunnen stellen voor diermodellen waarvoor met prioriteit alternatieven ontwikkeld zouden moeten worden. Rademaker komt tot vier criteria. Rademaker stelt voor om met voorrang te zoeken naar mogelijkheden voor het vervangen van diermodellen die:

- Een lage mate van transleerbaarheid hebben en/of
- Perifere doelen dienen en/of
- Ernstig ongerief veroorzaken en/ of
- De integriteit (soort specifieke vermogen) van de dieren in ernstige mate aantasten.



2. Zienswijze NCad

Het rapport van Rademaker is een heldere uiteenzetting van criteria op grond waarvan het wenselijk is om bepaalde diermodellen uit te faseren en waar mogelijk te (gaan) vervangen door alternatieven in de brede zin, de zogenaamde NAMs (zie noot 1). Vooral de uitwerking van de begrippen transleerbaarheid en integriteitsaantasting vormen een verrijking voor het debat over de aanvaardbaarheid van specifieke diermodellen. Ze dragen ook bij aan de aanscherping van de schade-batenafweging.

Rademaker bespreekt acht criteria en concludeert dat vier ervan geschikt zijn om te gebruiken bij het prioriteren van diermodellen die verworpen of vervangen zouden moeten worden. Hij bespreekt elk criterium afzonderlijk. In de praktijk wordt een schade-batenafweging gemaakt en gaan de criteria met elkaar interfereren. Soms kan blijken dat een model dat je op grond van één van de criteria zou willen verwerpen, op grond van een ander criterium toch zou moeten worden gehandhaafd bijvoorbeeld als daarvoor nog geen dierproefvrij alternatief voorhanden is en het er niet naar uitziet dat deze op redelijk korte termijn ontwikkeld kan worden en het belang van de proef groot is. Voor die gevallen benadrukt het NCad het belang van de V's van Vermindering en Verfijning.

In deze Zienswijze reflecteert het NCad op het rapport van Rademaker (bijlage 3), en geeft het een korte aanvulling op het criterium integriteit en een nadere precisering van het criterium transleerbaarheid. Vervolgens wordt besproken wat deze criteria betekenen voor de schade-batenanalyse (de ethische afweging) en geeft het NCad een handreiking voor hoe de criteria die Rademaker heeft ontwikkeld, kunnen bijdragen aan de schade-batenafweging door IvDs, DEC's en de CCD en de discussie daarover. De laatste paragraaf is gericht op beleidsmakers en geldverstrekkers en geeft advies over hoe zij diermodellen kunnen prioriteren voor vervanging en richting kunnen geven aan de ontwikkeling van NAMs. Tenslotte stelt het NCad nog een aantal vervolgstappen voor.



Reflectie van het NCad op de criteria van Rademaker

Gebruikte diersoort

Rademaker betoogt dat het op ethische gronden lastig is om onderscheid te maken tussen diersoorten en dat er onder filosofen geen consensus is over deze kwestie. Hij ziet om die redenen af van een nadere uitwerking van diersoort als criterium. Wetgeving wordt in de regel opgevat als de neerslag van de normen en waarden in een samenleving¹⁶. De EU-Richtlijn (artikel 8, 2010/63/EU, zie noot 2) en de Wet op de dierproeven (artikel 1, lid 5 en artikel 10e, Wod, zie noot 4) maken wel onderscheid tussen diersoorten. Een proef met een ongewerveld dier is geen dierproef, kopotigen uitgezonderd. Voor proeven met ongewervelde dieren (kopotigen uitgezonderd) is dus geen projectvergunning nodig of toezicht geregeld. Aan primaten is in de EU-Richtlijn en de Wod een bijzondere status toegekend met strengere eisen voor het gebruik voor wetenschappelijke doeleinden ervan. In het verlengde van de EU-Richtlijn en de Wod zou daarom het vervangen van diermodellen met niet-humane primaten prioriteit moeten hebben. Dit onderscheid tussen diersoorten is de wettelijke neerslag van de notie dat het bezwaar tegen het gebruik van bepaalde diersoorten toeneemt naarmate die diersoorten meer overeenkomsten vertonen met de mens.

Rademaker betoogt dat vanuit ethische overwegingen vraagtekens bij dit aspect van de Europese Richtlijn kunnen worden gezet. Daar sluit het NCad bij aan. Ook het NCad meent dat het verschil in gevoeligheid c.q. de mate van ongerief die het dier ervaart, gebaseerd op objectiveerbare criteria, een belangrijker criterium is dan diersoort. Een bepaalde handeling kan meer impact hebben op de ene diersoort of binnen de diersoort op de ene foklijn dan op de andere. In geval er geen objectiveerbare criteria zijn, geldt het voorzorgsprincipe¹⁷.

¹⁶ N.B.: Normen en waarden zijn steeds, maar geleidelijk, in beweging en de wetgeving komt daar vertraagd achteraan. Dat er een debat is onder ethici kan erop wijzen dat zo'n verandering gaande is in de samenleving, waarbij dieren als steeds gelijkwaardiger worden gezien aan elkaar en aan mensen.

¹⁷ Het voorzorgsprincipe is een moreel en politiek principe dat stelt dat als een ingreep of een beleidsmaatregel ernstige of onomkeerbare schade kan veroorzaken aan de samenleving of het milieu, de bewijslast ligt bij de voorstanders van de ingreep of de maatregel als er geen wetenschappelijke consensus bestaat over de toekomstige schade.

De ontwikkelingsstatus van alternatieven

Rademaker vindt de ontwikkelingsstatus van een dierproefvrij alternatief voor een bepaald diermodel wel van strategisch belang, maar ziet die ontwikkelingsstatus niet als een ethische overweging om een diermodel te prioriteren voor uitfasering. In zijn rapport focust hij op de diermodellen die het meest problematisch zijn, niet op 'laaghangend fruit'. Het NCad kan daar deels in meegaan. Ieder model heeft goede en minder goede eigenschappen c.q. beperkingen. Het is aan de wetenschap die te onderkennen en te streven naar continue verbetering van de modellen die gebruikt worden en om nieuwe, betere modellen te ontwikkelen. Daarbij moeten keuzes gemaakt worden en gaat het om het prioriteren van middelen waaronder die voor de ontwikkeling en validatie van NAMs. De ontwikkelingsstatus en beschikbaarheid van een NAM is dan wel degelijk een criterium. Zeker als het een alternatief is voor een type dierproef waarbij veel dieren worden gebruikt (zie hierna).

Uiteindelijk beslist de vergunningverlener, de CCD, of een dierproef wel of niet wordt toegestaan en onder welke voorwaarden. De aanvrager moet overtuigend bewijs leveren dat met het voorgestelde (dier)model het antwoord op de wetenschappelijke vraag redelijkerwijs kan worden gevonden en dat er geen dierproefvrij alternatief of een model met minder dieren en/of minder ongerief is waarmee dezelfde wetenschappelijke doelen kunnen worden behaald. Beschrijving van de goede en minder goede eigenschappen van de verschillende modellen en welke inspanningen er worden gepleegd om modellen verder te ontwikkelen, waaronder ook de NAMs, moet daarom deel uitmaken van de aanvraag. Als een onderzoeksteam van plan is om dieren te gebruiken in fundamenteel onderzoek, moet het idealiter duidelijk maken hoe de onderzoeksresultaten naar verwachting zullen bijdragen aan toekomstige studies in (verschillende stadia van) toegepast onderzoek (zie hieronder bij 'transleerbaarheid').

Aantallen dieren en diermodellen

Met de afwijzing door Rademaker van het criterium aantallen dieren en diermodellen kan het NCad niet meegaan. Niet voor niets wordt in de Wod en de achterliggende EU-Richtlijn gesproken over 3V's, namelijk vervangen, verfijnen én verminderen. 'Aantal als criterium' betreft dan niet zozeer het aantal dieren per individuele projectvergunningaanvraag waar het minimumaantal dieren moet worden gemotiveerd, (V van Vermindering), maar eerder het cumulatieve aantal dieren van een diermodel dat vaak en breed wordt ingezet in bepaalde onderzoeksgebieden. Daartoe behoren ook de diermodellen, die jaar in jaar uit



op grote schaal worden gebruikt voor het testen van de veiligheid en werkzaamheid van chemicaliën, geneesmiddelen en vaccins. Aanpassing van de regelgeving rondom veiligheid zou daar verandering in kunnen brengen. Het NCad wil daarom ook het aantal gebruikte dieren meenemen als criterium voor het prioriteren van diermodellen voor vervanging in de brede zin of voor verfijning.

Het concept transleerbaarheid van onderzoeksresultaten in het fundamentele onderzoek

De relevantie van het criterium transleerbaarheid van onderzoeksresultaten van een diermodel naar bijvoorbeeld de mens voor het als translationeel gelabelde onderzoek is eenvoudiger te beargumenteren dan voor het basale of fundamenteel onderzoek waarin, zoals eerder gesteld, het met name gaat om het ontrafelen van biologische mechanismen met als vervolgvraag in hoeverre het gevonden mechanisme meer of minder algemeen geldend is c.q. translationele waarde heeft¹⁸. Het werkdocument over projectevaluatie en beoordeling achteraf van de Europese Commissie geeft een aantal handvatten hoe invulling gegeven kan worden aan de schade – batenanalyse in het kader van de ethische toets van projectvergunningaanvragen (zie noot 18). Het vermeerderen van kennis als baten in de afweging wordt daarin gezien als voldoende rechtvaardiging voor de uitvoering van dat onderzoek mits voldoende wetenschappelijk onderbouwd en de koppeling wordt gemaakt met de verspreiding van resultaten (rekening houdend met intellectueel eigendom rechten) en mogelijke baten op langere termijn waarin transleerbaarheid steeds belangrijker wordt. Deze aspecten moeten geadresseerd worden bij het formuleren van de doelstelling van het project waarin ook beschreven moet worden wat de huidige stand van de kennis is waarop het project beoogt voor te bouwen (zie ook bijlage 4). Laat onverlet dat ook voor fundamenteel onderzoek geldt dat ingezet moet worden op de ontwikkeling en inzet van NAMs.

¹⁸ In bijlage 4 zijn de teksten overgenomen over fundamenteel onderzoek uit het *Werkdocument over projectevaluatie en beoordeling achteraf*, dat is samengesteld door de werkgroep van Nationale instanties die bevoegd zijn voor de tenuitvoerlegging van Richtlijn 2010/63/EU.

Schade-batenanalyse en de criteria van Rademaker

De schade-batenanalyse gaat in de eerste plaats over de afweging van de relevantie van de dierproef en het ongerief dat de proefdieren wordt aangedaan. Deze afweging wordt door de CCD na advies van een DEC gemaakt op basis van een projectvergunningaanvraag¹⁹. De projectvergunning is opgesteld door de aanvragende onderzoeker daarbij geadviseerd door de IvD. Deze werkwijze zorgt ervoor dat de kwaliteit van de vergunningaanvragen hoog is en dierwelzijnsaspecten gewaarborgd zijn.

Het Werkdocument voor Projectevaluatie van de Europese Commissie (zie noot 18), dat door de CCD en de DEC's als leidraad wordt gebruikt, beveelt aan om de aangepaste Bateson-kubus (Aanhangsel 2 van het Werkdocument) te gebruiken voor het afwegingskader (figuur 1). Deze kubus heeft drie assen één voor relevantie van het doel (de baten), één voor het ongerief (de schade) en een derde voor de waarschijnlijkheid dat de baten ook gerealiseerd worden. Uitzicht op transleerbaarheid maakt onderdeel uit van deze waarschijnlijkheid en maakt daarom in de praktijk onderdeel uit van de overwegingen in de schade-batenanalyse.

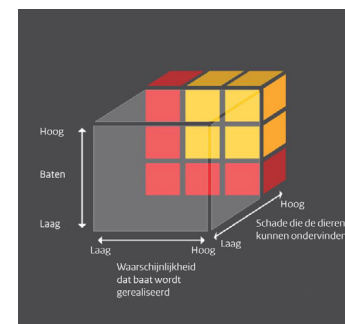
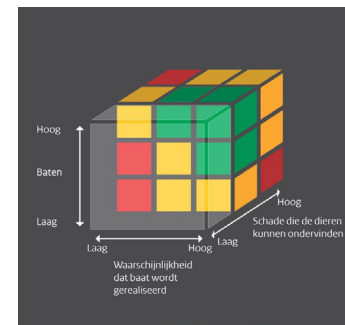
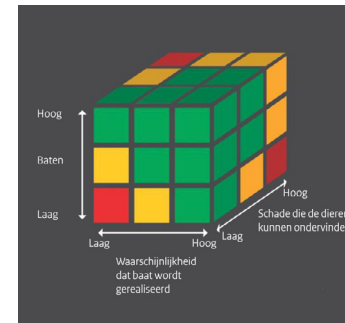
Rademaker bespreekt in zijn rapport dat ook integriteitsaantasting onderdeel zou moeten uitmaken van de ethische afweging. Toevoeging van integriteitsaantasting aan de schade-batenafweging kan het begrip 'schade' verrijken en bijdragen aan verfijning van het afwegingskader. Dit sluit aan bij de intentie van de wetgever. De Centrale Commissie Dierproeven (CCD) werkt al met dit verrijkte schadebegrip, zoals blijkt uit het Ethisch Toetsingskader waarmee de CCD werkt (zie noot 19).

Rademaker laat ook zien dat het begrip integriteitsaantasting nog weinig geoperationaliseerd is. Hij noemt het begrip aantasting van soort-specifieke vermogen als een mogelijke start van het gesprek.

¹⁹ <https://www.centralecommissiedierproeven.nl/binaries/ccd/documenten/formulieren/16/6/6/ethisch-toetsingskader/praktische-handreiking-ethisch-toetsingskader-versie-september-2019.pdf>

Er zijn nog geen handvatten voorhanden op grond waarvan je onderscheid zou kunnen maken in milde, matige of ernstige integriteitsaantasting. De discussie die onder ethici en filosofen speelde in de jaren '90, geeft daar ook geen handvatten voor. Rademaker adviseert om een start te maken met die operationalisering. Net als in het verleden voor welzijnsaantasting is gebeurd, zou toegegroeid kunnen worden naar een intersubjectieve overeenstemming over wanneer sprake is van milde, matige of ernstige integriteitsaantasting. De enige weg om tot een dergelijke indeling te komen is te beginnen met het gesprek erover, dus: als mogelijk sprake is van integriteitsaantasting, probeer dan te benoemen wat die integriteitsaantasting inhoudt en ga na hoe dat mee zou moeten wegen in de ethische afweging. Een werkgroep, zoals Rademaker aanbeveelt, zou zich kunnen buigen over concrete casussen vanuit het veld.

Figuur 1: Door EU-werkgroep aangepaste Bateson kubus (zie noot 18)





Prioriteren in de ontwikkeling van NAMs

Bij het doen van onderzoek staat de onderzoeksvraag centraal. De onderzoeksstrategie om die vraag te beantwoorden vereist een zorgvuldige keuze voor het te gebruiken model of combinatie van modellen, waaronder NAMs (zie noot 1). De vier criteria beschreven in het rapport van Rademaker wijzen de richting waar met prioriteit NAMs (in de brede definitie die eerder is gegeven) moeten worden ontwikkeld. Het NCad voegt aan de vier criteria van Rademaker de criteria 'aantal dieren' en ontwikkelingsstatus van dierproefvrije alternatieven toe. Echter zolang er ook nog dierproeven gedaan moeten worden, blijft inzet op de twee andere V's van vermindering en verfijning relevant. Zo zet het NCad zich in een apart project in voor het voorkómen van ernstig ongerief. De 'Routekaart Ernstig Ongerief' van de RSPCA²⁰ is vertaald naar het Nederlands. Het NCad organiseert samen met het IvD-Platform, de RSPCA en de IvD Leiden een symposium over dit onderwerp in het najaar van 2023.

Volgende vragen zijn dan of tussen die criteria ook nog geprioriteerd kan worden, hoe criteria elkaar mogelijk versterken en wat die criteria betekenen voor het onderzoek en de besteding van middelen voor de ontwikkeling van alternatieven.

Hierboven hebben we gezien dat het in de rede ligt om in elk geval de schade-batenanalyse nader te bekijken. Als dierproeven weinig relevantie hebben omdat ze een 'perifeer' doel dienen of omdat de kennis die de dierproef oplevert weinig relevant is voor het gestelde doel bijvoorbeeld omdat de transleerbaarheid van het diermodel laag is. Dan ligt het niet uitvoeren van het onderzoek voor de hand. Mochten de onderzoekers toch de vraag willen beantwoorden, dan zouden zij moeten uitwijken naar een dierproefvrij alternatief of een diermodel met heel weinig impact op de dieren en als deze er niet zijn, dan zouden deze ontwikkeld moeten worden.

Bij 'basale' of 'serieuze' doelen ligt dat anders en krijgt de afweging van de baten tegen de schade (ongerief en integriteitsaantasting) een prominentere rol. Zeker als het ongerief ernstig is of als de integriteit van de dieren ernstig wordt aangetast, dan zou voor deze onderzoeksvragen met voorrang gewerkt moeten worden aan de ontwikkeling van alternatieven in de brede zin, waaronder verfijnde modellen.

Ook voor veel gebruikte diermodellen, waar door het frequente gebruik, in de loop van de tijd veel dieren nodig zijn, is het wenselijk om met voorrang NAMs beschikbaar te krijgen. Per projectaanvraag leiden deze - vaak voorgeschreven - modellen niet tot een verwerping, maar het is wenselijk als deze modellen niet meer gebruikt hoeven te worden.

Tenslotte zou ook prioriteit moeten worden gegeven aan het verder ontwikkelen en valideren van NAMs die al in ontwikkeling zijn, in het bijzonder als deze een alternatief vormen voor veel gebruikte modellen.

Conclusies en aanbevelingen

- Om te zorgen dat de middelen voor de ontwikkeling en implementatie van NAMs en de vervanging van diermodellen optimaal worden ingezet, is prioritering wenselijk. Het rapport van Rademaker en deze zienswijze kunnen daarbij helpen. Het NCad zal beide documenten onder de aandacht brengen van de belangrijkste subsidieverstrekende organisaties: NWO, ZonMw, de gezondheidsfondsen en ProefdierVrij.
- Het NCad vraagt de minister van LNV om zijn collega van OC&W op de hoogte te stellen van deze zienswijze en daarbij de suggestie aan de hand te doen om samen met het ministerie van LNV en andere ministeries aanvullende middelen voor de ontwikkeling en implementatie van NAMs vrij te maken. Ook kan de minister van OC&W worden verzocht om aan de hand van deze zienswijze en het rapport in gesprek te gaan met NWO en ZonMw waar deze middelen met prioriteit zouden kunnen worden ingezet.
- Het verdient aanbeveling om het begrip transleerbaarheid uit te werken en toe te voegen aan de as 'waarschijnlijkheid dat de baten worden gerealiseerd' van de aangepaste Bateson kubus. Daarom raadt het NCad de minister van LNV aan om te bevorderen dat transleerbaarheid onderdeel gaat uitmaken van het 'Werkdocument over projectevaluatie en beoordeling achteraf' (zie noot 18) en op termijn wordt meegenomen bij de evaluatie van Richtlijn 2010/63/EU (zie noot 2).
- Het verdient aanbeveling om het begrip integriteitsaantasting uit te werken en toe te voegen aan de as 'schade' van de aangepaste Bateson kubus. Dit zal de schade-batenanalyse verrijken. Daarom raadt het NCad de minister van LNV aan om te bevorderen dat integriteitsaantasting onderdeel gaat uitmaken van het 'Werkdocument over projectevaluatie en beoordeling achteraf' (zie noot 18) en op termijn worden meegenomen bij de evaluatie van Richtlijn 2010/63/EU (zie noot 2).

²⁰ Lilley, E, Hawkins, P, and Jennings, M (2014) A 'road map' toward ending severe suffering of animals used in research and testing. *Altern Lab Anim* 2014 Vol. 42 Issue 4 Pages 267-72



- Er kunnen vraagtekens gezet worden bij het onderscheid dat tussen diersoorten wordt gemaakt in Richtlijn 2010/63/EU. Het NCad is van mening dat niet de relatie tot of gelijkenis met de mens, maar de mate van ongerief die het dier ervaart (de gevoeligheid), gebaseerd op objectiveerbare criteria, relevant is voor het maken van een keuze voor een diersoort of foklijn. Een bepaalde handeling kan meer impact hebben op de ene diersoort of binnen de diersoort op de ene foklijn dan op de andere. Het NCad adviseert de minister van LNV om deze overwegingen mee te geven aan de Europese Commissie bij de evaluatie van Richtlijn 2010/63/EU.
- Vanwege regelgeving rondom veiligheid worden jaar in jaar uit op grote schaal dieren gebruikt voor het testen van de veiligheid en werkzaamheid van chemicaliën, geneesmiddelen en vaccins. Het NCad raadt de minister van LNV aan zich internationaal in te spannen om aanpassing van de regelgeving rondom veiligheid om vermindering van het gebruik van dieren voor het testen van veiligheid en werkzaamheid van stoffen mogelijk te maken.

Bijlage 1

Auteurs



NCad-werkgroepleden

Dr. Monique Janssens (tot juli 2022)

Prof. dr. ir. Jan-Bas Prins (vanaf juli 2022)

Dr. ir. Elmar Theune (staf NCad)

Externe deskundige

Dr. Corné Rademaker – opsteller van het rapport Prioriteren van dierproeven voor vervanging.



Bijlage 2

Lijst van gebruikte afkortingen

3Vs	Vervanging, Vermindering en Verfijning
CCD	Centrale Commissie Dierproeven
DEC	Dierexperimentencommissie
EU	Europese Unie
GM	Genetisch gemodificeerd
IvD	Instantie voor Dierenwelzijn
LNV	Landbouw, Natuur en Voedselkwaliteit
NAM	New Approach Methodology (ook wel: New Approach Method)
NCad	Nationaal Comité advies dierproevenbeleid
NWO	Nederlandse Organisatie voor Wetenschappelijk Onderzoek
OC&W	Onderwijs, Cultuur en Welzijn
RSPCA	Royal Society for the Prevention of Cruelty to Animals
Wod	Wet op de dierproeven
ZonMw	Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie.



Bijlage 3

Rapport 'Prioriteren van dierproeven voor vervanging' van Corné Rademaker

Samenvatting

Dit rapport is geschreven in opdracht van de Nationale Commissie advies dierproevenbeleid (NCad). Het doel is om te voorzien in criteria die gebruikt kunnen worden om prioriteiten te stellen in het volledig of relatief vervangen van dierproeven.

Op basis van interviews en literatuuronderzoek zijn vier criteria uitgewerkt:

1. de mate van transleerbaarheid van het onderzoek;
2. de mate van relevantie van het doel van een dierproef;
3. de mate van ongerief en
4. de mate van aantasting van integriteit dat de proefdieren berokkend wordt.

Transleerbaarheid, ten eerste, gaat over het bepalen van de reproduceerbaarheid en veralgemeniseerbaarheid van diermodellen die gebruikt worden in dierproeven. Dierproeven waarbij diermodellen gebruikt worden die een lage mate van transleerbaarheid hebben, komen daarbij als eerste in aanmerking om vervanging voor te ontwikkelen. Voor de inschatting van de mate van transleerbaarheid kan hierbij voortgebouwd worden op tools die al zijn ontwikkeld. Bij de relevante van het doel van onderzoek, ten tweede, is een indeling te maken naar basale, serieuze, en perifere doelen, waarbij proefdieronderzoek dat perifere doelen dient als eerste in aanmerking komt om vervanging voor te ontwikkelen. De mate van ongerief, ten derde, gaat over pijn, lijden, angst of blijvende schade die een individueel dier binnen een dierproef naar verwachting zal ondervinden. Proefdieronderzoek dat ernstig en langdurig ongerief oplevert komt als eerste in aanmerking voor vervanging, gevolgd door proefdieronderzoek met ernstig ongerief. De mate van aantasting van integriteit, als laatste, verwijst naar iedere lichamelijke interventie (smalle interpretatie van integriteit) of iedere interventie die leidt tot een verlies aan soortspecifieke vermogens (brede interpretatie van integriteit). Dierproeven waarbij een ernstige mate van aantasting van integriteit betrokken is, komen daarbij als eerste in aanmerking om vervangen te worden. Analoog aan wat gedaan is voor ongeriefinschatting, zou een classificatie voor mate van integriteitsaantasting echter verder uitgewerkt moeten worden door bijvoorbeeld een werkgroep. Aanbevelingen voor stakeholders met betrekking tot dierproeven worden aan het einde van het rapport gepresenteerd. Geconcludeerd wordt dat de genoemde vier criteria bruikbaar zijn om dierproeven en diermodellen in beeld te krijgen die voorrang moeten krijgen in de ontwikkeling van alternatieven.



Inleiding

Het staande beleid in Europa en Nederland met betrekking tot dierproeven berust op de implementatie van de 3V's – vervanging, vermindering en verfijning – en een schadenbatenanalyse. Het 3V-beleid berust op het werk van Russell en Burch (1959). Binnen de 3V's neemt vervanging een bijzondere plaats in. Volgens Sandøe et al. (2015) formuleerden Russell en Burch vervanging niet voor niets als eerste van de 3V's: vermindering en verfijning zijn alleen relevant als mogelijkheden voor vervanging uitgeput zijn. Ook is (volledige) vervanging de enige van de 3V's die volledig in overeenstemming is met de idealen en doelstellingen van dierrechtenorganisaties en actiegroepen. In de samenleving kan vervanging dan ook op brede steun rekenen (vgl. Smith en Boyd 1991, 134–36).

De Europese richtlijn 2010/63/EU die het gebruik van dieren in dierproeven reguleert, ziet volledige vervanging zelfs als het uiteindelijke doel van de richtlijn: “Deze richtlijn vormt evenwel een belangrijke stap in de richting van het bereiken van het einddoel om de tests op levende dieren voor wetenschappelijke en onderwijskundige doeleinden volledig door andere procedures te vervangen zodra dit wetenschappelijk mogelijk zal blijken” (Europees Parlement en de Raad van de Europese Unie 2010, 34). Recent heeft het Europees Parlement ook een resolutie aangenomen waarin aangezet wordt tot een versnelling van de transitie naar proefdiervrije methoden, om op die manier ook daadwerkelijk het uiteindelijke doel van richtlijn 2010/63/EU te verwirkelijken (European Parliament 2021).

In Nederland reguleert de Wet op de dierproeven (Wod) het gebruik van dieren “voor experimentele of andere doeleinden, waarvan het resultaat bekend of onbekend is, of onderwijskundige doeleinden, die bij het dier evenveel of meer pijn, lijden, angst of blijvende schade kan veroorzaken als, dan wel dan het inbrengen van een naald volgens goed diergeneeskundig vakmanschap” (Wet op de dierproeven 2021, §1, Artikel 1a). De Wod schrijft voor dat projectenvoorstellen ethisch getoetst moeten worden door de Centrale Commissie Dierproeven (CCD), daarbij geadviseerd door dierexperimentencommissies (DEC's). De DEC's en de CCD doen daarmee een toetsing aan de hand van een wettelijk voorgeschreven kader, inclusief mogelijkheden ter vervanging van dierproeven. Bij andere partijen, zoals wetenschappers, industrie, ministerie, en het Nationaal Comité advies dierproevenbeleid (NCad) ligt de opgave om het ontwikkelen en erkend krijgen van proefdiervrije alternatieven te blijven aanjagen, versnellen en bevorderen. Vanuit deze rol ziet het NCad een behoefte om prioriteiten te stellen bij het vervangen van dierproeven. Hoewel uiteindelijk het streven is om dierproeven volledig te vervangen, is er behoefte aan

een strategie om die dierproeven te identificeren die ethisch gezien het meest problematisch zijn, om zodoende de inspanningen te richten op het volledig of relatief vervangen van die dierproeven.²¹ In het maatschappelijke en wetenschappelijke debat staan bepaalde dierproeven meer ter discussie dan andere, om verscheidene redenen. Zo deed de Universiteit Maastricht in 2015 afstand van het onderzoek naar hartfalen in labradors na protesten tegen dit type onderzoek (Animals Today 2015; NOS 2015), en ligt de in het depressieonderzoek gebruikte Gedwongen Zwemtest in toenemende mate onder vuur vanwege het betrokken ongerief en validiteitsproblemen (PETA 2022; Molendijk en De Kloet 2019). In dit onderzoek worden vanuit dit maatschappelijke en wetenschappelijke debat criteria gedestilleerd en uitgewerkt die behulpzaam kunnen zijn bij het prioriteren in het vervangen van dierproeven. De leidende gedachte hierbij is dat dierproeven die ethisch gezien het minst acceptabel zijn, het eerst vervangen zouden moeten worden. De gezochte criteria moeten hier dan ook uitdrukking aan geven.

Het uiteindelijke doel van de NCad is het leveren van een zienswijze waarin aan de hand van criteria geschetst wordt wat prioriteiten kunnen zijn in het streven naar vervanging van bepaalde dierproeven. Het doel van voorliggend rapport is een eerste bijdrage te leveren aan dit uiteindelijke doel door middel van interviews en literatuuronderzoek.

²¹ Russell en Burch (1959, hfdst. 5) onderscheiden naar volledige of absolute en relatieve vervanging. Absolute vervanging is het volledig vervangen van proefdieren door non-sentient materiaal, en vormt het absolute ideaal. Relatieve vervanging vatten Russell en Burch op als het gebruik van dieren “exposed, probably or certainly, to no distress at all” (Russell en Burch 1959, hfdst. 5, Modes of Absolute and Relative Replacement). Beide opties van vervanging zijn in dit rapport geïmpliceerd, waarbij aangetekend moet worden dat naast ongerief (distress) ook de aantasting van integriteit meespeelt. Vanuit het oogpunt van relatieve vervanging hebben niet alleen dierproeven met minder ongerief de voorkeur, maar ook dierproeven met een mindere mate van integriteitsaantasting.



Methodologie

Voor de bepaling van de scope van voorliggend onderzoek is eind 2021 contact gezocht met belanghebbenden uit het Nederlandse proefdierkundige veld. Acht personen zijn per e-mail benaderd met de vraag of zij wilden meewerken aan een “onderzoek naar dierproeven waar goede redenen voor zijn om vanaf te willen” (zie Appendix A). De vorm was een semigestructureerd online interview van maximaal een uur via Zoom of Teams, waarbij de leidende vragen waren: Worden er momenteel diermodellen gebruikt die eigenlijk ongewenst zijn? Waarom worden deze diermodellen nog gebruikt? Welke mogelijkheden zijn er om deze diermodellen te vervangen? Van de acht benaderde personen reageerden er zeven, allen positief. Twee van deze zeven personen brachten nog een extra persoon vanuit hun organisatie mee, zodat er twee duo-interviews gehouden zijn. Daarnaast zijn dierenbeschermingsorganisaties die zouden participeren in een al gepland rondetafelgesprek met het NCad uitgenodigd voor een groepsinterview aansluitend aan dit gesprek. Allen reageerden positief en stuurden één vertegenwoordiger. Dit groepsinterview duurde bijna een uur en veertig minuten.

De interviews met één persoon en het groepsinterview zijn met toestemming opgenomen en vervolgens woordelijk getranscribeerd. Respondenten van een van de duo-interviews hadden bezwaar tegen opname; daarom zijn bij dat interview aantekeningen gemaakt. Respondenten van het andere duo-interview hadden geen bezwaar tegen opname, maar wel tegen het woordelijke transcript. Het bezwaar was dat men alleen hoofdlijnen wilde duidelijk maken, geen details. Het woordelijke transcript is daarom omgevormd tot een samenvattend transcript.

Twee hoofdconclusies op basis van de interviews waren dat het genereren van een lijst met concrete ‘ongewenste’ diermodellen of dierproeven niet eenvoudig is, en dat de vraagstelling van het project specifiek moest. Zowel in de interviews als in het groepsinterview met dierenbeschermingsorganisaties werden weinig concrete (typen) dierproeven genoemd die op de gewenste prioriteitenlijst zouden moeten komen te staan. Op basis van deze informatie is besloten dat de focus moest komen te liggen op het beargumenteerd ontwikkelen van criteria ter prioritering van het vervangen van dierproeven, in elk geval in deze eerste fase.

In eerste instantie is getracht om een systematisch literatuuronderzoek van de wetenschappelijke literatuur te doen, gebaseerd op de zoektermen “priorit* AND replace* AND

“animal experiment*” OR “animal research” binnen de zoekmachines PubMed en Google Scholar. Dit leverde echter weinig resultaten op. Er is daarom gekozen voor een andere aanpak. Op basis van informatie uit interviews met belanghebbenden ter bepaling van de scope van dit literatuuronderzoek, en kennis en kunde van de auteur van dit rapport en vanuit het NCad, inclusief ondersteunend bureau, is een lijst van *mogelijke* criteria opgesteld. Deze criteria betroffen:

1. Mate van **transleerbaarheid** van het onderzoek
2. Mate van **relevantie** van het **doel** van een dierproef
3. Mate van **ongerief** die de proefdieren berokkend wordt
4. Mate van aantasting van **integriteit** van proefdieren
5. Mate van **proportionaliteit**
6. **Ontwikkelingsstatus** van **alternatieven**
7. **Aantallen dieren** of **procedures**
8. Gebruikte **diersoort**

Bij nadere beschouwing zijn criteria 1 tot en met 4 *wel*, maar criteria 5 tot en met 8 *niet* geschikt bevonden om nader te verkennen. In de paragraaf ‘Ongeschikte criteria ter prioritering vervanging dierproeven’ wordt eerst toegelicht waarom criteria 5 tot en met 8 ongeschikt bevonden zijn om nader uit te werken, waarna in de paragraaf ‘Geschikte criteria ter prioritering vervanging dierproeven’ criteria 1 tot en met 4 nader uitgewerkt worden. Ieder van deze laatste vier criteria is nader verkend middels een narratief review met behulp van de zoekmachines PubMed, Google Scholar en Google Search, aangevuld met *snowball sampling* en aangedragen stukken door experts uit het netwerk van de auteur. Binnen de beschikbare tijd is daarbij de meeste nadruk gelegd op de criteria van transleerbaarheid en integriteitsaantasting, gebaseerd respectievelijk op het gegeven dat er steeds meer stemmen opgaan om hier meer aandacht aan te geven bij de (ethische) beoordeling van dierproeven (Eggel en Würbel 2021; Garner e.a. 2017; Meijboom, Kostrzewa, en Leenaars 2020; Olsson en Sandøe 2021; Pound en Nicol 2018; Pound en Ritskes-Hoitinga 2018; Wehling 2009; Würbel 2017; Landi, Everitt, en Berridge 2021) en de behoefte aan nadere duiding van wat integriteitsaantastingen behelzen.



Ongeschikte criteria ter prioritering vervanging dierproeven

In deze paragraaf wordt toegelicht waarom mate van proportionaliteit, ontwikkelingsstatus van alternatieven, aantallen dieren of procedures, en de gebruikte diersoort ongeschikte criteria vormen ter prioritering in het vervangen van dierproeven.

Mate van proportionaliteit

Mate van proportionaliteit is beoordeeld als ongeschikt als *afzonderlijk* en *zelfstandig* criterium bij het prioriteren in het vervangen van dierproeven. Proportionaliteit zegt immers iets over de geschiktheid, noodzakelijkheid en adequaatheid van een dierproef in de context van schadeberokkening (Würbel 2017), en verwijst daarbij naar andere *criteria*. De geschiktheid en adequaatheid van dierproeven worden deels ondervangen respectievelijk onder de criteria van transleerbaarheid (zie 'Transleerbaarheid van het onderzoek: externe validiteit') en de relevantie van het doel van een dierproef (zie 'Relevantie van het doel van een dierproef'). Noodzakelijkheid houdt daarnaast verband met de ontwikkelingsstatus van alternatieven.

Ontwikkelingsstatus van alternatieven voor proefdieronderzoek

De ontwikkelingsstatus van alternatieven voor een bepaald type dierproef is ook overwogen als criterium in het vervangen van dierproeven. Het idee bij dit mogelijke criterium is dat prioriteit in het vervangen gegeven moet worden aan die typen dierproeven waarvoor de ontwikkeling van een alternatief al enigszins nabij is – even afgezien van de 'score' op andere criteria. Het gaat er dan om, om het laaghangend fruit te oogsten. Zo is in het algemeen de gedachte dat de ontwikkeling van alternatieven wetenschappelijk en technisch gezien bijvoorbeeld eenvoudiger is voor toxiciteitsonderzoek – waar veel gestandaardiseerd en herhaald wordt – dan voor hypothese-gedreven wetenschappelijk onderzoek. In tegenstelling tot de andere criteria heeft de status van de ontwikkelingsfase van alternatieven vooral een strategisch karakter. Dit criterium zegt op zichzelf niet waar meer inspanningen geleverd moeten worden in het vervangen van dierproeven *omdat daar de grootste ethische problemen liggen*, maar het wijst erop dat resultaten sneller geboekt kunnen worden bij alternatieven die al enigszins in ontwikkeling zijn. Er is daarom voor gekozen om dit criterium niet verder uit te werken. Dit neemt niet weg dat ook 'dichtbij' oplossingen aandacht verdienen.

Aantallen dieren of procedures

Ook de aantallen betrokken dieren of procedures vormen *op zichzelf* geen goed criterium voor het prioriteren in het vervangen van dierproeven. Aantallen dieren zijn wel relevant,

maar altijd in relatie tot andere karakteristieken, zoals ongerief en onderzoeksdoel. Zo is het bijvoorbeeld relevant dat in 2017 de grootste aantallen procedures met ernstig ongerief (meer dan 264.000) gerelateerd waren aan *batch potency testing* (European Commission 2020). Dit komt echter aan de orde bij het criterium van ongerief (zie 'Ongerief van de proefdieren').

Gebruikte diersoort

Binnen de Wod hebben de primatensoorten chimpansee, bonobo, orang-oetan en gorilla een bijzondere status. Dierproeven met deze diersoorten zijn zonder meer verboden (*Wet op de dierproeven* 2021, §10e, Artikel 1). Onderzoek met andere 'niet-menselijke primaten' mag alleen uitgevoerd worden onder voorwaarden die stringenter zijn dan voor andere diersoorten (*Wet op de dierproeven* 2021, §10e, Artikel 2 en 3). In het Verenigd Koninkrijk gelden ook strengere eisen voor het gebruik van katten, honden en paardachtigen (*Animals (Scientific Procedures) Act 1986* 2013; Schedule 2B Additional conditions for the grant of certain project licences). In de praktijk is het bovendien zo dat onderzoekinstellingen de keuze maken om onderzoek met bepaalde diersoorten niet (meer) uit te voeren, zoals in de Inleiding al benoemd in verband met het hartfalenonderzoek aan de Universiteit Maastricht. Uit het bekende Britse Ipsos Mori-onderzoek blijkt verder dat mensen dierproeven met bijvoorbeeld honden, katten en primaten minder acceptabel vinden dan onderzoek met ratten en muizen (Cameron, Clemence, en Xypolia 2018).²²

Inhoudelijk gezien kleven er echter problemen aan het gebruik van diersoort als zelfstandig prioriteringscriterium. Diersoort kan relevant zijn omdat het gaat om een soortelijk verschil in het vermogen tot het ervaren van ongerief, of om verschillen in homologe validiteit en daarmee transleerbaarheid naar de mens, of om de morele beschermwaardigheid van dieren op basis van lidmaatschap van een soort. Het nader formuleren van het criterium van gebruikte diersoort tot *beschermwaardigheid* van de diersoort, lost de problemen echter nog niet op. Hoewel sommigen de maatschappelijke hiërarchie in diersoorten gebruiken in de ethische oordeelsvorming (Hursthouse 2011; Kagan 2016; 2018), is het toekennen van een hogere beschermwaardigheid op basis van lidmaatschap van een soort ook bekritiseerd als *speciesistisch* en daarmee onbruikbaar (Olsson en Sandøe 2021; Singer 1993). Er lijkt hier daarom niet genoeg consensus te bestaan om beschermwaardigheid van een diersoort te kunnen laten dienen als criterium in het prioriteren bij het vervangen van dierproeven.

²² Wel vonden Deense studies dat diersoort als criterium minder gewicht heeft voor mensen dan bijvoorbeeld het betrokken ongerief (T. B. Lund, Lassen, en Sandøe 2012; T. B. Lund e.a. 2014).



Geschikte criteria ter prioritering vervanging dierproeven

Binnen het literatuuronderzoek zijn de volgende mogelijke criteria ter prioritering in het vervangen van dierproeven nader uitgewerkt: mate van transleerbaarheid van het onderzoek, mate van relevantie van het doel van een dierproef, mate van ongerief die de proefdieren berokkend wordt, en de mate van aantasting van integriteit van proefdieren. In de volgende vier paragrafen worden deze criteria nader uitgewerkt.

Transleerbaarheid van het onderzoek

We believe that more rigorous and extensive evaluations of animal models are necessary. (...) there are fields where it is widely recognized that existing animal models are insufficient. (...) We propose that, for improved ethical animal experimentation, animal research in these situations should be avoided until improved models are developed. (Vieira de Castro en Olsson 2015)

De transleerbaarheid of vertaalbaarheid van dierproeven is al geruime tijd een punt van aandacht in het biomedisch onderzoek (Ferreira e.a. 2020; Garner 2014; Garner e.a. 2017; Greek en Kramer 2019; A. Knight 2019; Leenaars e.a. 2019; Pound en Ritskes-Hoitinga 2018; Ram 2019; Van der Worp e.a. 2010; Zeiss 2015). Het punt is hier dat data verkregen uit dierexperimenten maar beperkt vertaalbaar zijn naar de mens.

Het percentage medicijnen die veelbelovend zijn in fase I klinische trials, maar het niet halen tot de markt ligt al enige tijd tussen de 86% en 93% (Alliance for Human Relevant Science 2020; Garner 2014; Hay e.a. 2014; Kola en Landis 2004; Thomas e.a. 2016; Wong, Siah, en Lo 2019). Redenen waarom medicijnen het niet halen tot de markt betreffen voornamelijk gebrek aan effectiviteit en veiligheid, maar ook bedrijfsredenen als gebrek aan financiering, portfolioprioritering en Intellectual Property-kwesties spelen mee (Cook e.a. 2014; Hay e.a. 2014; Thomas e.a. 2016). Veel minder eenduidigheid is er echter over de stap van onderzoek met proefdieren naar klinische trials met mensen. Een recent review van reviews met betrekking tot de vertaalbaarheid van bevindingen in proefdieren naar die in fase I/II-trials met mensen, wees uit dat succespercentages variëren van 0% tot 100% (Leenaars e.a. 2019). Dit is niet erg bemoedigend, omdat dit betekent dat de studies mogelijk niet voorspellend zijn voor de mens. Dat wil zeggen dat aan het criterium van voorspellende validiteit (*predictive validity*) mogelijk niet voldaan wordt.

Diermodellen en modeldieren

Naast voorspellende validiteit kunnen andere vormen van validiteit van diermodellen geïdentificeerd worden. Alvorens verschillende validiteitscriteria van diermodellen te bespreken, is het echter goed om helder te maken wat we onder diermodellen verstaan. Als eerste maken we onderscheid tussen diermodellen en modeldieren (vgl. Insel 2007). Soms komt men in de praktijk uitspraken tegen als: 'De diermodellen zullen gedood worden in het kader van het experiment.' Dit is een voorbeeld van een reïficatie: de drogreden dat een abstract construct gehanteerd wordt als ware het iets concreets.²³ Een dier dat gebruikt wordt als model voor een menselijke ziekte of aandoening is echter niet zelf een gesimplificeerde versie van die ziekte of aandoening (Hauskeller 2007).

Aangezien een diermodel als model of construct abstract is, is het niet iets wat gedood kan worden, zoals een concreet (proef)dier wel gedood kan worden. Naast diermodellen moeten we dus concrete proefdieren of modeldieren onderscheiden.

Verder moet opgemerkt worden dat wanneer in de praktijk gesproken wordt over het gebruik van een diermodel, dit enerzijds eigenlijk altijd gaat over het gebruik van proefdieren als model voor fysiologische, biologische of gedragsfenomenen in de mens. In die zin is er dus geen sprake van het gebruik van diermodellen binnen bijvoorbeeld landbouwkundig onderzoek, omdat daar vrijwel altijd het proefdier tot dezelfde soort behoort als het doeldier. Bijvoorbeeld, nieuwe voeders ten bate van de varkenshouderij (doeldieren) worden getest op varkens (proefdieren) en niet op een andere diersoort. Anderzijds kan gesteld worden dat ook de proefvarkens in genoemd voorbeeld model staan, en wel voor (gedomesticeerde) varkens van een bepaald ras of een bepaalde foklijn. Het gaat hier niet om de extrapolatie van de ene soort naar de andere soort zoals in biomedisch onderzoek, maar van individuele varkens in de proefsetting naar het ras of de foklijn waar ze toe behoren en die gebruikt wordt in de doelpraktijk (hier: varkenshouderij).

In het navolgende sluiten wij ons aan bij het in de praktijk gangbare begrip van een diermodel als extrapolatie van dier naar mens. Bijgevolg vatten we diermodellen op als simulaties van fysiologische, biologische of gedragsfenomenen in de mens binnen een

²³ Zie bijvoorbeeld ook de veelgeciteerde definitie van Held (1980), geciteerd in Hau en Schapiro (2011, 2): "a living organism in which a normative biology or behavior can be studied, or in which a spontaneous or induced pathological process can be investigated, and in which the phenomenon in one or more respects resembles the same phenomenon in humans."



diersoort, op basis van een van nature voorkomende, genetisch gemodificeerde of experimenteel geïnduceerde karakteristiek binnen die diersoort, met het doel om kennis te genereren die geëxtrapoleerd kan worden naar de menselijke doelgroep.²⁴ Voorbeelden op basis van een van nature voorkomende, genetisch gemodificeerde of experimenteel geïnduceerde karakteristiek zijn respectievelijk het hondenleukocytenadhesiedeficiëntie (CLAD)-model (Bauer e.a. 2004), het ClockΔ19-muismodel (Kristensen, Nierenberg, en Østergaard 2018), en het experimentele auto-immuun-encefalomyelitis-model in muizen, ratten en makaken (Procaccini e.a. 2015). Modeldieren daarentegen zijn concrete, levende dieren binnen een experiment die *model staan voor* bepaalde fysiologische, biologische of gedragsfenomenen in de mens. De manier *waarop* zij model staan is gespecificeerd in het diermodel. Dus, Ierse Setters met CLAD binnen een experimentele context zijn honden die model (kunnen) staan voor leukocytenadhesiedeficiëntie in kinderen.²⁵ De manier waarop zij model staan heeft betrekking op de (veronderstelde) overeenkomsten in symptomen en ziekteverloop van, en mechanismen achter, leukocytenadhesiedeficiëntie in honden en mensen.

Als het nu gaat over de transleerbaarheid van proefdieronderzoek ligt de focus dus direct op die onderzoeksgebieden waarbij dieren ingezet worden ten bate van kennisverwerving met betrekking tot fysiologische, biologische of gedragsfenomenen in de mens, en niet bijvoorbeeld op veterinaire onderzoek, landbouwkundig onderzoek of fundamenteel onderzoek in andere diersoorten.²⁶

Klassieke validiteitscriteria van diermodellen

“Transleerbaarheid is de validiteit van het model in het dupliceren van een specifiek aspect van targets, mechanismen, en/of ziekte in de mens” (Landi, Everitt, en Berridge 2021). Transleerbaarheid heeft dus alles te maken met de validiteit van diermodellen. Een diermodel (van ziekte) heeft validiteit in zoverre het een goed model vormt van de menselijke ziekte of aandoening (Belzung en Lemoine 2011). De vraag is dan natuurlijk

²⁴ Dit ligt dichtbij Sjöberg's (2017) definitie, gecorrigeerd voor de reïficatie in Sjöberg's definitie en deels de beschrijving van Hau en Schapiro (2011) volgend voor verschillende typen diermodellen.

²⁵ Ze kunnen natuurlijk ook model staan voor honden als zodanig, zoals in veterinaire onderzoek.

²⁶ Met name dit laatste kan wezenlijk inzichten opleveren ook voor de menselijke situatie, hoewel van tevoren geen doel van het onderzoek. Een voorbeeld is de ontwikkeling van optogenetica waarin de response van specifieke cellen in levend weefsel, zoals neuronen, op licht bestudeerd wordt in soorten als fruitvliegen, klauwkikkers en muizen (voor een review, zie Deisseroth 2015). Transleerbaarheid als criterium is hier (nog) niet aan de orde, omdat de probleemstelling vertrekt vanuit een bepaald biologisch proces en niet vanuit een bepaalde maatschappelijke context zoals 'de kliniek'.

wanneer een diermodel een goed model is. Hierbij kunnen meerdere aspecten onderscheiden worden, waarvan de al genoemde voorspellende validiteit er een is: de mogelijkheid om voorspellingen te doen met betrekking tot een menselijk fenomeen op basis van het presteren van een diermodel (Geyer en Markou 1995). Het kan dan gaan om het voorspellen van het effect van therapieën en behandelingen, van specifieke markers van een ziekte of van het ziekteverloop (Belzung and Lemoine 2011; Holmes 2003; van der Staay, Arndt, and Nordquist 2009). Voorspellende validiteit is een van de drie klassieke vormen van validiteit ter evaluatie van diermodellen (Willner 1984) en ter evaluatie van de vertaalbaarheid van proefdieronderzoek (Landi, Everitt, en Berridge 2021). Naast voorspellende validiteit is er fenotypische validiteit (*face validity*) en constructvaliditeit (*construct validity*).

Fenotypische validiteit heeft betrekking op de mate van overeenkomst tussen het diermodel en de onderzochte ziekte of aandoening bij een zo breed mogelijke range van symptomen (Belzung en Lemoine 2011; Geyer en Markou 1995; Van der Staay, Arndt, en Nordquist 2009; Willner en Mitchell 2002). Fenotypische validiteit poogt dus vooral om een of meerdere klinische diagnostische kerncriteria te reproduceren in een diermodel. Hoewel sommigen hierbij symptomen beperken tot het niveau van cognitie, emotie en gedrag (Belzung en Lemoine 2011; Van der Staay, Arndt, en Nordquist 2009), is dit niet noodzakelijk. Een overeenkomst in symptomen kan ook betrekking hebben op neurobiologische en fysiologische abnormaliteiten (Geyer en Markou 1995; Holmes 2003; Willner en Mitchell 2002). Voorbeelden zijn gewichtsverlies en toegenomen activiteit in de hypothalamus-hypofyse-bijnier-as bij depressie (Holmes 2003; Willner en Mitchell 2002).

In de literatuur kunnen grofweg drie opvattingen van constructvaliditeit gevonden worden, waarbij de eerste twee waarschijnlijk voortkomen uit de klassieke formulering van Willner (1984). In de eerste opvatting is constructvaliditeit de eis dat het experiment of model meet wat het moet meten en ondubbelzinnig geïnterpreteerd kan worden (Geyer en Markou 1995; Sjöberg 2017). De Gedwongen Zwemtest bijvoorbeeld werd ontwikkeld als een test voor het wanhoopssymptoom van depressie, maar een nieuwere interpretatie is dat het veeleer drijfgedrag als copingstrategie test (Commons e.a. 2017; Molendijk en De Kloet 2015; 2019; De Kloet en Molendijk 2016; Molendijk en De Kloet 2021).

De tweede opvatting van constructvaliditeit behelst de mate waarin een diermodel overeenkomt met de theorie achter de gemodelleerde ziekte of aandoening in de mens (Holmes 2003; Willner en Mitchell 2002). In het geval van depressie, bijvoorbeeld, kunnen



in diermodellen getheoretiseerde neurobiologische, etiologische of psychologische mechanismen getoetst worden.²⁷ Het probleem hierbij is dat, zeker in neuropsychiatrisch onderzoek, er niet één theorie is maar er meerdere zijn, en dat bijgevolg het ook niet eenduidig is aan welke theorie het diermodel getoetst moet worden.

De derde opvatting van constructvaliditeit richt zich direct op de mate van overeenkomst in mechanismen die ten grondslag liggen aan de ziekte of aandoening in de mens en de geobserveerde abnormaliteiten in het diermodel (Sams-Dodd 2006; Van der Staay, Arndt, en Nordquist 2009). Met Belzung en Lemoine (2011) zouden we dit ook *mechanistische validiteit* (*mechanistic validity*) kunnen noemen. Deze staat in contrast met fenotypische validiteit, die zich richt op de *effecten* van de (veronderstelde) mechanismen.

Nieuwere criteria van validiteit van diermodellen

Na de definiëring van voorspellende validiteit, fenotypische validiteit en constructvaliditeit zijn er verdere ontwikkelingen geweest. Hierbij zijn de drie klassieke validiteitscriteria aangevuld met andere criteria, en deels ook geherinterpreteerd. Overigens zijn deze criteria niet altijd 'nieuw'. Convergente validiteit (convergent validity) en onderscheidende validiteit (discriminant validity), bijvoorbeeld, werden al in 1959 geformuleerd in een psychologische context (D. T. Campbell en Fiske 1959). De incorporatie ervan in de expliciete evaluatie van diermodellen vond echter pas later plaats (Geyer en Markou 1995).

Validiteitscriteria volgens Garner

Garner (2014; 2017) plaatst de drie klassieke criteria van validiteit in één, wat hij noemt, dimensie. Daarnaast onderscheidt hij nog twee dimensies. De tweede dimensie gaat over interne (*internal validity*) en externe validiteit (*external validity*) en de derde dimensie gaat over convergente en onderscheidende validiteit. Garner geeft geen definities of omschrijvingen van de verschillende typen van validiteit, maar formuleert sleutelvragen.

Bij interne validiteit gaat het om de vraag of de methodologie en resultaten van de test of het model consistent zijn met zowel de theorie als de bestaande data van het model. Het gaat hierbij zowel om betrouwbaarheid (*reliability*) als reproduceerbaarheid (*replicability or reproducibility*) (Van der Staay, Arndt, en Nordquist 2009). Een betrouwbaar diermodel geeft bij herhaling consistente uitkomsten, en is reproduceerbaar in verschillende contexten, i.e., laboratoria. Bij externe validiteit gaat het om de vraag of de resultaten van de test of het model breed toepasbaar zijn (*generalizability*). Hierbij gaat het er allereerst om

of de verkregen inzichten ook van toepassing zijn op de onderzochte populatie of foklijn, maar daarna ook over toepassing op soortniveau en uiteindelijk ook op de menselijke situatie.

Bij convergente validiteit is de centrale vraag of de test of het model in brede zin overeenstemming vertoont met de eigenschappen van datgene wat gemeten wordt, of met de eigenschappen van de gemodelleerde menselijke situatie. Het gaat er hierbij om of een diermodel of test correleert met andere diermodellen of tests die eenzelfde fenomeen modelleren. Bij onderscheidende validiteit, als laatste, gaat het om de vraag of de test of het model alternatieve processen of andere diagnoses uitsluit. Dit is de keerzijde van convergente validiteit, aangezien bij onderscheidende validiteit een lage correlatie van verschillende testen of diermodellen wenselijk is, omdat ze geacht worden verschillende fenomenen te modelleren en meten.

Validiteitscriteria volgens Belzung en Lemoine

Ook Belzung en Lemoine (2011) hebben andere validiteitscriteria geformuleerd naast de drie klassieke, en daarbij de drie klassieke validiteitscriteria ook geherinterpreteerd. De criteria die zij onderscheiden zijn fenotypische validiteit, mechanistische validiteit, voorspellende validiteit, homologe validiteit (*homological validity*), en pathogene validiteit (*pathogenic validity*).

Fenotypische validiteit vatten Belzung en Lemoine op in de klassieke lijn. Wel specificeren ze deze naar ethologische validiteit (*ethological validity*) en biomarker-validiteit (*biomarker validity*). De eerste heeft betrekking op de mate van overeenkomst in gedragingen tussen modeldieren en mensen, terwijl de laatste betrekking heeft op de mate van overeenkomst in biotische markers. Een voorbeeld van dat laatste is de mate van overeenkomst in afwijkende concentraties van corticosteron in het bloed van ratten en van cortisol in mensen. Hiermee is gelijk aangegeven dat het niet gaat om de absolute concentratie van deze glucocorticoiden in het bloed – die verschilt tussen rat en mens – noch om een identieke chemische samenstelling, maar om eenzelfde functie van de marker: corticosteron in de rat vervult dezelfde functie als cortisol bij de mens (Belzung en Lemoine 2011).

Mechanistische validiteit als interpretatie van de klassieke constructvaliditeit is reeds eerder besproken.

²⁷ Soms wordt de etiologie echter opgevoerd middels een separaat criterium, namelijk etiologische validiteit (*etiological validity*) (bijvoorbeeld Belzung en Lemoine 2011; Geyer en Markou 1995).



Voorspellende validiteit specificeren Belzung en Lemoine nader als de overeenkomst in relatie tussen enerzijds de triggerfactoren en het optreden van ziekte in modeldieren (inductievaliditeit; *induction validity*), en anderzijds de therapie of medicatie en de ziekte in de mens (vereffeningsvaliditeit; *remission validity*).

Homologe validiteit heeft betrekking op de gekozen diersoort en foklijn. Een voorbeeld dat Belzung en Lemoine geven is dat *Caenorhabditis elegans* een slechte keuze is om de reductie in hippocampusvolume bij depressie te modelleren – *C. elegans* heeft immers geen hippocampus – maar een betere keuze wanneer de focus is op basale serotonerge fenomenen onder stresscondities.

Pathogene validiteit gaat over de overeenkomst tussen diermodel en de situatie in de mens met betrekking tot processen die leiden tot ziekte. Hierbij onderscheiden Belzung en Lemoine naar ontopathogene validiteit (*ontopathogenic validity*) en trigger-validiteit (*trigger validity*). Ontopathogene validiteit heeft betrekking op de overeenkomst in vroege omgevingsfactoren die een organisme kwetsbaar maken in latere fasen van het leven. Trigger-validiteit gaat vervolgens over de overeenkomst in factoren die een organisme pathologisch maken.

Deze vijf criteria met bijhorende subcriteria presenteren Belzung en Lemoine als een generiek raamwerk ter validatie van diermodellen van psychiatrische aandoeningen.

Tabel 1 overzicht van validiteitscriteria

Klassieke validiteitscriteria	Validiteitscriteria volgens Garner		Validiteitscriteria volgens Belzung en Lemoine	
	Dimensie	Subtype	Hoofdcriterium	Subcriterium
Fenotypische validiteit	Fenotypisch construct voorspellend	vs. Fenotypische validiteit	Fenotypische validiteit	Ethologische validiteit Biomarker validiteit
Constructvaliditeit		Constructvaliditeit	Mechanistische validiteit	-
Voorspellende validiteit		Voorspellende validiteit	Voorspellende validiteit	Inductievaliditeit Vereffingsvaliditeit
	Intern vs. Extern		Interne validiteit Externe validiteit	
	Convergent onderscheidend	vs. Convergente validiteit	Onderscheidende validiteit	
			Homologe validiteit	Soortgelijke validiteit Foklijn validiteit
			Pathogene validiteit	Ontopathogene validiteit Trigger validiteit



Toepassingsbereik van validiteitscriteria

De genoemde validiteitscriteria zijn ontwikkeld binnen het veld van neurologische en psychiatrische aandoeningen zoals depressie, angst en schizofrenie (Sams-Dodd 1999; Willner 1984; Willner, Muscat, en Papp 1992). Het beoordelen van diermodellen aan de hand van validiteitscriteria is ook grotendeels beperkt gebleven tot dit veld van onderzoek (bijvoorbeeld, Belzung en Lemoine 2011; Garner 2014; Meijboom, Kostrzewa, en Leenaars 2020; Morrice, Gregory-Evans, en Shaw 2018; Treit, Engin, en McEown 2010; Tricklebank en Garner 2012). Hoewel Sams-Dodd (2006) een generiek raamwerk voor de evaluatie van de klinische validiteit van ziektemodellen ontwikkeld heeft, blijven de voorbeelden hierbij ook beperkt tot de neuropsychiatrie. Denayer et al. (2014) hebben de stap gemaakt buiten de neuropsychiatrie en een modificatie van Sams-Dodd's raamwerk van validiteitscriteria toegepast binnen de oncologie.

Het is niet helder waarom de toepassing van verschillende validiteitscriteria beperkt zijn gebleven tot de neuropsychiatrie. Een mogelijkheid is dat de neuropsychiatrie te kampen heeft met de grootste vertaalbaarheidsproblemen. Er is tot op heden niet genoeg informatie om te kunnen beoordelen of er inderdaad significante verschillen zijn in vertaalbaarheid tussen onderzoeksvelden (Leenaars e.a. 2019). Wel is bekend dat niet medicijnen voor neurologische en psychiatrische aandoeningen, maar oncologische medicijnen het laagste succespercentage hebben in klinische trials (93-97% faalt, zie Hay e.a. 2014; Thomas e.a. 2016; Wong, Siah, en Lo 2019).²⁸ Dit suggereert dat niet de neuropsychiatrie maar de oncologie te kampen heeft met de grootste vertaalbaarheidsproblemen (vgl. Mak, Evaniew, en Ghert 2014). Dit kan verklaren waarom de validiteitscriteria nu ook toegepast worden binnen de oncologie. Anderzijds scoort echter ook cardiovasculair onderzoek laag op succes (93% faalt; Hay e.a. 2014; Thomas e.a. 2016)²⁹, maar is het toepassen van de validiteitscriteria daar tot op heden beperkt.

Hoe dit ook zij, er zijn geen principiële redenen om validiteitscriteria niet breder toe te passen dan alleen op de neuropsychiatrie en oncologie. Sterker nog, ook in bijvoorbeeld infectieonderzoek speelt een criterium als fenotypische validiteit impliciet een rol. Een onderdeel in de zoektocht naar een valide diermodel van SARS-CoV-2-infecties in mensen

²⁸ Therapieën zoals verschillende vormen van bestraling vallen niet onder de FDA-classificatie van 'drugs' en zijn hier dus ook niet meegenomen. Immunotherapie valt er wel onder, omdat het daarbij gaat om medicijnen als anti-PD-1 checkpoint inhibitors (durvalumab, pembrolizumab, nivolumab) en anti-CTLA-4 antilichamen (ipilimumab).

²⁹ Het is goed om op te merken dat op basis van andere publicaties het percentage medicijnen voor cardiovasculaire aandoeningen dat faalt vele malen lager ligt, namelijk 74-80% (Kola en Landis 2004; Wong, Siah, en Lo 2019). Mogelijk heeft dit te maken met het verschil in berekening van succespercentages. Zie voor een discussie over methodologie, Wong et al. (2019).

betrof de mate van overeenkomst in klinische verschijnselen als de productie van snot (*nasal discharge*), hoesten, verhoging van de lichaamstemperatuur, en verhoogde cytokine-profielen in het bloed (Bertzbach e.a. 2021; A. Singh e.a. 2020; Zhai e.a. 2021). Singh et al. (2020) roepen dan ook op tot een validatie van diermodellen voor SARS-CoV-2 met behulp van bijvoorbeeld de drie klassieke validiteitscriteria (vgl. Swearingen 2018). Recentere tools van Ferreira et al. (2019; 2020) en Wehling (Wehling 2009; Wendler en Wehling 2012; 2017) om de validiteit van diermodellen in te schatten hebben dan ook een generieke focus (zie 'Tools om de validiteit van diermodellen te bepalen').

Hierbij moet nog opgemerkt worden dat transleerbaarheid als criterium relevant is van zodra er sprake is van een diermodel (zie ook 'Diermodellen en modeldieren'). Ook bij fundamenteel onderzoek binnen het biomedische domein, bijvoorbeeld naar de achterliggende mechanismen van MS, is transleerbaarheid van belang. Het gaat hier echter om de transleerbaarheid van *kennis* met betrekking tot een bepaald biologisch of etiologisch fenomeen, verworven middels een diermodel, naar de menselijke situatie. Bij toegepast onderzoek betreft het echter de transleerbaarheid van de *effectiviteit* van een ontwikkeld technisch product zoals een implantaat of een medicijn. Fundamenteel en toegepast onderzoek binnen het biomedische domein hebben dus wel een verschillende focus – die overigens ook gecombineerd kunnen worden in een individueel onderzoeksproject – maar transleerbaarheid is relevant voor beide.³⁰

Transleerbaarheid als criterium

In de literatuur wordt interne validiteit vaak besproken in de context van de *experimentele proefopzet*: wordt er gerandomiseerd, wordt er geblindeerd, wordt er van tevoren een powerberekening uitgevoerd, worden de juiste controlegroepen geselecteerd, etc. (Henderson e.a. 2013)? Dit zijn belangrijke procedures in het licht van goede wetenschapsbeoefening. Het gaat hierbij om het borgen van de reproduceerbaarheid van dierproeven (betrouwbaarheid vormt daar een onderdeel van in die zin dat betrouwbaarheid iets zegt over de gebruikte meetinstrumenten). Reproduceerbaarheid wordt echter niet alleen geborgd door te randomiseren, blinderen, etc., oftewel door goede wetenschapsbeoefening, maar is ook afhankelijk van de methode waarmee een specifiek diermodel gegeneerd wordt. Bijvoorbeeld, de transleerbaarheid van een onderzoek naar multiple sclerose (MS) in resusapen op basis van het experimentele auto-immune encefalomyelitismodel (EAE-model) is niet alleen mede afhankelijk van een juiste powerberekening, randomisering en blinding, maar ook van de reproduceerbaarheid (en betrouwbaarheid) van het

³⁰ Fundamenteel biomedisch onderzoek is daarom een voorbeeld van *oriented basic research*, in onderscheid van *pure basic research* zoals in het voorbeeld van optogenetica in voetnoot 6 (Eurostat 2021).

EAE-model als zodanig.³¹ Dit laatste wordt bepaald door de methode waarmee het desbetreffende diermodel gerealiseerd wordt. Dus, bij interne validiteit is er sprake van zowel een algemeen-experimenteel aspect, als van een aspect eigen aan het diermodel als zodanig.

Hetzelfde geldt voor externe validiteit. Ferreira et al. (2020) wijzen erop dat bij het criterium van externe validiteit niet alleen de experimentele proefopzet een rol speelt – zoals bijvoorbeeld relevant eindpunten – maar ook de diermodellen als zodanig. En ook Henderson et al. zeggen iets soortgelijks:

Terwijl bedreigingen met betrekking tot externe validiteit geadresseerd kunnen worden door replicatiestudies onder variërende experimentele condities, worden bedreigingen met betrekking tot constructvaliditeit gereduceerd door het articuleren, adresseren, en bevestigen van theoretische vooronderstellingen die onder de klinische generalisatie liggen. (Henderson e.a. 2013, 2)

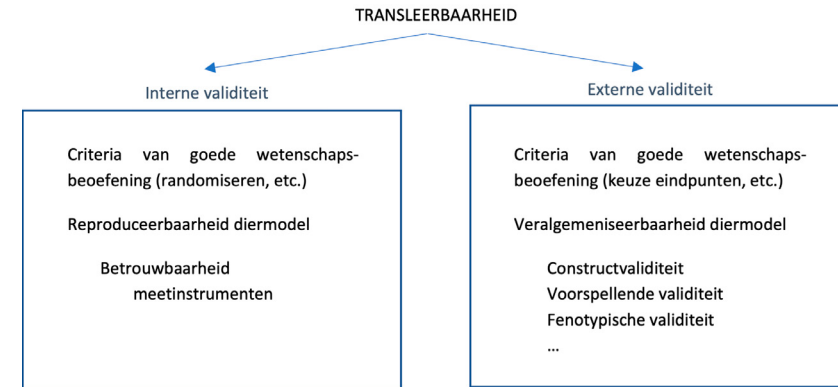
Bij de externe validiteit van dierproeven spelen dus algemeen-experimentele aspecten een rol, zoals de juiste keuzen van de eindpunten en het uitvoeren van replicatiestudies onder variërende experimentele condities (vgl. Voelkl e.a. 2021).³² Dit vergroot de kans op veralgemeeniseerbaarheid van de betreffende dierproef. Deze veralgemeeniseerbaarheid is echter ook afhankelijk van het gebruikte diermodel als zodanig: het zijn de drie klassieke validiteitscriteria van fenotypische, voorspellende en constructvaliditeit, en nieuwere validiteitsprincipes als homologe validiteit, die een beoordeling van dit laatste aspect mogelijk maken.^{33,34} Een en ander is samengevat in Figuur 1.

³¹ Daar komt dan nog het punt van externe validiteit bij.

³² Dit laatste is bijvoorbeeld relevant om te corrigeren voor de grote mate van genetische homogeniteit binnen lijnen van proefdieren als muizen en ratten, in tegenstelling tot de grote mate van genetische variëteit binnen de menselijke populatie.

³³ Hoewel Garner meerdere dimensies van validiteit onderscheidt, wordt het bij hem niet duidelijk hoe deze dimensies zich tot elkaar verhouden. Dit is wel van belang als transleerbaarheid als criterium gehanteerd wordt bij het prioriteren van dierproeven. Immers, staan de verschillende validiteitscriteria simpelweg naast elkaar en zou een diermodel idealiter goed moeten scoren op ieder validiteitscriterium? Of is er een bepaalde hiërarchie aan te brengen tussen de verschillende validiteitscriteria? Op basis van wat hier geschreven wordt, lijkt dat laatste mogelijk.

³⁴ In termen van Pound en Ritskes-Hoitinga (2018) betreffen de algemeen-experimentele aspecten 'overkomelijke problemen' bij het verbeteren van de translatie van resultaten van dierproeven, en zijn het de meer theoretische aspecten van het diermodel, en dan met name soortverschillen, die 'onoverkomelijke problemen' genereren (vgl. PETA 2020).



Figuur 1. Samenhang tussen validiteitscriteria met betrekking tot de transleerbaarheid van dierproeven.

De inzet van dit rapport is het opstellen van criteria die behulpzaam zijn bij het stellen van prioriteiten in het vervangen van dierproeven. Een van deze criteria is de transleerbaarheid van onderzoek. De transleerbaarheid is afhankelijk van de interne en externe validiteit van dierproeven. Hierbij valt zowel bij interne als bij externe validiteit een algemeen-experimenteel aspect te onderscheiden en een aspect eigen aan het diermodel als zodanig. Voor het praktisch prioriteren bij het vervangen van dierproeven, impliceert dit twee zaken:

1. In algemene zin is het zo dat dierproeven die niet voldoen aan de algemeen-experimentele eisen aan onderzoek, zoals randomisering en blinding, ongewenst zijn. De algemeen-experimentele eisen zijn belangrijk als het gaat om goede wetenschapsbeoefening en richtlijnen die dit voorschrijven zouden dan ook geaccordeerd moeten worden door iedere wetenschapsinstelling (vgl. Henderson e.a. 2013; Hooijmans, Leenaars, en Ritskes-Hoitinga 2010). Toch vormen ze geen criterium als het gaat om het stellen van prioriteiten in het vervangen van dierproeven, juist omdat ze zo generiek zijn voor het doen van onderzoek. Met andere woorden, algemeen-experimentele criteria kunnen niet onderscheiden tussen diermodellen.
2. Dit is anders voor het diermodel eigen aspect. De transleerbaarheid van een diermodel hangt af van de reproduceerbaarheid (intern) en veralgemeeniseerbaarheid (extern). Hierbij hangt de reproduceerbaarheid van een diermodel mede af van de methode waarmee een specifiek diermodel gerealiseerd wordt, terwijl de veralgemeeniseerbaarheid van een diermodel mede afhangt van de mate waarin het model de

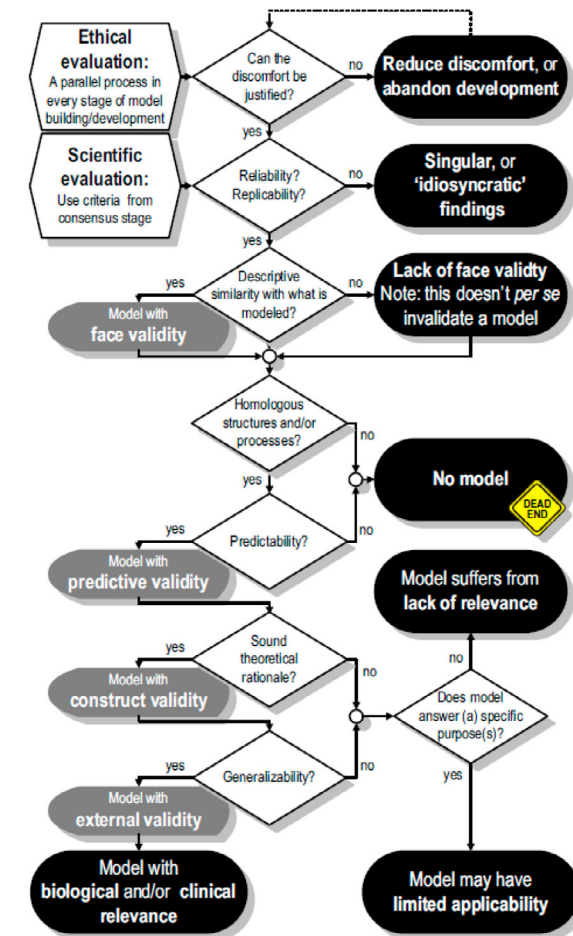
veronderstelde achterliggende mechanismen van een biologisch of etiologisch fenomeen in de mens simuleert.³⁵ Deze reproduceerbaarheid en veralgemeniseerbaarheid, als subcriteria van transleerbaarheid, verschillen dan ook op een generiek niveau per diermodel en kunnen bijgevolg gehanteerd worden als subcriteria onder transleerbaarheid bij het prioriteren in het vervangen van dierproeven (zie opnieuw Figuur 1).

Tools om de validiteit van diermodellen te bepalen

Een belangrijke vraag is hoe de validiteit van diermodellen in het algemeen, en uitgesplitst voor de verschillende deelgebieden (fenotype, mechanisme, etc.), bepaald kan worden. Willner (1991) was van mening dat dit een (waarde)oordeel betrof en niet iets wat gemeten kan worden. Recent zijn er echter tools ontwikkeld om de validiteit en transleerbaarheid van diermodellen op een gestructureerde manier te bepalen. Dit betreft achtereenvolgens de modellen van Van der Staay, Arndt en Nordquist (2009), Sams-Dodd (2006) – verder ontwikkeld door Denayer et al. (2014) –, Ferreira et al. (2019; 2020) en Wehling (Wehling 2009; Wendler en Wehling 2012; 2017). Van der Staay, Arndt en Nordquist's tool betreft een *kwantitatieve* tool, in die zin dat de uiteindelijke beoordeling van een diermodel aan de hand van vooraf gespecificeerde criteria plaatsvindt op basis van consensus. Binnen de tools van Sams-Dodd/Denayer et al., Ferreira et al., en Wehling vindt de uiteindelijke beoordeling plaats op basis van de *kwantitatieve* score die een diermodel toegekend krijgt aan de hand van vooraf gespecificeerde criteria. De verschillende tools zullen nu kort toegelicht worden.

Tools van Van der Staay, Arndt en Nordquist (2009)

Van der Staay, Arndt en Nordquist (2009) voorzien in een stroomdiagram waarin de wetenschappelijke evaluatie begint met het bepalen van de betrouwbaarheid en reproduceerbaarheid van een diermodel (de interne validiteit), gevolgd door het bepalen van de fenotypische validiteit, de voorspellende validiteit, de constructvaliditeit, en de externe validiteit (zie Figuur 2).



Figuur 2. Stroomdiagram ter evaluatie van een diermodel. Overgenomen op basis van Creative Commons Attribution License CC by 2.0 uit Van der Staay, Arndt en Nordquist (2009).

³⁵ Dit maakt construct- of mechanistische validiteit wetenschappelijk gezien tot het belangrijkste criterium, maar het is tegelijkertijd het moeilijkst te beoordelen aspect van een diermodel (Holmes 2003). Dat laatste is de reden waarom dikwijls afgegaan wordt op fenotypische validiteit, aangezien symptomen vergeleken kunnen worden tussen modeldieren en humane patiënten.



Tool van Sams-Dodd (2006)/Denayer et al. (2014)

Mede vanwege het feit dat het klassieke gebruik van de verschillende validiteitscriteria het niet toelaat om diermodellen gradueel te beoordelen (zie bijvoorbeeld het zojuist besproken stroomdiagram van Van der Staay, Arndt en Nordquist) en het niet mogelijk is om verschillende modellen te vergelijken (bijvoorbeeld *in vivo* versus *in vitro* versus *in silico*) heeft Sams-Dodd (2006) een voorstel gedaan voor het gradueel beoordelen van verschillende modellen.

Sams-Dodd werkt met drie variabelen: het gebruikte organisme (bijvoorbeeld muis of mens), de complexiteit van de gebruikte structuur (bijvoorbeeld cel of orgaan of geheel organisme) en de manier waarop de ziekte zich ontwikkelt (bijvoorbeeld gesimuleerd door chemische toediening of 'als vanzelf' door genetische aanleg). Validiteit beoordeelt Sams-Dodd nu in termen van geen, laag, gemiddeld, en hoog. Het gebruik van primaten krijgt bijvoorbeeld de beoordeling 'gemiddeld' voor organisme, het gebruik van cellijnen 'laag' voor complexiteit, en onderzoek met humane patiënten 'hoog' voor het proces van ziekteontwikkeling.

Om te komen tot een totaalbeoordeling over de drie variabelen heen, heeft Sams-Dodd een algoritmisch scoresysteem ontwikkeld. Het gebruik van weefsel van muizen *ex vivo* waarbij de ziekte geïnduceerd is met behulp van chemische middelen, krijgt bijvoorbeeld een totaalscore 3 omdat het op alle drie de criteria afwijkt van de klinische situatie (organisme is muis in plaats van mens, complexiteit is weefsel in plaats van geheel organisme, en ziekte-inductie is kunstmatig in plaats van natuurlijk/werkelijk). De hoogste score (0) betreft onderzoek met de humane patiënt zelf (humaan, intact organisme, werkelijk ziekteproces). Alle andere mogelijke combinaties variëren daarmee tussen 0 en 3. Denayer et al. (2014) hebben Sams-Dodd's algoritmische scoresysteem verder ontwikkeld door de criteria van fenotypische validiteit en voorspelbaarheid (i.e., voorspellende validiteit) toe te voegen, en de hoogst mogelijke score per criterium te verhogen tot vier. Bij fenotypische validiteit, bijvoorbeeld, geeft geen overeenkomst in symptomen tussen de klinische praktijk en het diermodel een score 1, één overeenkomst een 2, één kernsymptoom een 3, en meer dan één kernsymptoom een 4. Hoe hoger de score over alle criteria, hoe groter de waarschijnlijke validiteit van het diermodel.

Tool van Ferreira et al. (2019; 2020)

Vanwege het feit dat de klassieke validiteitscriteria multi-interpretabel zijn, beogen Ferreira et al. (2019; 2020) naar eigen zeggen een objectief scoresysteem om verschillende diermodellen te kunnen vergelijken in termen van validiteit. Het bestaande systeem van

Sams-Dodd en Denayer et al. ziet volgens hen meerdere aspecten over het hoofd, zoals histologie en biomarkers. Op basis van literatuurstudie en gesprekken met stakeholders stellen Ferreira et al. een Framework to Identify Models of Disease (FIMD) voor met daarbinnen acht domeinen: epidemiologie, symptomatologie en natuurlijke geschiedenis, genetica, biochemie, etiologie, histologie, farmacologie, en eindpunten (zie Tabel 2). Bij ieder van deze domeinen worden gesloten kernvragen geformuleerd (ja/nee). Per vraag kan bij een positief antwoord een vaste score behaald worden; de maximale score per domein is gelijk aan die van de andere domeinen. Hoe hoger de totaalscore, hoe groter de waarschijnlijke validiteit van het diermodel.

Tabel 2. Het Framework to Identify Models of Disease (FIMD). Overgenomen op basis van Creative Attribution Licence CC by 4.0 uit Ferreira et al. (2019). Vertaling C.J. Rademaker.

	Gewicht
1. Epidemiologische validatie	12,5
1.1 Kan het model de ziekte simuleren in de relevante sekse?	6,25
1.2 Kan het model de ziekte simuleren in de relevante leeftijdsgroepen (e.g., juveniel, volwassen of verouderend)?	6,25
2. Symptomatologie en natuurlijke geschiedenis validatie	12,5
2.1 Kan het model de symptomen en comorbiditeiten die gewoonlijk aanwezig zijn bij de ziekte repliceren? Zo ja, welke?	2,5
2.2 Is de natuurlijke geschiedenis van de ziekte hetzelfde als bij de mens, met betrekking tot:	
2.2.1 Tijd tot aanvang	2,5
2.2.2 Ziekteprogressie	2,5
2.2.3. Duur van de symptomen	2,5
2.2.4 Ernst	2,5
3. Genetische validatie	12,5
3.1 Heeft dit diersoort ook orthologe genen en/of eiwitten die betrokken zijn in de menselijke ziekte?	4,17
3.2 Zo ja, zijn de relevante genetische mutaties or veranderingen ook aanwezig in de orthologe genen/eiwitten?	4,17
3.3 Zo ja, is de expressie van zulke orthologe genen en/of eiwitten hetzelfde als in de menselijke situatie?	4,16
4. Biochemische validatie	
4.1 Als er bekende farmacodynamische biomarkers gerelateerd aan de pathofysiologie van de ziekte zijn, zijn deze ook aanwezig in het model?	3,125
4.2 Gedragen deze farmacodynamische biomarkers zich hetzelfde als die in mensen?	3,125
4.3 Als er bekende prognostische biomarkers gerelateerd aan de pathofysiologie van de ziekte zijn, zijn deze ook aanwezig in het model?	3,125
4.4 Gedragen deze prognostische biomarkers zich hetzelfde als die in mensen?	3,125
5. Etiologische validatie	12,5
5.1 Is de etiologie van de ziekte hetzelfde als in mensen?	12,5

	Gewicht
6. Histologische validatie	12,5
6.1 Zijn de histopathologische structuren in de relevante weefsels hetzelfde als die gevonden worden in mensen?	12,5
7. Farmacologische validatie	12,5
7.1 Zijn effectieve medicijnen in mensen ook effectief in dit model?	4,17
7.2 Zijn ineffectieve medicijnen in mensen ook ineffectief in dit model?	4,17
7.3 Zijn medicijnen met verschillende werkingsmechanismes en aangrijpend op verschillende <i>pathways</i> getest in dit model? Zo ja, welke?	4,16
8. Validatie van eindpunten	12,5
8.1 Zijn de eindpunten die gebruikt worden in preklinische studies hetzelfde of vertaalbaar naar de klinische eindpunten?	6,25
8.2 Zijn de methoden die gebruikt worden om de preklinische eindpunten te beoordelen vergelijkbaar met degenen die gebruikt worden om klinische eindpunten te beoordelen?	6,25

Tool Van Wehling (Wehling 2009; Wendler en Wehling 2012; 2017)

Een laatste scoringsysteem betreft het model van Wehling dat focust op het in een vroege fase beoordelen van het translatiepotentieel van farmacologisch onderzoek (Wehling 2009; Wendler en Wehling 2012; 2017). Hierbij worden vijf dimensies onderscheiden: startbewijs (bijvoorbeeld: wat is al bekend vanuit *in vitro* data?), humaan bewijs (bijvoorbeeld uit klinische trials); biomarkers voor het voorspellen van effectiviteit en veiligheid (bijvoorbeeld: hoe 'dicht bij' het ziektebeeld ligt een biomarker?); testen van *proof-of-mechanism*, *proof-of-principle* en *proof-of-concept* (bijvoorbeeld: helderheid en acceptatie van surrogaat-eindpunten); en aspecten gelinkt aan gepersonaliseerde geneeskunde (bijvoorbeeld: wat is bekend over de impact van genetische variatie op farmacokinetiek en farmacodynamiek). De subonderdelen binnen de verschillende dimensies hebben allen een verschillend gewicht, en in de recentste publicatie is dit gespecificeerd naar ziektegebied. Zo kennen Wendler en Wehling (2017) op basis van de recentste inzichten het meeste gewicht toe aan de dimensie 'biomarkers voor het voorspellen van effectiviteit en veiligheid' bij oncologisch onderzoek, maar heeft bij psychiatrisch onderzoek de dimensie 'menschelijk bewijs' het grootste gewicht. Het gewicht per subonderdeel wordt vermenigvuldigt met een score die gegeven kan worden tussen 1 en 5, waarbij 1 staat voor zwak en 5 voor sterk. Wehling (2009) geeft daarbij kenmerken van wat telt als eigenschappen voor respectievelijk zwakke en sterke scores. Na optellen van alle afzonderlijke scores wordt dit getal gedeeld door 100, wat een eindscore geeft tussen de 1 en 5. Daarbij wordt gesteld dat een score van 4 of hoger staat voor een goede vertaalbaarheid.



Keuze voor een tool

In de context van dit rapport voert het te ver om de verschillende tools ter bepaling van de validiteit van diermodellen te vergelijken en te komen tot een voorstel voor het gebruik van een bepaalde tool. Het doel was een overzicht te geven van tools die behulpzaam kunnen zijn bij het prioriteren in het vervangen van dierproeven als het gaat om transleerbaarheid. Idealiter zal in samenspraak met wetenschappers uit het veld een nadere keuze gemaakt worden voor het in de praktijk op gestructureerde wijze bepalen van de transleerbaarheid van dierproeven.

Is translatie überhaupt mogelijk?

Sinds de publicatie van LaFollette en Shanks's *Brute Science: Dilemmas of Animal Experimentation* (1996) staat de *mogelijkheid* van het transleren van inzichten verworven middels een diermodel naar de mens fundamenteel ter discussie. LaFollette en Shanks geven aan dat het louter presenteren van succesvolle dan wel falende gevallen van translatie – waarbij voorstanders van proefdieronderzoek succesvolle gevallen presenteren en tegenstanders van proefdieronderzoek falende gevallen – onvoldoende is om te oordelen over de mogelijkheid van translatie überhaupt. Het is namelijk inherent aan het doen van onderzoek dat negatieve resultaten gevonden worden. Er zal doorgestoten moeten worden naar wat falende gevallen van translatie betekenen: reflecteren ze een methodologisch defect of zijn ze onderdeel van 'normale wetenschap' (Kuhn)? LaFollette en Shanks zijn van mening dat er inderdaad zo'n methodologisch defect aanwezig is in proefdieronderzoek. Er is een ontologisch probleem dat bestaat in de causale disanalogieën tussen modeldier en mens: causale mechanismen zijn vrijwel nooit identiek tussen modeldier en mens. Volgens LaFollette en Shanks kunnen diermodellen daarom alleen dienen als *hypothetische diermodellen* en niet als *causale diermodellen*. Omdat LaFollette en Shanks stellen dat er geen causale disanalogieën mogen bestaan tussen modeldier en mens om de translatiestap te kunnen maken, trekken zij de conclusie dat diermodellen hoogstens kunnen dienen ter generatie van hypothesen voor humaan onderzoek.³⁶ Dit stelt mede de ethische aanvaardbaarheid van proefdieronderzoek op losse schroeven. Daar komt nog een epistemologisch punt bij: of een diermodel externe validiteit bezit, kan niet vastgesteld worden zonder al te weten wat we juist hopen te leren van de extrapolatie. Met andere woorden, of aan de ontologische claim van vergelijkbaarheid voldaan wordt, kunnen we volgens LaFollette en Shanks pas weten als we de resultaten van

proefdieronderzoek vergelijken met de resultaten van humaan onderzoek.³⁷ Dit wordt wel de *extrapolator's circle* genoemd (Steel 2010).

Hoewel het te ver voert om uitgebreid op deze materie in te gaan, is het goed om op te merken dat LaFollette en Shanks' argument niet onweersproken is gebleven. Nordgren (2010), bijvoorbeeld, haalt werk van Steel over *comparative process tracing* aan, dat aangeeft dat een causaal mechanisme niet volledig identiek hoeft te zijn tussen modeldier en mens, maar dat gekeken moet worden naar de overeenkomsten in fasen binnen dit causale mechanisme (vgl. Steel 2010). Verder suggereert Nordgren dat een combinatie van *in vivo*, *in vitro* en *in silico* methoden de mogelijkheden voor extrapolatie kan versterken, en dat ook het gebruik van alleen proefdiervrije methoden serieuze beperkingen kent. Als laatste merkt Nordgren op dat LaFollette en Shanks' benadering te sterk theoriegedreven is en te weinig recht doet aan het pragmatische en voorlopige karakter van wetenschapsbeoefening. Aansluitend hieraan bekritisieren Degeling en Johnson (2013) specifiek LaFollette en Shanks' onderscheid tussen causale diermodellen en hypothetische diermodellen, omdat het de waarde van hypothetische diermodellen devalueert en de ethische agenda van LaFollette en Shanks verraadt. Degeling en Johnson stellen voor om in plaats daarvan het onderscheid – dat in de dagelijkse proefdierpraktijk al gemaakt wordt (Hau en Schapiro 2011) – tussen verkennende, verklarende en voorspellende diermodellen te hanteren. In dit laatste zijn ook de validiteitscriteria van respectievelijk fenotypische, construct- en voorspellende validiteit te herkennen. In deze lijn van Degeling en Johnson kan dan ook het criterium van transleerbaarheid gehanteerd worden bij het prioriteren in het vervangen van dierproeven.

Samenvatting

Transleerbaarheid (vertaalbaarheid) gaat over de vraag in hoeverre de data verkregen uit een dierexperiment vertaalbaar zijn naar de mens. In het geval de probleemstelling van een onderzoek vertrekt vanuit een maatschappelijke praktijk als de geneeskunde, is transleerbaarheid zowel voor toegepast onderzoek als voor fundamenteel onderzoek binnen dat domein een relevant criterium. Bij toegepast onderzoek gaat het om de transleerbaarheid van de effectiviteit van een product (is een medicijn of hulpmiddel dat werkzaam is bij een proefdier dat ook bij de mens?). Bij fundamenteel onderzoek gaat het om de transleerbaarheid van kennis met betrekking tot een specifiek mechanisme naar de klinische situatie (heeft het model klinische relevantie?).

³⁶ In de wetenschappelijke vakliteratuur wordt dit punt van ontologische verschillen ook aangeroerd, zie bijvoorbeeld A. Knight (2019) en Pound & Ritskes-Hoitinga (2018).

³⁷ Voor LaFollette en Shanks is dit vooral een probleem omdat volgens hun wetenschappers van mening zijn dat humaan onderzoek "scientifically second-rate" is (LaFollette en Shanks 1996, 23).



Transleerbaarheid heeft met validiteit te maken: een diermodel (van een ziekte) is valide wanneer het een goed model vormt van de menselijke ziekte of aandoening, en meer specifiek wanneer onderliggende mechanismen overeenkomen. Als het gaat om de validiteit van proefdieronderzoek worden veelal experimentele aspecten als de noodzaak tot randomiseren en blinderen genoemd. Hoewel deze kwaliteitseisen uiteraard van belang zijn - dierproeven die niet voldoen aan hoge wetenschappelijke standaards zijn sowieso ongewenst -, zeggen deze eisen zeggen nog niet iets over de bruikbaarheid van het *diermodel* mocht het experiment in overeenstemming met hoge wetenschappelijke standaards zijn uitgevoerd. Reproduceerbaarheid en veralgemeniseerbaarheid als subcriteria van validiteit zijn echter *wel* relevant voor het beoordelen van het diermodel zelf. Bij *reproduceerbaarheid* gaat het om de vraag of het model consistente en betrouwbare resultaten oplevert. Dus: wordt er gemeten wat er gemeten moet worden? Bij *veralgemeniseerbaarheid* draait het om de vraag of de resultaten behaald met het diermodel te extrapoleren zijn naar de menselijke situatie. Dus: vormt het diermodel inderdaad een goed model voor het te bestuderen fenomeen in de mens? Onderdeel van veralgemeniseerbaarheid zijn de drie klassieke criteria van fenotypische, construct- en voorspellende validiteit.

Verschuillende tools zijn ontwikkeld om de validiteit van diermodellen in te schatten dan wel te meten. Het beste kan in samenspraak met het werkveld een keuze gemaakt worden voor een tool of een combinatie van tools.

Dierproeven waarbij diermodellen gebruikt worden die een lage mate van transleerbaarheid hebben, komen als eerste in aanmerking om vervangen te worden, mede afhankelijk van de beoordeling op andere criteria, zoals de mate van relevantie van het doel van een dierproef.

Voorbeeld transleerbaarheid

GEDWONGEN ZWEMTEST

De Gedwongen Zwemtest (Forced Swim Test; FST) wordt sinds de jaren '70 van de vorige eeuw gebruikt binnen depressieonderzoek. De test bestaat uit het plaatsen van een knaagdier in een bak met lauwwarm water van waaruit het niet kan ontsnappen. Het dier zal actief rond gaan zwemmen en proberen te ontsnappen, maar zal dit op een gegeven moment opgeven en zich passief laten drijven. Deze tijdsduur van immobiliteit is de primaire uitkomstparameter van de test. Omdat klinisch gebruikte antidepressiva een reductie geven in deze tijdsduur van immobiliteit wanneer toegepast in knaagdieren, wordt een verminderde tijdsduur van immobiliteit geïnterpreteerd als bewijs van de werkzaamheid van nieuwe geteste middelen (Porsolt, Le Pichon, en Jalfre 1977).

De test staat sterk ter discussie, voornamelijk vanwege de beperkte transleerbaarheid naar de mens voor wat betreft depressie (voor een review, zie Trunnell en Carvalho 2021). Of de FST fenotypische validiteit toegekend kan worden, hangt af van de interpretatie van het drijfgedrag van de knaagdieren. Wordt deze geïnterpreteerd in de lijn van Porsolt, Le Pichon, en Jalfre (1977) als uiting van wanhoop – een signaal van depressie in de mens – dan bezit de FST fenotypische validiteit. Deze interpretatie staat echter onder druk; onder de alternatieve interpretatie dat het drijfgedrag een copingstrategie is, bezit de FST geen fenotypische validiteit voor wat betreft depressie (Commons e.a. 2017; Molendijk en De Kloet 2015; 2019; De Kloet en Molendijk 2016; Molendijk en De Kloet 2021). Verder heeft de FST geen etiologische validiteit, want de test van 15 minuten (rat) of 6 minuten (muis) verschilt sterk van de trage pathogenese van depressie in mensen (Molendijk en De Kloet 2015). Bovendien worden antidepressiva gegeven aan gezonde dieren, wat niet de klinische praktijk reflecteert. Nauwverwant hieraan is het ontbreken van constructvaliditeit, want de achterliggende mechanismen van depressie zijn onbekend (Sewell e.a. 2021). Of de FST (enige mate van) voorspellende validiteit heeft staat sterk ter discussie (hoge mate: Slattery en Cryan 2012; geen: Trunnell en Carvalho 2021). Op basis van het criterium van transleerbaarheid is de FST binnen depressieonderzoek dus een voorbeeld van een test die versneld vervangen zou moeten worden. Dit is al gaande binnen onderzoeksinstellingen en de industrie (PETA 2022; Ritskes-Hoitinga, persoonlijke communicatie, 30 september, 2021).



Relevantie van het doel van een dierproef

Proefdieronderzoek wordt uitgevoerd met een doel. Hierbij wordt onderscheid gemaakt tussen het directe en het uiteindelijke doel (CCD 2019).³⁸ Het directe doel betreft het doel dat verwekelijkt kan worden binnen de context van een dierproef, terwijl het uiteindelijke doel verder weg in de toekomst ligt. Het directe doel van alzheimeronderzoek kan bijvoorbeeld zijn het nader ontrafelen van het mechanisme achter de vorming van amyloïd-beta-plaques tussen zenuwcellen in de hersenen. Het uiteindelijke doel betreft dan het bijdragen aan het voorkomen, genezen, of dragelijk maken, van de ziekte van Alzheimer. Het zijn de vele experimenten en onderzoeken tezamen in een vakgebied die een substantiële bijdrage kunnen leveren aan dit uiteindelijke doel.

Typen en varianten van proefdieren onderzoek

Bij de verschillende uiteindelijke doelen die onderzoek kan dienen, spelen twee factoren een rol. Dit is ten eerste het *type* doel dat nagestreefd wordt met het onderzoek. Een precedent is in 2013 gezet met een verbod op dierproeven ten bate van cosmetische doeleinden (hoewel uitzonderingen daarop mogelijk zijn) (European Parliament and the Council 2019). Het belang van proefdieronderzoek voor cosmetische doeleinden scoort ook in enquêtes en onderzoeken laag, en niet alleen in een Europese context (Cameron, Clemence, en Xypolia 2018; S. Knight en Barnett 2008; T. B. Lund e.a. 2014; Sandgren e.a. 2020). Uit deze onderzoeken blijkt ook dat respondenten veiligheidstesten van chemicaliën ter gebruik in huishoudens en bedrijfscontexten maar nauwelijks hoger waarderen dan cosmetisch onderzoek. Onderzoek ten bate van de gezondheid van mens en dier scoort consistent het hoogst (zie ook Ormandy en Schuppli 2014). Sandgren et al.'s (2020) enquête onder faculteitspersoneel en studenten van de University of Wisconsin-Madison in de Verenigde Staten van Amerika nam ook landbouwkundig onderzoek mee als categorie.³⁹ Respondenten vonden dergelijk proefdieronderzoek beter te rechtvaardigen dan cosmetisch onderzoek, maar slechter te rechtvaardigen dan onderzoek ten bate van de gezondheid van mens en dier. Nu zegt de categorie van

landbouwkundig onderzoek niet alles. Het ligt namelijk voor de hand om te veronderstellen dat de mate van acceptatie van landbouwkundig onderzoek verschilt per *variant* van landbouwkundig onderzoek. In Nederland heeft in 2018 bijvoorbeeld de toenmalige staatssecretaris van Economische Zaken, Van Dam, op instigatie van de CCD een advies gevraagd aan de Raad voor Dieraangelegenheden (RDA) met betrekking tot de wenselijkheid van het gebruik van dierproeven ten behoeve van de conventionele en intensieve veehouderij. De RDA (2018b) heeft daarbij geadviseerd dat proefdieronderzoek ten behoeve van de veehouderij idealiter zou moeten bijdragen aan gezondheid en welzijn van landbouwhuisdieren en verduurzaming van de veehouderij. Hierin is geïmpliceerd dat bijvoorbeeld onderzoek naar het gebruik van regionaal gewonnen eiwitbronnen in de voeders van melkkoeien beter te rechtvaardigen valt dan onderzoek naar de inclusie van een tropisch gewas in de voeders van zeugen in kraamhokken. Meer in het algemeen gesteld, verschilt de mate van acceptatie van onderzoek dus ook naar gelang de variant binnen een bepaald type onderzoek. Dit betreft de tweede factor.

Beoordeling van verschillen typen en varianten van proefdierenonderzoek

Ook binnen het medische onderzoek kunnen verschillende waarderings van varianten van onderzoek onderscheiden worden. Zo laat onderzoek zien dat mensen het gebruik van proefdieren ten behoeve van bijvoorbeeld onderzoek naar ziekten als de ziekte van Alzheimer en kanker acceptabel vinden dan onderzoek gericht op obesitas (T. B. Lund, Lassen, en Sandøe 2012; T. B. Lund e.a. 2014). Wat hierbij voor respondenten een rol speelt, is dat obesitas een aandoening is die in principe te voorkomen is – door middel van aanpassingen in levensstijl – en dat het voorkomen van obesitas in het beeld van mensen een persoonlijke verantwoordelijkheid betreft (Freriks e.a. 2005; T. B. Lund, Lassen, en Sandøe 2012). Andere factoren zijn de mortaliteit en morbiditeit verbonden aan een ziekte, en in hoeverre er al bestaande middelen zijn om het leven met een ziekte of aandoening draaglijk te maken. Migraine bijvoorbeeld wordt in dat opzicht een minder zwaar gewicht toegekend dan de ziekte van Alzheimer en kanker (T. B. Lund, Lassen, en Sandøe 2012).

³⁸ Doelen zijn nauwverwant aan de baten van onderzoek (Niemi 2021). Analooq aan het onderscheid tussen directe en uiteindelijke doelen, spreken Brønstad et al. (2016) bijvoorbeeld van actuele en beloofde baten. Het voert te ver om de precieze verhouding tussen baten en doelen hier nader uit werken.

³⁹ In het achterhoofd moet gehouden worden dat dergelijk onderzoek onder faculteitspersoneel en studenten ernstige beperkingen met zich meebrengt voor wat betreft generaliseerbaarheid naar een breder publiek (Ormandy en Schuppli 2014).



Het voorbeeld van obesitas als een zogenaamde *lifestyle disease* laat echter ook zien dat dit een lastige discussie is. Zoals Freriks et al. (2005, 68) terecht opmerken met betrekking tot de vraag of dierproeven gedaan zouden mogen worden voor het ontwikkelen van geneesmiddelen tegen overgewicht en obesitas:

Op het eerste gezicht lijkt het moreel gepaste antwoord niet moeilijk. Dierenwelzijnsorganisaties zullen de vraag ontkennend beantwoorden, maar andere partijen zullen geneigd zijn de nodige nuances aan te brengen. Niet alle gevallen van overgewicht, zo zullen zij aanvoeren, zijn op het conto van de eigen verantwoordelijkheid te schuiven. Het zou wel heel cru zijn om de 'onschuldige' slachtoffers van deze aandoening een potentieel middel te onthouden alleen omdat anderen de kwaal aan zichzelf te wijten hebben.

Hier komt bij dat sociale en verslavingsfactoren een rol kunnen spelen. Deze moeilijkheid kwam ook naar voren in de discussie aan het begin van dit millennium over de vraag of proefdieronderzoek ten bate van aandoeningen gekoppeld aan roken ethisch wel te rechtvaardigen valt. In het jaarverslag *Zo doende 2003* van de toenmalige Voedsel- en Warenautoriteit wordt melding gemaakt van een casus die speelde binnen een DEC (VWA 2004). Het betrof onderzoek naar de schadelijke effecten van aldehydes in sigarettenrook. Dergelijk onderzoek is (potentieel) "ethisch problematisch omdat roken een slechte gewoonte is die vermeden kan worden en maatschappelijk bestreden wordt" (Freriks e.a. 2005, 69). Dit kan echter wederom genuanceerd worden: het is namelijk de vraag "in hoeverre men bij een verslavingsziekte zoals roken, welke een 'lifestyle' disease wordt genoemd, nog kan spreken van een eigen verantwoordelijkheid" (Schurgers 2005, 35). Bovendien:

In hoeverre hebben mensen een ziekte aan hun eigen, bewust risicovolle gedrag te wijten? Belangrijk is dat er in gevoelige periodes in het leven, zoals de puberteit, andere redenen zijn om te gaan roken dan wanneer men hier pas op volwassen leeftijd mee begint. Daarnaast zijn er natuurlijk nog meerokers. Wanneer de DEC besluit [sic] dat er geen dierproeven ten behoeve van rokers gedaan moeten worden, ontnemt ze ook deze twee laatste groepen het zicht op een behandeling (Schurgers 2005, 37).

Terwijl dierrechtenorganisaties wellicht sneller geneigd zijn om te oordelen dat dergelijk onderzoek voor levensstijlgerelateerde aandoeningen onacceptabel is, is er altijd wel een nuance te vinden waarom een dergelijke variant van onderzoek toch doorgang zou moeten vinden (Freriks e.a. 2005).

Een andere moeilijkheid met betrekking tot het bepalen van de relatieve waarde van onderzoek heeft te maken met het feit dat inzichten verkregen uit een type onderzoek, of variant van een type onderzoek, breder relevant kan zijn. Zo kan proefdieronderzoek naar de etiologie van obesitas inzichten opleveren die breder relevant zijn dan alleen voor mensen die momenteel lijden aan obesitas. Misschien moet proefdieronderzoek ten bate van humane voeding zelfs wel hoger aangeslagen worden dan medisch onderzoek, omdat voedingsinterventies ingezet kunnen worden ter versterking van het immuunsysteem en daarmee mogelijk mede beschermend kunnen zijn tegen allerhande ziekten en aandoeningen (de zogenaamde nutraceuticals).

Een ander voorbeeld is onderzoek naar het in gevangenschap kweken van paling (CCD 2011). De huidige palinghouderij is afhankelijk van in het wild gevangen glasaal. Deze glasaal wordt op palingboerderijen afgemest tot consumptieformaat. Het grote probleem met de paling is echter dat deze zowel een bedreigde diersoort is als dat natuurlijke reproductie in gevangenschap niet mogelijk is. Er is echter een hormoonbehandeling ontwikkeld om eieren en hom te verkrijgen van ouderdieren, en larven tot ontwikkeling te laten komen. Mocht het daadwerkelijk lukken om glasaal op deze manier op te kweken, dan dient dit zowel het doel van behoud van de diersoort als de duurzaamheid van de palinghouderij. Het probleem is hier dat het behoud van de diersoort gewaardeerd kan worden als van groot belang, en dat de productie van glasaal ter afmesting in palinghouderijen lager gewaardeerd kan worden (het eten van paling is immers niet noodzakelijk en ook zijn er alternatieven). Tegelijk zijn beiden doel van het onderzoek.

Meer in het algemeen concluderen Olsson et al. (2015) dan ook dat het relatieve belang van verschillende (uiteindelijke) doelen van onderzoek moeilijk te evalueren is, omdat er zoveel vraagstukken bij betrokken zijn, en er geen overeenstemming is over hoe verschillende doelen gerangschikt zouden moeten worden. En ook in het Europese werkdocument 'Project Evaluation and Retrospective Assessment' wordt gesteld dat er geen overeenstemming is over het relatieve belang van het gebruik van dieren ten behoeve van bijvoorbeeld educatieve doeleinden, wettelijk voorgeschreven veiligheidstesten, menselijke gezondheid of dierlijke gezondheid, en dat er daarom geen eenvoudige hiërarchische orde te geven is (Expert Working Group for Project Evaluation and Retrospective Assessment 2011).



Het is waar dat een eenvoudige en precieze hiërarchie van de uiteindelijke doelen van onderzoek moeilijk te geven is. Toch kan er wel degelijk een trend onderscheiden worden. Het onderscheid dat VanDeVeer (1995) maakt tussen basale, serieuze en perifere belangen kan hierbij behulpzaam zijn. *Basale belangen* gaan over specifieke condities of goederen die benodigd zijn voor mens en dier om “minimaal adequaat” functioneren mogelijk te maken, zoals voedsel, water, zuurstof of de afwezigheid van chronische, intense pijn (VanDeVeer 1995, 88).⁴⁰ Voor traditioneel levende Inuit betreft de jacht op zeehonden bijvoorbeeld een basaal belang, terwijl voor mensen in moderne landbouwrijke gebieden het houden van dieren voor de slacht dat niet is. *Serieuze belangen* betreffen goederen of condities zonder welke een mens of dier kan overleven, maar welke wel moeilijk of kostbaar voor diens welzijn zijn. Kippen hebben bijvoorbeeld een serieus belang bij stofbaden en varkens bij wroeten. *Perifere belangen*, als laatste, betreffen een goed of conditie zonder welke leven mogelijk is, maar een mens of dier niet floreert. Voorbeelden hier zijn het (cultuurrelatieve) gebruik van decoraties als ivoren hangers en tot op zekere hoogte het gebruik van verzorgingsartikelen.

Als we deze indeling van het gewicht van belangen vertalen naar doelen van onderzoek, dan zijn mogelijke voorbeelden van onderzoek met ‘basale doelen’ onderzoeken die gericht zijn op de preventie, genezing en het draagbaar maken van levensbedreigende ziekten en ziekten die ernstig belemmerend zijn voor de kwaliteit van leven van mens⁴¹ en dier, en het behoud van diersoorten en ecosystemen. Voorbeelden van serieuze doelen zijn onderzoeken gericht op duurzame vormen van landbouw en veehouderij en veiligheids-testen van chemicaliën benodigd in duurzame voedselproductie. Voorbeelden van perifere doelen betreffen onderzoeken naar de veiligheid van cosmetische producten (zijn al verboden in de EU), de productie van bont (bontproductie is sinds 2021 verboden in Nederland), productiedoeleinden verbonden aan de intensieve veehouderij, en veiligheidstesten voor chemicaliën ter gebruik in de plasticindustrie.

Het relatieve belang van deze doelen van onderzoek moet gezien worden in het licht van de huidige, moderne samenleving, en tegen de achtergrond van het gebruik van proefdieren. de huidige, moderne samenleving, en tegen de achtergrond van het gebruik van proefdieren. Al de doelen van onderzoek zijn op zich legitiem, maar *in het licht van het gebruik van proefdieren* kan een bepaalde orde aangebracht worden. Verder is het voor te stellen dat, bijvoorbeeld, in een tijd en samenleving waarin geen alternatieven bestaan voor het dragen van bont, en mensen sterven van de kou, onderzoek naar de productie van bont een basaal doel wordt.

Bij het wegen van het doel van onderzoek en de daaraan verbonden waarden is het ook van belang om uit te gaan van wat we nu weten en wat plausibel is om te veronderstellen. Zo is het redelijk om te stellen dat onderzoek naar de etiologie van obesitas in de eerste plaats ten goede zal kunnen komen aan mensen die lijden aan obesitas. De eventuele bredere toepassing van verworven inzichten van dergelijk onderzoek behoren tot de categorie ‘toevalstreffer’ (*serendipity*). De kleine kans daarop zou niet mee moeten wegen in het belang van het doel van een voorgesteld onderzoek.

Natuurlijk is het zo dat onderzoek meerdere doelen tegelijk kan dienen, zoals het voorbeeld van de paling duidelijk maakte. Dit is op zich echter nog geen probleem voor een ordening van de doelen van onderzoek. Een praktisch oplossing hier is om uit te gaan van het ‘zwaarste’ doel bij de beoordeling of bij een bepaalde vorm van onderzoek versneld vervanging van dierproeven gezocht moet worden. Dat wil zeggen, vormen van onderzoek met een ‘zwaarder’ doel, zullen minder hoog op de prioriteringsagenda gezet worden (mede afhankelijk natuurlijk van hoe ze scoren op andere prioriteringscriteria).

Als laatste moet in het oog gehouden worden dat het hier gaat om *prioriteringscriteria* bij de zoektocht naar vervanging van dierproeven, niet om bijvoorbeeld de wenselijkheid van een direct verbod. Dat maakt de minder precieze vaststelling van het belang van verschillende categorieën van doelen van onderzoek een relatief klein probleem.

⁴⁰ Zulke condities zijn soort- en rassespecifiek: voor een melkgeit bijvoorbeeld is huisvesting een must in de winter, terwijl dat voor veel schapenrassen niet zo is.

⁴¹ Hierbij maken we verder geen onderscheid tussen levensstijlgerelateerde ziekten en andere ziekten of aandoeningen, omdat het verwerven van een ziekte of aandoening altijd mede bepaald wordt door de omgeving. Bovendien zullen er altijd meer en minder verwijtbare groepen mensen kunnen profiteren van de eventuele uitkomsten van proefdieronderzoek.



Samenvatting

Proefdieronderzoek wordt uitgevoerd met een doel. Het directe doel van een concrete dierproef, bijvoorbeeld het ontrafelen van een mechanisme, kan worden onderscheiden van het uiteindelijke doel, bijvoorbeeld een therapie voor Multiple Sclerosis, dat verder in de toekomst ligt en waarvoor meerdere experimenten en onderzoeken nodig zijn. Sommige van die uiteindelijke doelen staan ter discussie. Zo is onderzoek ten behoeve van cosmetica in Europa verboden en velen vragen zich af of het nog nodig is om nog veiligheidstesten van chemicaliën op dieren te doen.

Met behulp van VandeVeer's onderscheid in basale, serieuze en perifere belangen kunnen de verschillende uiteindelijke doelen van onderzoek geordend worden. Een onderzoek met een basaal doel is bijvoorbeeld gericht op de preventie, genezing en het draagbaar maken van levensbedreigende ziekten en ziekten die ernstig belemmerend zijn voor de kwaliteit van leven van mens en dier en het behoud van diersoorten en ecosystemen. Bij serieuze doelen gaat het bijvoorbeeld om onderzoek gericht op duurzame vormen van veehouderij en veiligheidstesten van voedingsmiddelen. In de huidige moderne samenleving zouden bijvoorbeeld productiedoelen in de veehouderij perifeer genoemd kunnen worden.

Als doelen van onderzoek kunnen al deze doelen legitiem zijn, maar op het moment dat proefdieren nodig zijn om het doel te bereiken, kan een hiërarchie worden aangebracht. Voor projecten en diermodellen die primair perifere doelen dienen, zouden met voorrang alternatieven moeten worden ontwikkeld.

Voorbeeld relevantie doel dierproef

SPEENDIARREE BIJ BIGGEN

Al meer dan twee decennia wordt onderzoek gedaan naar de aanpak van gezondheidsproblemen van jonge biggen. Het gaat dan om biggen die in de periode kort na het weghalen bij de zeug diarree krijgen. Door een complex van factoren zijn biggen in deze periode na het spenen gevoeliger voor infecties, vaak met diarree als gevolg. Vooral sinds het bekend worden van de keerzijde van antibioticagebruik in de veehouderij, is veel onderzoek gericht op het bestrijden van speendiarree, zoals de inzet van verbeterde voeding en aanpassing van management- en huisvestingssystemen (zie voor enkele reviews: J. M. Campbell, Crenshaw, en Polo 2013; Rhouma e.a. 2017; Wensley e.a. 2021).

Hoewel de onderzoeken al geleid hebben tot diverse aanpassingen in de praktijk, blijft speendiarree een groot probleem voor de varkenshouderij. Het roept de vragen op: in hoeverre hebben de dierproeven tot dusver bijgedragen aan een oplossing van het probleem; betreft de onderzoekaanvraag voor het effect van het zoveelste voedings-supplement op speendiarree niet enkel symptoombestrijding; zijn er geen fundamentele alternatieven voorhanden, bijvoorbeeld de biggen later spenen of in lagere dichtheden houden? Er kan ook gekozen worden voor duurzame oplossingen, die echter op termijn een verandering in de gehele sector vragen.

Hoe moet dit type onderzoek beoordeeld worden in het licht van het criterium van de relevantie van het doel van een dierproef? In het licht van het onderscheid tussen basale, serieuze en perifere belangen, kan gesteld worden dat dit type onderzoek naar speendiarree bij biggen een perifeer belang heeft: het betreft onderzoek dat vooral een economisch belang dient en maar in hele beperkte mate de gezondheid van de biggen. Bovendien zijn er alternatieven beschikbaar. Het vervangen van dergelijk proefdieronderzoek heeft dan ook hoge prioriteit.

Het voorbeeld is met toestemming in gemodificeerde vorm overgenomen uit: RDA (2018b).



Ongerief van proefdieren⁴²

Proefdieronderzoek gaat gepaard met ongerief, ofwel aantasting van het welzijn van proefdieren. Gegeven de definitie van een dierproef in de Wod is zelfs inherent minimaal licht ongerief betrokken bij een dierproef.⁴³ Binnen de Europese Unie is het ongerief dat proefdieren ondergaan de voornaamste vorm van schade binnen de voorgeschreven schaden-baten-analyse. Daarnaast geldt er Europees een bovengrens aan toegestaan ongerief⁴⁴:

Onverminderd de vrijwaringsclausule in artikel 55, lid 3, dragen de lidstaten er zorg voor dat een procedure niet wordt uitgevoerd indien deze leidt tot een ernstige mate van pijn, lijden of angst die waarschijnlijk langdurig zal zijn en niet kan worden verzacht (Europees Parlement en de Raad van de Europese Unie 2010, Artikel 15.2).⁴⁵

De genoemde vrijwaringsclausule bepaalt dat voor buitengewone en wetenschappelijk te rechtvaardigen redenen, langdurig ernstig ongerief toch toegestaan kan worden door individuele lidstaten.

De Europese richtlijn 2010/63/EU schrijft ook een classificatie van de ernst van procedures voor in termen van terminaal, licht, matig, en ernstig ongerief. De ernst van procedures wordt hierbij “bepaald aan de hand van mate van pijn, lijden, angst of blijvende schade die een individueel dier tijdens de procedure naar verwachting zal ondervinden” (Europees Parlement en de Raad van de Europese Unie 2010, Bijlage VIII). Hoewel de Europese richtlijn pijn, lijden en angst in een adem noemt, betogen Olsson et al. (2019) dat er een kwalitatief onderscheid is tussen verschillende gradaties van pijn (en ziekte) enerzijds, en (angst en) lijden anderzijds. De mate van pijn is afhankelijk van duur en intensiteit, maar de

cruciale vraag is of het dier in staat is tot *coping*. Op het moment dat de pijn overweldigend wordt voor het dier “*compensation cannot occur, normal life cannot be experienced, and the animal cannot fully recover and will be fundamentally changed even if the external situation improves*” (Olsson e.a. 2019, sec. How to Measure Severe Suffering). Mogelijk dat daarom de Europese richtlijn ook een aparte clausule opgenomen heeft over langdurig ernstig ongerief, naast de voorgeschreven schaden-baten-analyse.

In het kader van het stellen van prioriteiten in het vervangen van dierproeven, staan theoretisch gesproken dierproeven die ernstig en langdurig ongerief, ofwel lijden, met zich meebrengen hoog op de agenda. Het geval wil echter dat dierproeven met ernstig en langdurig ongerief niet uitgevoerd worden binnen de Europese Unie. Als een lidstaat een dierproef met ernstig en langdurig ongerief uit wil laten voeren, moet dit binnen een maand gerapporteerd worden aan de Europese Commissie. Olsson et al. (2019) vermelden echter dat op 1 juli 2019 nog geen enkele notificatie van een dierproef met ernstig en langdurig ongerief binnengekomen was bij de Europese Commissie. Hoewel dit reden tot voorzichtig optimisme kan zijn, moet wel in oegenschouw genomen worden dat de inschatting van de betrokken mate van ongerief ter verantwoordelijkheid van afzonderlijke commissies komt, en dat hier slechts algemene regels voor zijn (Olsson e.a. 2019). Er kan dus variatie zijn in hoe regels geïnterpreteerd en toegepast worden. Bijgevolg kan de afwezigheid van ernstig en langdurig ongerief louter een administratieve waarheid zijn. Sowieso zou nader onderzoek moeten worden of in recente jaren wel notificaties binnengekomen zijn bij de Europese Commissie.

Na ernstig en langdurig onderzoek, betreft proefdieronderzoek dat ernstig ongerief met zich meebrengt een logische focus als prioriteit voor vervanging. In het Verenigd Koninkrijk betrof 4% van de dierproeven ernstig ongerief in 2020 (Home Office 2021). In Nederland was dit in hetzelfde jaar 0,9% (NVWA 2022). Volgens de laatste Europese cijfers was in 2016 en 2017 echter bij 11% van de dierproeven ernstig ongerief betrokken (European Commission 2020).

Wat hier verder nog van belang is, is dat Europees gezien het zogenaamde *batch potency testing* verreweg de grootste categorie met ernstig ongerief vormt, gevolgd op afstand door onderzoek naar het zenuwstelsel en diagnostische testen (European Commission 2020).

⁴² Deze sectie is kort gehouden vanwege een parallel NCad onderzoek naar ernstig ongerief en de bekendheid in het werkveld met de ongeriefinschatting.

⁴³ Voor de definitie, zie de Inleiding.

⁴⁴ Recent hebben ook DeGrazia en Beauchamp (2020) een raamwerk van zes criteria gelanceerd ter bepaling van de morele aanvaardbaarheid van proefdieronderzoek, waarvan één een bovenlimiet aan schade betreft.

⁴⁵ Zie ook *Wet op de dierproeven* (2021, §3, Artikel 10b, lid 2).



In het kader van het stellen van prioriteiten in het vervangen van dierproeven, kan dus de schaal die binnen de Europese Unie gehanteerd wordt om ongerief te 'meten' gebruikt worden. De prioritering verloopt daarbij van ernstig en langdurig, tot ernstig, tot matig, tot mild ongerief.

Samenvatting

Proefdieronderzoek gaat gepaard met ongerief, ofwel aantasting van het welzijn van proefdieren. De Europese richtlijn 2010/63/EU schrijft een classificatie van de ernst van procedures voor in termen van minimaal, licht, matig, en ernstig ongerief. Deze classificatie van ongerief kan gebruikt worden als prioriteringsschema in het vervangen van dierproeven, waarbij gestart wordt met het stimuleren van alternatieven voor dierproeven en diermodellen die (langdurig) ernstig ongerief met zich mee brengen.

Voorbeeld ongerief

ROUTINEMATIGE BATCH-TESTEN CLOSTRIDIUMVACCINS

In Europa vormt batch potency testing verreweg de grootste categorie dierproeven met ernstig ongerief. De Europese Farmacopee schrijft voor dat iedere geproduceerde batch van een vaccin getoetst moet worden op werkzaamheid. Er is weinig bekend over de specifieke vaccins waarbij ernstig ongerief optreedt. Een studie van de RSPCA uit 2008 levert de belangrijkste, hoewel gedateerde, informatie (J. Cooper en Jennings 2008).

Eén van de voorbeelden waar de grootste zorgen leven met betrekking tot het ongerief van proefdieren, betreft de toetsing van de werkzaamheid van batches veterinaire vaccins (J. Cooper en Jennings 2008; Cruelty Free Europe 2022; Hendriksen, persoonlijke communicatie, 9 februari, 2022). Eén voorbeeld daarbinnen zijn de Clostridiumvaccins. In het geval van vaccins tegen boutvuur (*Clostridium chauvoei*) in schapen en runderen, en botulisme (*Clostridium botulinum*) in schapen, runderen, paarden en vogels, is de Farmacopee weinig directief en laat de mogelijkheid open blootstellingstesten uit te voeren (J. Cooper en Jennings 2008). Het zijn met name deze blootstellingstesten – naast toxine-neutraliserende testen – die voor veel ongerief zorgen bij routinematige batch-testen. Muizen (*C. botulinum*) en hamsters (*C. chauvoei*) worden hierbij verdeeld in gevaccineerde en ongevaccineerde groepen, en vervolgens blootgesteld aan de betreffende ziekteverwekker. Het betrokken ongerief voor de proefdieren neemt hierbij ernstige vormen aan. In het geval van *C. chauvoei* gaat het hierbij om oedeemvorming en necrose van weefsel, en uiteindelijk de dood van de hamsters in de ongevaccineerde groep (K. P. Singh, Parihar, en Tripathi 1992, geciteerd in J. Cooper en Jennings 2008). In het geval van *C. botulinum* gaat het om uitdroging, verlies van sensomotorische reflexen en hart- of ademhalingsfalen van muizen in de ongevaccineerde groep (Luvisetto e.a. 2003, geciteerd in Cooper en Jennings 2008).

Vanuit het oogpunt van ongerief heeft het vervangen van routinematige batch-testen van Clostridiumvaccins dus hoge prioriteit. Belangrijk daarbij is dat vervangende in vitro testen in ontwikkeling zijn voor zowel *C. botulinum* (Rust e.a. 2017) als *C. chauvoei* (Nicholson e.a. 2019).



Integriteitsaantasting van de proefdieren

Achtergrond van het begrip integriteit

Meerdere auteurs hebben betoogd dat het dierethische debat gebaat zou zijn bij een expliciete bespreking van de aantasting van de integriteit van proefdieren in onderzoeksprojecten (De Vries 2004; Röcklinsberg, Gamborg, en Gjerris 2014; Verhoog 2004). Ook de CCD heeft de integriteit van proefdieren opgenomen als evaluatiepunt in het adviesformulier dat DEC's aanleveren bij de CCD (vraag C12) (zie ook CCD 2019).⁴⁶ Hoewel dus reeds toegepast op de proefdierpraktijk, is het begrip 'dierlijke integriteit' in eerste instantie opgekomen in de Nederlandse discussie in de jaren '90 rondom ingrepen als het onthoornen van vee, het couperen van staarten van rashonden, en genetische modificatie van dieren (Brom 1997; De Vries 2009). Het begrip integriteitsaantasting duidt daarbij op moreel problematische zaken 'boven' of 'achter' welzijnsaantastingen (Bovenkerk, Brom, en Van Den Bergh 2002; Röcklinsberg, Gamborg, en Gjerris 2014; De Vries 2006). Ook al is een ingreep als het couperen van een staart van een hond vrijwel pijnloos wanneer gedaan onder anesthesie en volgens goed diergeneeskundig vakmanschap, toch ervaren veel mensen een dergelijke ingreep als moreel problematisch. De begrippen integriteit en integriteitsaantasting bieden handvatten om dit bespreekbaar te maken.

De meeste studies verwijzen in hun bespreking van integriteit naar de definitie die Rutgers en Heeger daarvan gaven in 1999 (Bovenkerk, Brom, en Van Den Bergh 2002; Gjerris en Gamborg 2010; Grimm 2014; Nordgren 2010; Ortiz 2004; Schmidt 2008; Verhoog 2007)⁴⁷ Deze klassiek geworden definitie van integriteit van dieren luidt: "[the] wholeness and completeness of the animal and the species-specific balance of the creature, as well as the animal's capacity to sustain itself independently in an environment suitable to the species." (Rutgers en Heeger 1999, 45).⁴⁸ Drie noties spelen een centrale rol in deze definitie: (1) de heelheid van

het dier; (2) de soortspecifieke balans van het dier; en (3) de capaciteit om zich onafhankelijk te handhaven in een voor die soort gepaste omgeving. Als voorbeeld van de aantasting van de heelheid van het dier noemen Rutgers en Heeger het onthoornen van vee. Een voorbeeld dat Rutgers en Heeger geven met betrekking tot de aantasting van de soortspecifieke balans van het dier is het fokken op snelle spieraanzet in vleeskuikens, waardoor deze dieren op een gegeven ogenblik niet meer goed kunnen lopen en zelfs spontaan sterven. Bij deze vleeskuikens is de stofwisseling uit balans. Een voorbeeld, als laatste, van de aantasting van de capaciteit om zich onafhankelijk te handhaven in een voor die soort gepaste omgeving, is het fokken van dikbilkoeien als het Belgisch Witblauw (Rutgers en Heeger 1999, 47–48). Deze dieren hebben een buitengewone sterke spiergroei in combinatie met een smal geboortekanaal. Deze lichamelijke configuratie wijkt, als gevolg van domesticatie en selectie, af van het rund in een meer natuurlijke context. Hierdoor kunnen kalveren van dikbilrassen niet op een *natuurlijke* of *soorteigen* wijze geboren worden: een keizersnede is altijd vereist. Deze dikbilkoeien zijn dus altijd aangewezen op intensieve hulp van de kant van de mens.

De voorbeelden die Rutgers en Heeger zelf aangedragen hebben als illustratie van integriteitsaantastingen hebben allemaal betrekking op landbouwhuisdieren en gezelschapsdieren. Zonder op dit moment al direct in te stemmen met de definitie van Rutgers en Heeger, zijn er bij wijze van illustratie ook voorbeelden te geven van wat telt als integriteitsaantastingen bij *proefdieren* volgens Rutgers en Heeger's definitie. Een aantasting van de heelheid van het dier, als eerste, is bijvoorbeeld de teen- of oorknip bij genetisch gemodificeerde muizen ter identificatie en genetische karakterisatie (vgl. NCad 2020). Het meest spraakmakende 'voorbeeld', als tweede, van verstoring van de soortspecifieke balans van proefdieren is de muis met een menselijk oor op de rug (vgl. Trouw 1995). Een subtieler geval betreft het voorkomen van hoornvorming bij runderen middels CRISPR-technologie. Een voorbeeld, als laatste, van een aantasting van de capaciteit om zich onafhankelijk te handhaven in een voor de soort gepaste omgeving zijn de naakte muizen die veel gebruikt worden in oncologisch onderzoek. Vanwege de afwezigheid van een thymus en een vacht vereisen deze muizen voor een goed biologisch en mentaal functioneren een microbiologisch en klimatologisch sterk gecontroleerde laboratoriumomgeving, wat hun afhankelijkheid van de mens zeer groot maakt.

⁴⁶ In Denemarken werd het begrip integriteit in 2006 gebruikt binnen de wetgeving met betrekking tot het kloneren van dieren (Gjerris en Sandøe 2006). Of het momenteel nog steeds gehanteerd wordt is ons onbekend.

⁴⁷ Hoewel het begrip 'dierlijke integriteit' sterke wortels in de Nederlandse discussie heeft, wordt het al in 1995 ook gehanteerd in een document van een Zwitserse ethische commissie, zonder enige referentie naar Nederlandse auteurs (Ethik-Studienkommission des Eidgenössischen Volkswirtschaftsdepartements zur Gentechnologie im ausserhumanen Bereich 1995). Dit zou erop kunnen duiden dat 'dierlijke integriteit' als begrip niet louter een Nederlandse oorsprong heeft.

⁴⁸ De RDA volgt min of meer deze opvatting van integriteit: "Integriteit van het dier omvat de soortspecifieke heelheid en compleetheid van een dier en zijn vermogen om naar zijn aard onafhankelijk te functioneren in een bij de soort passende omgeving" (RDA 2018a, 25).



Samengevat gaat integriteit volgens de klassiek geworden definitie van Rutgers en Heeger over de heelheid van het dier, de soortspecifieke balans van het dier, en het vermogen om zich onafhankelijk te handhaven in een voor de soort gepaste omgeving.

Receptie van Rutgers en Heeger's definitie van integriteit

Met betrekking tot de receptie van de definitie van integriteit zoals door Rutgers en Heeger gegeven, zijn er drie zaken om op te merken. De eerste betreft de interpretatie van het eerste subcriterium, de heelheid van het dier. Het tweede betreft een verschuiving in de opvatting van integriteit in de richting van een focus op capaciteiten of vermogens van dieren. Het derde betreft een kritische kanttekening bij het derde subcriterium van onafhankelijk functioneren in een voor de soort gepaste omgeving. Deze drie punten zullen achtereenvolgens worden besproken.

Heelheid van het dier

De vraag is wat precies verstaan moet worden onder de heelheid van het dier. Is het onthoornen van vee problematisch omdat daarmee direct geïntervenieerd wordt in het dierlijk lichaam of is het problematisch omdat het rund dan een lichaamsdeel moet missen wat het van nature wel had?⁴⁹ Brom benoemt dat heelheid gaat over de "heelheid (gaafheid) van het dier in fysieke en fysiologische zin" (Brom e.a. 1996, 15). Hij vat dit echter breed op blijkens zijn voorbeeld van een naakte muis die een vacht moet missen die het van nature wel heeft (zie ook 'Voorbeeld integriteit: naakte muizen').⁵⁰ Dit sluit aan bij de opvatting dat het onthoornen van vee problematisch is omdat ze iets moeten missen dat ze van nature, als soort, fysiek wel hebben.

⁴⁹ Een andere optie is nog: of is het problematisch omdat het rund bepaalde vermogens moet missen als gevolg van het onthoornen? Dit komt ter sprake onder 'Dierlijke vermogens'.

⁵⁰ Het 'van nature' verwijst naar noties als 'natuurlijkheid' en 'telos' (Hauskeller 2005; Kramer en Meijboom 2021). Er liggen hier allerlei interessante verbanden tussen de begrippen van integriteit, natuurlijkheid en telos. Zo stelt Hauskeller bijvoorbeeld dat "biological integrity consists in the ability to live according to one's natural ends [or] telos" (Hauskeller 2016, 47). Dit vraagt om een nadere verkenning, die echter te ver voort voor dit rapport.

Schmidt (2008, 319) stelt echter dat het bij de operationalisering van de heelheid van het dier allereerst gaat om *overschrijding van de grenzen van het dierlijk lichaam*, zoals in het geval van een amputatie.⁵¹ Dat sluit aan bij de opvatting dat het couperen van staarten problematisch is omdat daarmee direct geïntervenieerd wordt in het dierlijk lichaam. Het fokken van een naakte muis vormt dan bijvoorbeeld geen aantasting van dierlijke integriteit, omdat niet in letterlijke zin lichamelijke grenzen overschreden worden (zie ook 'Integriteit van individu, foklijn of ras, en soort').

Er zijn dus minimaal twee interpretaties mogelijk van de heelheid van het dier. Hier wordt op terug gekomen in de paragraaf "Smalle en brede opvatting van integriteit".

Dierlijke vermogens

Om zich in een bepaalde omgeving fysiologisch en ethologisch in evenwicht te houden zonder intensieve hulp van de mens" (Brom e.a. 1996, 15). De eerste interpretatie "betreft vooral ethologische en functionele factoren in het beeld. De referentie is hier naar een 'ethogram' waarin alle meer of minder essentiële soort-specifieke gedragsuitingen van een dier worden beschreven" (Brom e.a. 1996, 15).⁵² Dieren zouden de mogelijkheid moeten hebben om soorteigen gedrag te vertonen. Varkens zouden in de modder moeten kunnen wroeten, kippen kunnen scharrelen, en knaagdieren tunnels kunnen graven. In de recente wijziging van de Wet Dieren heeft dit aspect van de mogelijkheid tot uitoefening van soorteigen gedrag ook meer gewicht gekregen (Tweede Kamer der Staten Generaal 2021). In de tweede interpretatie is de referentie niet naar de soort, maar naar de mate van onafhankelijk en succesvol *coping* van het dier, en ligt daarmee dicht bij bepaalde opvattingen van dierenwelzijn (bijvoorbeeld Broom 1993; Korte, Olivier, en Koolhaas 2007).

⁵¹ Bezien vanuit het wettelijk kader kan de vraag gesteld worden wat 'overschrijding van de grenzen van het lichaam' toevoegt aan het bestaande wettelijke kader dat immers spreekt van een dierproef wanneer sprake is van "experimentele of andere" en "onderwijskundige doeleinden, die bij het dier evenveel of meer pijn, lijden, angst of blijvende schade kan veroorzaken als, dan wel dan het inbrengen van een naald volgens goed diergeneeskundig vakmanschap" (Wet op de dierproeven 2021, §1, Artikel 1a). Dekt de huidige definitie van wat een dierproef behelst voor de wet, niet al iedere aantasting van de heelheid van het dier? Inderdaad is het zo dat de heelheid van het dier, opgevat in smalle zin, per definitie wordt aangetast als er sprake is van een dierproef. Dit zegt echter nog niets over *de mate* van aantasting van heelheid/integriteit (zie "Integriteit: absoluut of relatief?"). Verder is ook de relatie tussen de wettelijke definitie van een dierproef en een opvatting van heelheid in brede zin anders.

⁵² Vgl. "Wanneer in een dier essentiële eigenschappen van de diersoort zijn verdwenen, is de integriteit van dat dier aangetast" (RDA 2016, 37).



In lijn met deze focus op *vermogens* van dieren hebben Heeger en Brom in een wat latere fase een andere definitie gegeven van integriteit. De definitie luidt: “*the physical and mental ability of the animal to realise their well-being and to flourish*” (Heeger en Brom 2001, 245). In het algemeen kan gesteld worden dat vermogens “datgene [zijn] waartoe een mens (of niet-menselijk dier) in staat is of onder de juiste omstandigheden zou kunnen zijn” (Nussbaum 2006, 119).⁵³ Dierlijke vermogens gaan hier dus over dat wat een dier in staat stelt om zijn welzijn en floreren te realiseren. Heeger en Brom komen hier heel dicht bij definities van welzijn in termen van *coping*, zoals juist besproken. Dat roept dan wel de vraag op in hoeverre integriteit daadwerkelijk nog iets zegt ‘boven’ of ‘achter’ dierenwelzijn.⁵⁴

De Vries (2009, 88) wijst erop dat Heeger en Brom’s definitie van integriteit niet helemaal samenvalt met de klassieke definitie: het onthoornen van vee, bijvoorbeeld, is problematisch volgens de klassieke opvatting van integriteit omdat de heilheid van het dier dan aangetast wordt. Volgens de definitie van Heeger en Brom is een dergelijke interventie echter alleen maar problematisch als het leidt tot aantasting van het vermogen van het vee om diens welzijn en floreren te realiseren.⁵⁵ Aantasting van de (lichamelijke) heilheid van

⁵³ Hoewel de discussie over integriteit in termen van vermogens – voor zover we weten – geen expliciete verwijzing naar de vermogensbenadering van Nussbaum kent, toont zij formeel overeenkomsten. Nussbaum vertaalt tien belangrijke menselijke vermogens naar de context van dieren, te weten het vermogen tot leven, gezondheid, lichamelijke integriteit, het hebben van plezier, het ervaren van emoties door betrokkenheid bij anderen, praktisch denken (voor zover als vermogen aanwezig), het aangaan van sociale banden, relaties met andere diersoorten, spel, en het hebben van controle over de omgeving (Nussbaum 2006; Janssens 2020). Opvallend is dat Nussbaum ‘lichamelijke integriteit’ noemt als vermogen. Ze lijkt dit te bedoelen in de zin van ‘het vermogen tot het handhaven van lichamelijke integriteit’. Toch lijkt lichamelijke integriteit nog fundamenteler te zijn, in die zin dat het een voorwaarde is voor het überhaupt hebben van vermogens, zoals in het debat over dierlijke integriteit naar voren komt (vgl. Schmidt 2008).

⁵⁴ Zoals gezegd benoemen Heeger en Brom expliciet de relatie tot dierenwelzijn wanneer ze stellen dat integriteit bestaat in ‘*the physical and mental ability of the animal to realise their well-being and to flourish*’. Aangezien Heeger en Brom respect voor integriteit als een aparte plicht zien naast zorg voor een bevredigend leven voor dieren, impliceert dit dat zij integriteitsaantastingen niet reduceren tot welzijnsaantastingen.

⁵⁵ De Vries claimt dat het onthoornen van runderen geen negatieve invloed heeft op de mogelijkheden van runderen om een goed leven te leiden en te floreren, en dat daarom onthoornen geen aantasting van de integriteit van het dier betekent. Dit is aanvechtbaar. Welzijnsaantastingen in de vorm van pijn op de plaats waar de hoorns weggebrand zijn kunnen bijvoorbeeld langdurig zijn, wat een inbreuk is op een goed leven voor een rund.

het dier is, volgens de nieuwere definitie, niet altijd een vorm van aantasting van de integriteit van het dier. In de nieuwere definitie wordt ‘heilheid’ dus niet absoluut gehanteerd, maar wordt dat wat het dier in potentie ‘kan’ met zijn genoom of een bepaald lichaamsdeel meegewogen bij het bepalen of iets een integriteitsaantasting is.⁵⁶

Onafhankelijk functioneren

Zoals gezegd behelst het derde subcriterium van Rutgers en Heeger’s definitie de capaciteit om zich onafhankelijk te handhaven in een voor die soort gepaste omgeving. Schmidt werkt dit theoretisch op originele wijze uit door te stellen dat integriteit “niet alleen een *toestand* – waarin ze zich kunnen bevinden – betekent, maar... ook een specifiek *integratievermogen behelst*” (Schmidt 2008, 318, vertaling CJR). Het dier is zelf *actief* betrokken bij het in stand houden van de samenwerking tussen de verschillende delen van zijn lichaam,⁵⁷ en de inpassing van zijn lichaam als geheel in de omgeving (*Umwelt*) (Schmidt 2008). Een concreet voorbeeld is het bouwen van nesten door vogels:

Het bouwen van een nest bestaat uit een patroon van verschillende gedragingen, dat op zijn beurt onderdeel uitmaakt van een groter gedragspatroon, waar onder andere ook het sociaal gedrag tussen mannetjes- en vrouwtjesvogels, het broedgedrag en hun zorggedrag voor de jongen deel van uitmaken. Dit is op zijn beurt weer onderdeel van het gedragspatroon van de betreffende vogelsoort. De eenheid van de betreffende vogel bestaat dus, naast de soort-specifieke eenheid van organen, uit zijn of haar vermogen tot organisatie van het eigen gedrag (Lijmbach 1996, 68, nadruk toegevoegd).

⁵⁶ De Vries kiest uiteindelijk voor de nieuwere definitie, omdat die volgens hem (onder andere) recht kan doen aan vrijwel alle morele intuïties en oordelen die we hebben als het gaat om integriteitskwesies bij dieren.

⁵⁷ Verhoog (2007, 369) merkt op dat “*integrity presupposes the existence of an ‘organism’, a living whole with interconnected parts. It is the interconnectedness, the balanced harmony of the parts of the whole, which is somehow linked to the concept of integrity*”. Ook anderen merken op dat integriteit ‘heilheid’ impliceert, maar daarmee ook een integratie van ‘delen’ binnen een bepaalde ‘geheel’ (Holdrege 2002; geciteerd in Hauskeller, 2016; Schmidt 2008; Nordgren 2010). Juist wanneer de verschillende delen in een gebalanceerde verhouding ten opzichte van elkaar zijn, spreken we over heilheid. Zo bezien vormen de eerste twee subcriteria uit de klassieke definitie twee kanten van dezelfde medaille: de ‘heilheid’ uit het eerste subcriterium veronderstelt de ‘balans’ uit het tweede subcriterium. De dood van een dier is daarmee de ultieme aantasting van diens integriteit omdat “*the whole is completely broken into pieces and ceases to exist*” (Nordgren 2010, 70).



Hoewel op een eerste niveau integriteitaantasting dus duidt op een overschrijding van de min of meer statistische lichamelijke grenzen van het dier (zie boven, 'Heelheid van het dier'), duidt het op een tweede niveau op het actief in stand houden door het dier van de grenzen van zijn lichaam in relatie tot zijn omgeving. Het is deze 'actieve' component waarnaar verwezen wordt in het derde subcriterium in de definitie van Rutgers en Heeger.

Toch kan de vraag gesteld worden of Rutgers en Heeger's definitie niet verder gaat dat deze actieve component, doordat ze stelt dat dieren zich *onafhankelijk* moeten kunnen handhaven. Dit lijkt te absoluut gesteld. Ten eerste maakt dit het begrip integriteit moeilijk hanteerbaar in relatie tot gedomesticeerde dieren die per definitie niet meer in hun 'oorspronkelijke', natuurlijke habitat leven (Brom 1997, 139). Brom (1997, 139) stelt dan ook als meer hanteerbaar criterium voor: "de mate waarin een dier hulp nodig heeft van de zijde van de mens in het bevredigen van zijn behoeften en het bijvoorbeeld op 'onnatuurlijke' wijze moet worden gevoed en gehouden".

Ten tweede heeft de zorgethiek in relatie tot de *menselijke* situatie laten zien dat afhankelijkheid en kwetsbaarheid – al dan niet wederzijds – kenmerkend zijn voor het menselijk bestaan en niet noodzakelijkerwijs problematisch (bijvoorbeeld Comstock 1992; Kittay 2011; Engster 2019). Evenzo kan de vraag gesteld worden waarom het vereist is dat dieren onafhankelijk moeten kunnen functioneren; afhankelijkheidsrelaties kunnen ook daar een legitieme plaats innemen, afhankelijk van de situatie (Engster 2006). Wel kan gesteld worden dat er moreel gesproken een verschil is tussen het actief afnemen van bepaalde vermogens van mens en dier en daarmee het afhankelijker maken van deze mensen en dieren, en het eenvoudigweg bestaan in en erkennen van (wederzijdse) afhankelijkheidsrelaties. In het eerste geval is het gevaar van instrumentalisatie en een aantasting van de *waardigheid* van mens en dier aanwezig (Balzer, Rippe, en Schaber 2000; Ortiz 2004; Nussbaum 2006; Hauskeller 2007). Het tweede geval betreft een gegeven werkelijkheid.

Smalle en brede opvatting van heelheid

In relatie tot het subcriterium van 'heelheid' werd al vermeld dat er minimaal twee interpretaties mogelijk zijn van de heelheid van het dier. In de eerste interpretatie wordt heelheid, en daarmee integriteit, in *smalle zin* opgevat als overschrijding van de lichamelijke grenzen van het dier. Hierbij telt bijvoorbeeld een teenknip – zonder verdere gevolgen voor de locomotie en ander gedrag van de muis – als een aantasting van dierlijke integriteit, maar niet de fok van naakte muizenlijnen die sterk beperkt zijn in hun

aanpassingsvermogen aan de microbiële omgeving en het klimaat.⁵⁸ In het laatste geval is namelijk geen sprake van een directe lichamelijke interventie.

Heelheid kan echter ook geïnterpreteerd worden binnen een *bredere* opvatting van integriteit, waarbij referentie gemaakt wordt naar enerzijds de vermogens van dieren en anderzijds de soort waartoe het behoort. Kort samengevat is een integriteitsaantasting dan *iedere interventie die leidt tot een verlies aan soortspecifieke vermogens*. Integriteit bestaat daarmee in *het bezit van het set van vermogens die kenmerkend voor de soort is*.⁵⁹ Het gaat hierbij expliciet om *de set*, ofwel combinatie, van vermogens die kenmerkend zijn voor een soort. Sommigen van deze vermogens zullen breed gedeeld zijn, maar anderen kunnen exclusief zijn voor een soort. Een min of meer soortspecifiek vermogen van knaagdieren is bijvoorbeeld het knagen aan relatief harde (organische) voorwerpen, terwijl het vermogen tot voortplanten een breed gedeeld vermogen is binnen het dieren- en plantenrijk. Lammerts van Bueren et al. (2003) spreken in dat laatste geval van de 'integriteit van het leven'.

Omdat het vermogens zijn, vraagt het activiteit van het dier om die vermogens te *realiseren*.⁶⁰ Vermogens 'vragen' dus om een geëigende context waarbinnen ze gerealiseerd kunnen worden. Leghennen in een kooihuisvesting ontberen bijvoorbeeld de juiste context voor het realiseren van stofbaden als soortspecifiek gedrag. Evenzo ontnemt het individueel huisvesten van mannelijke muizen hen de mogelijkheid om elkaars vacht te poetsen als soortspecifieke activiteit.⁶¹ Toch is het goed om op te merken dat het op deze

⁵⁸ Een lastig geval betreft het afknippen van een stukje van de vin van een vis, de zogenaamde vinknip (Berger Eriksen e.a. 2011). Hierbij wordt de heelheid van het lichaam aangetast, maar niet op permanente basis, want de vin van een vis groeit weer aan.

⁵⁹ Het onderscheid dat de RDA (2016) in haar zienswijze op fokkerij en voortplantingstechnieken maakt tussen indicatoren voor soorteigen gedrag en indicatoren voor integriteit is dus niet erg zinnig. Inderdaad weet zij geen concrete indicatoren te benoemen voor integriteit. Het voorbeeld dat zij noemt van kippen gefokt op blindheid kan evengoed benaderd worden vanuit de mogelijkheid tot uitoefening van soorteigen gedrag (scharrelen, etc.).

⁶⁰ Nussbaum (2006) noemt dit *functionings*.

⁶¹ Stereotypieën kunnen in die lijn worden gezien als onnatuurlijk gedrag en in zekere zin als het spiegelbeeld van natuurlijk of soorteigen gedrag: het stangbijten van zeugen wordt bijvoorbeeld gezien als omgericht wroet- en nestbouwgedrag. Stereotypieën zijn herhaalde, vaste gedragspatronen die gekenmerkt worden door hun (ogenschijnlijke) doelloosheid (Broom 1991). Onderzoek naar het voorkomen van stereotypieën en het bevorderen van soorteigen gedrag middels kooiverrijking bij proefdieren als muizen is nog beperkt maar gaande (Bailoo e.a. 2018; Gross e.a. 2012; Mason e.a. 2007; Novak e.a. 2016).



wijze niet voorzien in een geëigende omgeving voor (proef)dieren niet een aantasting van de integriteit van deze dieren betreft. De dieren behouden hun vermogens, alleen wordt niet voorzien in de mogelijkheid tot realisering van deze vermogens. Dit is wel een punt van zorg⁶², maar geen zaak van integriteitsaantasting.

Belangrijk om nu op te merken is dat deze brede opvatting van integriteit recht kan doen aan het morele onbehagen dat opgeroepen wordt door procedures als de fok van naakte muizenlijnen en het gebruik van gevasectomeerde mannelijke muizen. Deze praktijken zijn niet noodzakelijkerwijs problematisch volgens een smalle opvatting van integriteit.⁶³ Wel vereist een operationalisering van een brede opvatting van integriteit veel kennis van soortspecifieke vermogens van dieren en in hoeverre deze aangetast kunnen worden (meer over operationalisering onder 'Operationalisering van integriteit').

Hier staat tegenover dat de smalle opvatting van integriteit het gemakkelijkst te operationaliseren is, en aansluit bij hoe lichamelijke integriteit van mensen vastgelegd is in de Nederlandse Grondwet en gehanteerd wordt in de humane medische ethiek.

Empirische bruikbaarheid van het begrip integriteit

In de wetenschappelijke en ethische literatuur is er veel discussie (geweest) over de vraag of integriteit een empirisch bruikbaar begrip is. De vraag is of integriteit een *ideaal* beschrijft of (ook) een relatie tot de *empirie* heeft. In de bespreking hiervan zal gestart worden met het weergeven van de standpunten binnen het debat, en geleidelijk overgegaan worden tot een meer kritische bespreking van deze standpunten en een voorstel gedaan worden voor een constructieve opvatting van integriteit.

Een reden om te ontkennen dat integriteit een relatie tot de empirie heeft is ingegeven door de claim dat een ingreep bij dieren om veterinaire redenen – zoals een keizersnede bij een schaap in geval het lam niet in een goede positie ligt – geen aantasting van de integriteit van het dier behelst, in tegenstelling tot een ingreep bij dieren voor gebruiksdoeleinden – zoals een keizersnede bij een schaap om de effecten van vroeggeboorte te

⁶² Dit punt wordt binnen bredere opvattingen van dierenwelzijn geadresseerd als 'natural living' (Fraser e.a. 1997b; Fraser 2008; V. Lund 2006; Green en Mellor 2011).

⁶³ Tegelijkertijd moet gezegd worden dat dergelijke procedures ook problematisch gevonden kunnen worden op andere gronden, bijvoorbeeld op basis van de notie dat de mens ingevoegd is in een gegeven orde in de werkelijkheid die hij alleen tot zijn schade kan manipuleren, en daarom bescheidenheid moet oefenen en hoogmoed vermijden (Brom 1997, 152–56; D. E. Cooper 1998; Scruton 1998).

simuleren (Vorstenbosch 1993; Rutgers en Heeger 1999). In het eerste geval is de ingreep primair in het belang van de gezondheid van het schaap en het lam zelf. In het tweede geval is de ingreep primair in het belang van de mens: het schaap met lam dient als model voor vroeggeboorte in de mens. Volgens Vorstenbosch en Rutgers en Heeger speelt de *intentie* van de handeling dus mee als we oordelen dat een bepaalde ingreep de integriteit van dieren aantast of niet.⁶⁴

Een probleem met deze opvatting is echter dat het resultaat van twee handelingen met verschillende intenties lichamelijk hetzelfde resultaat kan geven. Het couperen van de staart van een hond voor esthetische redenen geeft hetzelfde lichamelijke resultaat – namelijk een hond zonder staart – als het couperen van de staart om veterinaire redenen (Bovenkerk, Brom, en Van Den Bergh 2002). Volgens de opvatting van Vorstenbosch en Rutgers en Heeger zou het couperen om esthetische redenen de integriteit van de hond wel aantasten, maar het couperen om veterinaire redenen niet. Maar zou hieruit dan niet de conclusie getrokken moeten worden dat

Integrity is not a biological aspect of the animal itself after all. The concept then loses its objective, biological character and becomes a moral rather than an empirical notion. It does not refer to a notion of factual intactness or wholeness so much as to a perceived intactness. It refers to how we feel an animal should be (Bovenkerk, Brom, en Van Den Bergh 2002; nadruk in origineel)?⁶⁵

⁶⁴ Een andere, meer perifere, reden waarom aangevoerd wordt dat integriteit niet een empirisch begrip is, vindt zijn oorsprong in de opmerking van Vorstenbosch (1993) dat de integriteit van een dier alleen geschaad kan worden door een *menselijke* interventie. Zoals aangegeven door De Vries wil Vorstenbosch daarmee waarschijnlijk niet ontkennen dat de (lichamelijke) heelheid van een dier gewijzigd kan worden door biotische oorzaken, zoals bijvoorbeeld het optreden van kanker of door een predator, maar dat we alleen van een aantasting van de integriteit van dieren spreken als de heelheid van het dier aangetast wordt door de mens. Het is echter niet onmiddellijk evident waarom er niet gesproken kan worden van aantasting van de integriteit van een muis wanneer een poot afgerukt wordt door een kat.

⁶⁵ Overigens waren Rutgers en Heeger zelf van mening dat integriteit wel degelijk ook objectief bepaald kan worden omdat het refereert aan "*biological traits that are characteristic of the species in question*" (Rutgers en Heeger 1999, 49). Het blijft gissen hoe Rutgers en Heeger zelf dit objectieve karakter van integriteit wilden handhaven in combinatie met de intentie van de handeling.



In diezelfde lijn wordt door de CBD opgemerkt dat

De mate waarin de biologische kenmerken van een dier worden aangetast of aangepast, kan worden beschouwd als een indicator van de mate van aantasting van de integriteit van het dier. In filosofische termen heet het dan dat het begrip integriteit-aantasting is geoperationaliseerd door te kijken naar de mate waarin bepaalde biologische kenmerken, zoals uiterlijk, gedrag en zelfredzaamheid veranderen. Beide zaken kunnen echter niet aan elkaar gelijkgesteld worden: integriteit heeft een waarderende betekenis, terwijl de biologische kenmerken van een dier in objectieve termen kunnen worden beschreven (CBD 2008, 22).

Hoewel niet geëxpliciteerd, wordt in deze lijn integriteit *intersubjectief* geacht. In tegenstelling tot objectief waarneembaar uiterlijk, gedrag, etc. is integriteit een bepaalde, breed gedragen en door dialoog gevestigde waardering door mensen van die objectief waarneembare biotische fenomenen.

Deze gedachtelijn kan om meerdere redenen bekritiseerd worden. Laten we beginnen bij de koppeling van intenties aan het oordeel of een bepaalde dierkundige handeling de integriteit van een dier aantast. De Vries (2009, 89–90) betoogt, onzes inziens terecht, dat de aantasting van de integriteit van de dieren *niet* toegeschreven hoeft te worden aan de intenties waarmee gehandeld wordt.⁶⁶ Ook een ingreep voor veterinaire doeleinden kan gezien worden als aantasting van de integriteit van het dier. De (lichamelijke) heilheid van het dier wordt uiteindelijk toch aangetast en/of verliest bepaalde vermogens: een dier moet een poot missen, wordt levenslang onvruchtbaar, etc. Een dergelijke ingreep kan echter wel gerechtvaardigd zijn, omdat het principe van weldoen voorschrijft dat de gezondheid en welzijn van het dier bevorderd moet worden. Het weldoen weegt in dit

⁶⁶ Iets soortgelijks kan gezegd worden van Bovenkerk, Brom en Van den Bergh's (2002, 18–19) voorbeeld met betrekking tot menselijke integriteit. Zij zijn huiverig om de conclusie te trekken dat een transgender persoon zijn eigen lichamelijke integriteit aantast als deze zou besluiten om middels medische weg van geslacht te veranderen. De reden hiervoor is dat Bovenkerk et al. menen dat wanneer een ingreep door de persoon zelf gewild is, er geen sprake kan zijn van aantasting van integriteit. Ook hier zou gesteld kunnen worden dat er wel degelijk sprake kan zijn van een aantasting van lichamelijke integriteit als een dergelijke ingreep uitgevoerd wordt, maar dat dit afgewogen moet worden tegen het (mogelijk) meer in harmonie komen met zichzelf van de transgender persoon. Zo gesteld hoeft de conclusie dan ook niet te zijn dat "integrity as a moral notion can therefore be diametrically opposed to integrity as an empirical notion" (Bovenkerk, Brom, en Van Den Bergh 2002, 19): het begrip integriteit heeft zowel een empirische als evaluatieve kant.

geval zwaarder dan het respect hebben voor de integriteit van het dier. Er is dus een moreel verschil tussen de twee handelingen – een ingreep voor veterinaire doeleinden en een handeling voor gebruiksdoeleinden – zonder dat deze bepaald wordt door de intentie van de handeling.

Hier kan nog aan toegevoegd worden dat het koppelen van intenties aan het begrip integriteit, de empirische bruikbaarheid van dit begrip niet ten goede komt. Het voorbeeld van de keizersnede om veterinaire reden kan dit duidelijk maken. Keizersnedes zijn in het belang van de gezondheid van schaaap en lam, maar ook in het belang van de boer: de boer heeft belang bij levende en gezonde dieren. Een handeling voor veterinaire doeleinden kan dus ingebed zijn in een gebruikerspraktijk. Als intenties onderdeel zouden zijn van onze beoordeling of, en in hoeverre, de integriteit van een dier aangetast is, dan is het in het geval van een keizersnede op een schapenboerderij niet duidelijk of de integriteit van de schapen inderdaad aangetast is of niet.⁶⁷

Een tweede punt van kritiek met betrekking tot deze gedachtelijn heeft betrekking op de opmerking van de CBD dat integriteit een (intersubjectief-)waarderende betekenis zou hebben terwijl biologische concepten objectief(-beschrijvend) zouden zijn.^{68,69} Niemand van ons zal echter betwijfelen of gezondheid (en ziekte) een biologisch begrip is. Toch is het een begrip dat niet alleen beschrijvend is – 'deze resusaap is gezond' – maar ook waarderend, of liever: evaluatief. Gezondheid wordt daarom ook wel een '*thick concept*' genoemd (Haverkamp, Bovenkerk, en Verweij 2018). Gezondheid geeft een bepaalde norm aan; als daarvan afgeweken wordt spreken we dan ook over ziekte. Over waarin die norm dan bestaat lopen de meningen uiteen, getuige ook de vele definities van gezondheid, maar dat er een norm ligt, is duidelijk.

⁶⁷ Vorstenbosch schrijft daarom waarschijnlijk dat "it is unusual - at least - to say that the integrity of the animal is violated when human (veterinarian) action, for instance surgery, is taken that is intended to benefit the animal - and only the animal." (Vorstenbosch 1993, 111; nadruk toegevoegd). Alleen als een handeling helemaal ten dienste van het dier zou zijn en niet ook ten dienste van bijvoorbeeld de boer of onderzoeker, zou er geen sprake zijn van aantasting van integriteit van het dier. Dit maakt de empirische bruikbaarheid van het concept echter minimaal, want in de praktijk zijn doeleinden en belangen van mens en dier vaak met elkaar vervlochten.

⁶⁸ Een ander punt van kritiek, van minder belang echter, is dat de CBD 'zelfredzaamheid' noemt als voorbeeld van een biologisch kenmerk. Zelfredzaamheid heeft echter ook een sterk waarderende betekenis, zodat het als voorbeeld van wat het CBD beoogt, niet erg geschikt is.

⁶⁹ Het is mogelijk dat begrippenparen als intrinsieke en extrinsieke waarde, en finale en instrumentele waarde, primair (inter)subjectief van aard zijn, hoewel dit ook bestreden wordt (bijvoorbeeld Van den Heuvel, Nullens, en Roothaan 2017). Hoe dit ook zij, integriteit moet wel onderscheiden worden van een begrip als intrinsieke waarde; dit gebeurt niet altijd (zie voor deze gelijkshakeling bijvoorbeeld: COGEM 2018).



Belangrijk in dit verband is dat (dieren)welzijn ook een *thick concept* is: “Any conception of animal welfare inherently involves values because it pertains to what is better or worse for animals” (Fraser e.a. 1997a, 188, nadruk in origineel; vgl. Fraser 1995; Tannenbaum 1991). Dierenwelzijn gaat niet alleen over de staat waar het dier momenteel in is – bijvoorbeeld of het vrij is van honger en dorst, een comfortabel onderkomen heeft, gezond is, de mogelijkheid heeft tot het uitoefenen van soorteigen gedrag, en vrij is van angst⁷⁰ – maar veronderstelt ook dat bijvoorbeeld de (chronische) aanwezigheid van honger en dorst ongerief voor het dier betekent, en daarmee problematisch en onwenselijk is.

Dit laat zien dat biologische concepten niet louter objectief-beschrijvend zijn, zoals gesteld door de CBD. De stelling die hier geponereerd wordt is dat integriteit een *thick concept* is net zoals gezondheid en welzijn (vgl. Visser en Verhoog 1999, iii). Integriteit heeft zowel een empirische ofwel beschrijvende, als een evaluatieve kant. Net als bij de begrippen gezondheid en welzijn is er daarbij discussie mogelijk over de vraag waarin de norm voor integriteit bestaat (zie bijvoorbeeld CBG 2002). Dit is echter geen reden om het begrip terzijde te schuiven, net zo min als bij de begrippen gezondheid en welzijn.

Gegeven het feit dat het bijdragen aan het welzijn van proefdieren breed gedragen en geoperationaliseerd is – denk aan de 3V's van vervanging, verfijning en verfijning (bijvoorbeeld Fenwick, Griffin, en Gauthier 2009) – kan beargumenteerd worden dat er bij analogie ook geen fundamenteel bezwaar hoeft te zijn tegen het gebruik van het begrip van integriteit als prioriteringsbeginsel omdat het evaluatief van aard zou zijn. Natuurlijk kan er wel op andere gronden bezwaar gemaakt worden tegen het gebruik van de mate van integriteitsaantasting als prioriteringsbeginsel. Een bezwaar is bijvoorbeeld dat integriteit een te absoluut concept is en geen gradatie toestaat. Op dit bezwaar zal in de paragraaf ‘Hoe kan integriteit geoperationaliseerd worden?’ ingegaan worden.

De conclusie van deze paragraaf is dat integriteit wel degelijk ook een empirische component heeft. Het is mogelijk om deze conclusie te handhaven als gesteld wordt dat de integriteitsaantasting niet bepaald wordt door de intentie van de handeling, maar een relatie heeft tot de gevolgen voor het dier. Tegelijkertijd heeft integriteit als begrip een evaluerende functie, net als de begrippen gezondheid en (dieren)welzijn. Er is dan ook geen reden om de integriteit van het dier uit te sluiten als prioriteringsbeginsel in het vervangen van dierproeven omdat het evaluatief van aard zou zijn.

⁷⁰ De befaamde vijf vrijheden van dierenwelzijn zoals geformuleerd in 1965 door een Britse commissie (Rogers Brambell e.a. 1965).

Operationalisering van integriteit

Integriteit: absoluut of relatief?

Verwant aan de discussie over de empirische bruikbaarheid van het begrip integriteit, heeft in de literatuur de vraag gespeeld of integriteit een absoluut of relatief begrip is. Vorstenbosch suggereert dat integriteit een soort ja/nee-begrip is, net zoals het concept van maagd zijn of niet (Vorstenbosch 1993). Evenzo betogen Bovenkerk, Brom en Van den Bergh (2002) dat integriteit een absoluut concept lijkt te zijn, net als zwanger zijn of niet. Integriteit is aanwezig, of niet. Gradaties zijn niet mogelijk.

Anderen hebben echter betoogd dat wel degelijk gesproken kan worden van gradaties in integriteitsaantasting, en daarmee van een relatief begrip van integriteit. Volgens Rutgers en Heeger (1999) kan gesproken worden van meer of minder ernstige aantastingen van integriteit al naar gelang (1) de heilheid van het dier; (2) de soortspecifieke balans van het dier; en (3) de capaciteit om zich onafhankelijk te handhaven in een voor die soort gepaste omgeving, aangetast wordt. Hierbij is niet helemaal duidelijk of de mate van aantasting van integriteit nu bepaald wordt door *hoeveel criteria* geschonden worden (waarbij ieder criterium in ja/nee-termen opgevat wordt), of dat bij ieder van deze drie criteria een mate van aantasting bepaald kan worden, wat vervolgens een cumulatieve mate van aantasting van integriteit geeft.

Ook De Vries meent dat het mogelijk is “to objectively assess the seriousness of the violation of integrity” (De Vries 2009, 91). De ernst van de aantasting is volgens De Vries een functie van de mate waarin de vermogens van het dier om zijn welzijn te realiseren gereduceerd zijn. Hij stelt daarbij dat het verwijderen van een poot weleens een minder ernstige aantasting van de integriteit van een hond kan zijn dan het couperen van de staart, omdat de staart een belangrijke sociale functie heeft.



Belangrijk om op te merken is echter dat zowel Rutgers en Heeger als De Vries spreken over de mate van *aantasting* van integriteit (zie ook Brom e.a. 1996). Dit is echter nog niet hetzelfde als gradaties in integriteit. Bovenkerk, Brom en Van den Bergh (2002) stellen namelijk óók dat over gradaties in *aantasting* van integriteit gesproken kan worden, maar – zoals gezegd – niet over gradaties in integriteit.⁷¹

De vraag kan evenwel gesteld worden hoe zinnig het is om de aanleiding of oorzaak van de integriteitsaantasting en de gevolgen daarvan voor het dier op deze manier uit elkaar te trekken. Een ‘schending’ of ‘aantasting’ is altijd een schending of aantasting van iets, in dit geval de integriteit *van het dier*. De daad valt dus niet los te zien van de gevolgen voor het dier. Wil men van ‘meer en minder’ spreken in relatie tot de schending of aantasting van integriteit van dieren, dan zal men toch moeten kijken naar de empirische gevolgen voor de dieren.⁷²

Inderdaad heeft Brom eerder gesteld dat “gradatie mogelijk is omdat het mogelijk is om een reeds in zijn integriteit aangetast dier, opnieuw in integriteit aan te tasten. De integriteit van een hond waarvan de staart gecoupeerd is, is aangetast, maar kan opnieuw aangetast worden door het couperen van de oren” (Brom 1997, 131–32). In die zin verschilt het van maagd of zwanger zijn: wie eenmaal geen maagd meer is, kan niet meer ontmaagd worden. En wie eenmaal zwanger is, kan op dat moment niet nog eens zwanger worden.

⁷¹ Helaas werken Bovenkerk, Brom en Van den Bergh dit niet verder uit, maar gesteld kan worden dat ze hiermee de gevolgen van de handeling buiten beschouwing laten en slechts focussen op de handeling zelf. Dit wordt in de ethiek wel een deontologische vorm van moreel redeneren genoemd, in tegenstelling tot bijvoorbeeld een gevolgen-ethische vorm van redeneren. Dit past in hun koppeling van integriteit met de intentie van de handeling. Toch is dat weer niet geheel consistent met het voorbeeld dat Bovenkerk, Brom en Van den Bergh (2002) noemen – het verwijderen van de poot van een hond versus het couperen van de staart – waarbij zij menen dat de eerste een grotere schending van integriteit behelst dan de laatste. Gezien vanuit de intentie van de handeling, zou namelijk het couperen van de staart een grotere schending van integriteit behelzen omdat deze ingreep doorgaans gepleegd wordt om esthetische doeleinden. Dit in tegenstelling tot het verwijderen van een poot dat doorgaans gedaan wordt om veterinaire redenen.

⁷² Vgl. “To say of someone that he has violated the integrity of an animal does not so much say that someone has the wrong attitude towards animals, but first and foremost that he has performed an intervention in the life of an animal with morally problematic consequences” (De Vries 2006, 48 1). En evenzo: “In this sense, the concept of animal or phenotypical integrity may be called a consequentialist concept, because arguments that are couched in its terms base the (*prima facie*) rightness or wrongness of an action, in this case a genetic modification, on the consequences of the action for the resulting animal” (De Vries 2009, 97).

Samenvattend kan gesteld worden dat integriteit ook als relatief begrip opgevat kan worden en dat dus gradaties in integriteitsstatus mogelijk zijn.⁷³ Hoe deze gradatie aangebracht kan worden zal in de paragraaf ‘Integriteit als criterium in de praktijk’ op ingegaan worden.

Integriteit van individu, foklijn of ras en soort

Eerder werd al benoemd dat dierlijke integriteit een rol is gaan spelen in de discussie over genetische modificatie omdat het handvatten bood om het morele onbehagen dat mensen hierbij ervoeren ter sprake te brengen (Macnaghten 2004). De vraag die hier rijst is in hoeverre het nu mogelijk is om over de integriteit van een foklijn, ras, of soort te spreken. Deze vraag zal behandeld worden aan de hand van een voorbeeld.

Het Green Fluorescent Protein (GFP) gen afkomstig uit de kwal *Aequorea victoria* en het luciferase-gen afkomstig uit de vuurvlieg *Photinus pyralis* worden veel gebruikt in proefdieronderzoek. Door middel van het inbrengen van deze genen in bijvoorbeeld muizen of zebravissen kan men de activiteit van fysiologische processen waar men in geïnteresseerd is, monitoren. Bij blootstelling aan licht (GFP) of luciferine zullen namelijk de betreffende weefsels oplichten.

Betreft het inbrengen van het dergelijke genen in bijvoorbeeld een zebravislijn een integriteitsaantasting? Wat in de beantwoording van deze vraag als eerste vastgesteld moet worden is om de integriteit *van wie of wat* het gaat. Gaat het om de integriteit van de individuele vissen van deze zebravislijn? Of gaat het om de integriteit van de zebravislijn? Laten we beginnen met de individuele vissen. Geconstateerd moet worden dat genetische

⁷³ Verscheidene auteurs hebben betoogd dat integriteit de norm ‘respect voor integriteit’ ‘oproept’ (Vorstenbosch 1993; Brom 1997, 135; Hauskeller 2016). En inderdaad, als integriteit als absoluut begrip gehanteerd wordt (zo Bovenkerk, Brom, en Van Den Bergh 2002), dan kan geen enkele menselijke handeling daar iets aan toedoen, alleen afdoen (Hauskeller 2016, 29–30). De enige gepaste respons voor de mens lijkt dan respect voor integriteit te zijn. Als echter integriteit opgevat wordt als iets dat in meer of mindere mate aanwezig kan zijn, dan hoeft niet alleen respect een moreel gepaste respons te zijn. Ook *weldoen* kan een gepaste respons zijn, namelijk wanneer de integriteit van het dier al aangetast is. Weldoen aan het dier behelst dan het (deels) herstellen van de integriteit van het dier. Bijvoorbeeld, van een hond die een poot mist door wat voor oorzaak dan ook, en een prothese aangemeten krijgt, kan gezegd worden dat het een stukje integriteit herwint.



modificatie plaatsvindt in hele vroege embryonale fases. In deze fases is nog geen sprake van een lichaam in de zin van een verdeling in en samenhang tussen organen. Er is dus geen sprake van een overschrijding van lichamelijke grenzen in strikte zin (de smalle interpretatie van integriteit). Wel echter kan gesproken worden over een aantasting van integriteit in brede zin *wanneer het inbrengen van dergelijke genen leidt tot een verlies aan bepaalde vermogens op latere leeftijd*. De referentie voor wat telt als een verlies aan vermogens is de set van vermogens die het dier gehad zou hebben als het dier genetisch niet gemodificeerd zou worden. In het geval van de zebra vis met GFP en/of luciferase-gen is het echter niet goed voor te stellen welke vermogens verloren gaan wanneer deze gehouden wordt onder laboratoriumomstandigheden.⁷⁴ Dit is het probleem dat gevolgen op het fenotypische niveau van te voren niet altijd gemakkelijk te voorspellen zijn (CBD 2002).

Er treedt nu echter een moeilijkheid op. Er kan namelijk onderscheiden worden naar de embryo's waarin geïntervenieerd wordt en toekomstige nakomelingen van deze genetisch gemodificeerde embryo's. Zojuist ging het over de integriteitsaantasting van de embryo's waarin geïntervenieerd werd. Maar hoe staat het met mogelijke toekomstige nakomelingen? Is daarvan ook de integriteit aangetast?

Er zijn hier minstens drie opties. De eerste is om het wild-type dier met zijn vermogens als referentie te nemen (zo Verhoog 1992; 2003; 2007). In dat geval zijn ook de nakomelingen van genetisch gemodificeerde zebra visen in hun integriteit aangetast. Het probleem met deze optie is dat niet (altijd) duidelijk is wat 'het' wild-type dier is – aangezien veel proefdieren al decennia tot millennia gedomesticeerd zijn – en waarom dat de relevante set van vermogens bezit.

De tweede optie is om de foklijn of het ras zoals het tot dusver ontwikkeld is als referentie te nemen. In dat geval 'tellen' de effecten in de huidige generatie van biotechnische handelingen die hebben plaatsgevonden in vorige generaties niet als integriteitsaantastingen van de huidige generatie. Wel zou weer gelden dat een genetische modificatie, met negatieve gevolgen voor de set van vermogens, in de huidige generatie een integriteitsaantasting betekent van de dieren in de huidige generatie. Ons is geen auteur uit de

⁷⁴ Dit zou anders *kunnen* zijn voor zebra visen die vervolgens in het wild terecht komen. Het zou zo kunnen zijn dat dergelijke zebra visen problemen hebben met voortplanten omdat het licht dat ze verspreiden andere zebra visen afschrikt. In dat geval is er wel sprake van een aantasting van integriteit van de zebra visen.

literatuur bekend die deze optie verdedigt. Een probleem met deze optie is in ieder geval dat het referentiepunt arbitrair is.

De derde optie is om te kijken naar de algehele lichaamsstructuur en het gedrag van de dieren. Dit is met name relevant wanneer vermogens weggenomen worden, zoals soms in het geval van een gen knock-out. De hele lichaamsstructuur inclusief fysiologie kan dan nog verwijzen naar een bepaalde functie, die echter gefrustreerd wordt door de gen knock-out. Hoewel Hauskeller (2005) en Kramer en Meijboom (2021) niet met het begrip integriteit werken maar met de notie van *telos*, is dit wel de strekking van hun betoog.

Zowel wanneer het wild-type dier als uitgangspunt genomen wordt, als wanneer de algehele lichaamsstructuur en het gedrag als uitgangspunt genomen wordt, kan – met een variant op de bekende uitdrukking 'fok met ongerief' – gesproken worden van 'fok met integriteitsaantasting'.

Tot zover hebben we gekeken naar de integriteit van individuele leden van een soort. Kan er echter ook gesproken worden over de integriteit van een foklijn, ras of soort? Zowel op basis van de smalle en de brede interpretatie van integriteit ligt het niet voor de hand om te stellen dat een foklijn, ras of soort een integriteit zou kunnen hebben. In beide gevallen is integriteit namelijk gebaseerd op lichamelijkeheid; de brede interpretatie van integriteit gaat alleen verder door te stellen dat deze lichamelijkeheid voorziet in bepaalde vermogens bij dieren en dat moreel gesproken het probleem zit in het afnemen van deze vermogens. Foklijnen, rassen en soorten hebben echter geen lichaam, en bijgevolg kan moeilijk over de integriteit van een foklijn, ras of soort gesproken worden.⁷⁵

Samengevat kan gesteld worden dat bij sommige foklijnen sprake is van inherente integriteitsaantasting, ofwel fok met integriteitsaantasting. Als het wild-type dier als uitgangspunt genomen wordt, zal dit sneller het geval zijn dan wanneer de uiteindelijke lichaamsstructuur van de resulterende dieren als uitgangspunt genomen wordt. Vanwege de koppeling van dierlijke vermogens aan lichamelijkeheid, en de niet-lichamelijke van soorten – alleen individuen van de soort hebben een lichaam – kan beter niet gesproken worden over integriteitsaantastingen van foklijn, ras of soort.

⁷⁵ Rolston (2002) neemt een wat ingewikkelde positie in want hij stelt enerzijds dat soorten integriteit kunnen bezitten en hij ziet individuele leden van een soort als delen van de soort als geheel. Anderzijds accepteert hij wel dat "a species ... is not a bounded singular" (Rolston 2002, 7), wat toch suggereert dat soorten geen lichaam hebben. Het voert te ver om hier nader op in te gaan.



Integriteit als criterium in de praktijk

Het begrip integriteit is gedurende zo'n tien jaar gebruikt als toetsingscriterium binnen de Commissie Biotechnologie bij Dieren (CBD), die het begrip jarenlang hanteerde bij het beoordelen van dierproeven waarin genetische modificatie van dieren plaatsvond (Brom e.a. 1996). Hierbij maakte de CBD nader onderscheid tussen genotypische en fenotypische integriteitsaantasting bij de genetische modificatie van dieren. De CBD beschouwde "elke wijziging die wordt aangebracht in het genoom van het dier als een aantasting van de genotypische integriteit, maar [deed] vervolgens een uitspraak over de mate van de aantasting van de integriteit door te kijken naar de fenotypische gevolgen van de biotechnologische handelingen" (CBD 2002, 11). Dit betekende in de praktijk bijvoorbeeld dat "met een relatief groot construct een bepaalde genetische modificatie conditioneel gemaakt kan worden of zo kan worden aangebracht dat deze slechts in één bepaald orgaan of weefsel tot expressie komt. Vaak kan daarmee de aantasting van de integriteit op fenotypische niveau juist beperkt worden". Met andere woorden, het inbrengen van een groter stuk DNA telde niet perse als een grotere aantasting van integriteit; dat was afhankelijk van het uiteindelijke fenotype en derhalve van de aantasting van fenotypische integriteit. Bij dit laatste ging het dan vooral om aberraties in "biologische kenmerken, zoals uiterlijk, gedrag en zelfredzaamheid" (CBD 2008, 22), waarbij de norm voor gedrag (en uiterlijk) ingegeven wordt door de vraag of het al dan niet "soortegen" is (CBD 2008, 13).⁷⁶

Uit de ervaringen van de CBD blijkt dat het in de praktijk nog niet altijd zo eenvoudig is om uit te maken wanneer en in welke mate er sprake is van de aantasting van integriteit van een dier:

⁷⁶ Uit wat de CBD (2008) schrijft wordt niet geheel duidelijk of het 'soortegen' alleen betrokken werd op 'gedrag' of ook op 'uiterlijk'. Heeft 'soortegen' geen betrekking op 'uiterlijk', dan is de vraag wat voor norm er gehanteerd werd voor 'uiterlijk'. Heeft het wel ook betrekking op 'uiterlijk', dan rijzen er vragen met betrekking tot de consistentie in beoordeling van de CBD, daar ze van mening was dat de fenotypische integriteit van dieren – zoals muizen of zebrafissen – waarbij een genconstruct uit een kwal of vuurvlieg is ingebracht, niet of minimaal aangetast wordt. Er kan echter moeilijk volgehouden worden dat een zebrafis die, blootgesteld aan een bepaalde stof, oplicht als een vuurvliegje, een soortegen uiterlijk vertoont.

Een voortdurend terugkerende moeilijkheid voor de Commissie is dat de aantasting van gezondheid, welzijn en integriteit van de genetisch gemodificeerde dieren in veel gevallen niet goed te meten is en daarom wetenschappelijk niet voldoende te onderbouwen is. Zelfs als deze dieren ter beschikking zouden staan van onderzoekers, die hiertoe de expertise en faciliteiten hebben, dan nog zou het aan gevalideerde wetenschappelijke instrumenten ontbreken om dit te doen op een wijze waar deskundigen het in grote lijnen over eens zijn. In veel gevallen echter gaat het om dieren die niet voor dergelijk onderzoek beschikbaar zijn of zelfs om dieren waarvan het fenotype nog niet bekend is. De Commissie kan daarom vooralsnog niet anders doen dan met de beschikbare gegevens en de in de Commissie aanwezige expertise naar beste weten een schatting geven van de vermeende aantasting van welzijn en integriteit. Nochtans verdient het beschikbaar komen van valide meetinstrumenten dienaangaande de volle aandacht (CBD 2002, 13).

De CBD adresseert het punt dat 'valide meetinstrumenten' wenselijk zijn om de aantasting van integriteit in te schatten. Het is voorstelbaar dat analoog aan de ongeriefinschatting die er in Europa bestaat een inschatting en classificatie van de aantasting van integriteit gemaakt wordt door een samen te stellen werkgroep.

Een vraag die hierbij rijst is in hoeverre gradaties aangebracht kunnen worden in de ernst van de aantasting van integriteit. Als voorschot kan het volgende gesteld worden. Als integriteit volgens de smalle interpretatie opgevat wordt als lichamelijke heilheid, dan is het moeilijk in te zien hoe verschillende ingrepen – zoals een vasectomie, een teenknip of het nemen van een bloedmonster – kunnen verschillen in ernst. In alle geval vindt immers een overschrijding van lichamelijke grenzen plaats. Gradatie zal dan vooral plaats hebben in relatie tot de frequentie van de hoeveelheid ingrepen die de integriteit aantasten. Met een variant op Brom (1997): de integriteit van een gevasectomeerde mannelijke muis is aangetast, maar wordt verder aangetast door een teenknip.

Dit ligt anders voor de brede interpretatie van integriteit. Ook hier speelt de frequentie van ingrepen een rol. Maar daarnaast hangt de ernst van de aantasting van integriteit vooral samen met de mate waarin een specifiek dierlijk vermogen aangetast wordt, en het aantal



vermogens dat aangetast wordt. Bij een vasectomie bijvoorbeeld wordt het reproductieapparaat als stelsel aangetast in zijn integriteit, verliest daarmee zijn reproductieve functie en tast bijgevolg ook de integriteit van het dier als geheel aan. Het is echter denkbaar dat een ingreep gedaan wordt die de reproductieve functie niet volledig wegneemt, maar slechts de vruchtbaarheid vermindert. In dat geval vindt er dus een beperktere aantasting van integriteit plaats.

Het is denkbaar dat een classificatie van integriteitsaantastingen ontwikkeld wordt met gradaties van licht, matig en ernstig. Net als bij de ongeriefinschatting zal dit geen kwantitatieve exercitie zijn, maar een kwalitatieve inschatting waarbij ook de inbreng van ethologen gewenst is. Ernstige integriteitsaantastingen vormen daarbij een speerpunt in het vervangen van dierproeven.

Samenvatting

In het voorgaande is betoogd dat de mate van integriteitsaantasting van proefdieren gehanteerd kan worden als criterium in het prioriteren in het vervangen van dierproeven. De kern van integriteit wordt gevormd door de samenhang van delen binnen een bepaald geheel; in het geval van dieren de samenhang tussen lichaamsdelen en het dier als geheel, dat zichzelf op zijn beurt verhoudt tot het geheel van zijn omgeving. Een smalle en een brede interpretatie van integriteit zijn besproken, waarbij de eerste betrekking heeft op de overschrijding van lichamelijke grenzen van het dier en het laatste betrekking heeft op iedere interventie die leidt tot een verlies aan soortspecifieke vermogens.

Op de brede opvatting van integriteit rijst wel de vraag wat als referentiepunt wordt genomen in geval nakomelingen worden geboren van dieren die middels genetische modificatie in hun integriteit zijn aangetast. Zijn die nakomelingen ook in hun integriteit aangetast? Zowel wanneer het wild-type dier als uitgangspunt genomen wordt, als wanneer de algehele lichaamsstructuur en het gedrag van het actuele dier als uitgangspunt genomen wordt, kan er inderdaad sprake zijn van integriteitsaantasting bij de nakomelingen. Het gaat hier dan – met een variant op de bekende uitdrukking ‘fok met ongerief’ – om ‘fok met integriteitsaantasting’.

Met betrekking tot gradaties in integriteitsaantasting, kan ingeval van de smalle definitie (interventie in dierlijk lichaam) alleen onderscheid gemaakt worden naar het aantal interventies binnen één dier. Ook bij de brede definitie speelt de frequentie een rol, maar

verdere gradatie in integriteitsaantasting is mogelijk omdat meer of minder soortspecifieke vermogens in meerdere of mindere mate weggenomen kunnen worden.

De brede opvatting van integriteit doet waarschijnlijk het meest recht aan het morele onbehagen van mensen met bijvoorbeeld vachtloze muizen of gevasectomeerde mannelijke muizen. Operationalisering van het brede begrip van integriteit is echter niet eenvoudig omdat dit veel kennis vraagt van de soort-specifieke vermogens van dieren en of, en wanneer, deze aangetast worden. Het is wenselijk om een werkgroep te laten werken aan een classificatie in termen van licht, matig en ernstig gekoppeld aan de frequentie van ingrepen, de mate waarin een specifiek dierlijk vermogen aangetast wordt, en het aantal vermogens dat aangetast wordt. Diermodellen en dierproeven met ernstige vormen van integriteitsaantastingen vormen hierbij een speerpunt in het vervangen van dierproeven. (zie ook ‘Voorbeeld integriteit: naakte muizen’)



Voorbeeld integriteit

NAAKTE MUIZEN

Naakte muizen ontberen een thymus en een vacht. Ze zijn ontstaan door een natuurlijke mutatie. De afwezigheid van een thymus betekent dat de muizen slechts in beperkte mate T-cellen bezitten en daarmee immuundeficiënt zijn. De naakte muizen worden veel gebruikt binnen oncologische diermodellen (Boven en Winograd 2018).

Hoewel de naakte muizen niet ontstaan zijn als gevolg van bewust menselijk handelen, kan er weldegelijk gesproken worden van een integriteitsaantasting. De dieren zijn niet 'heel' in die zin dat ze een vacht en thymus ontberen en daarmee beperkt aangepast zijn aan wisselende omgevingscondities: de naakte muizen lopen een grotere kans om af te koelen en ziek te worden. Deze dieren kunnen zichzelf dan ook niet in stand houden, maar vereisen een microbiologisch en klimatologisch sterk gecontroleerde laboratoriumomgeving. Hierin is de mate van aantasting van de integriteit van naakte muizen bijvoorbeeld groter dan van muizen die een thymectomie ondergaan hebben. Bij de naakte muizen is er niet alleen een probleem met de thymus, maar ook missen ze een vacht en het bijbehorende relatieve onvermogen om zich aan te passen aan wisselende klimaatcondities.

Een concreet voorbeeld van wat een speerpunt kan vormen in het vervangen van dierproeven vormt de Red Fluorescent Protein (RFP) Nude Mouse. Het betreft hier een naakte muis waarin het gen voor Red Fluorescent Protein uit het neteldier *Discosoma striata* in de kiembaan is ingebracht. Vervolgens worden cellen uit een humane kankercellijn – waarin dan weer het gen voor Green Fluorescent Protein uit de kwal (*A. victoria*) is ingebracht – orthotopisch getransplanteerd naar deze muizen. Op deze manier kan de tumor micro-omgeving in kaart gebracht worden.

Volgens een smalle opvatting van integriteit is hierbij vooral de orthotopische implantatie een probleem. Dat is het ook voor de brede opvatting van integriteit, omdat voor de hand ligt dat bepaalde vermogens aangetast gaan worden door de ontwikkelende tumor, bijvoorbeeld met betrekking tot locomotie. Maar ook de naaktheid en het ontbreken van de thymus van de muis zijn dan een probleem (zoals zojuist geschetst). Of RFP en GFP interfereren met integriteit is lastiger te zeggen en vereist nader onderzoek.

Beoordelen middels een kader van criteria

De inzet van deze studie is het identificeren van criteria ter prioritering in het vervangen van dierproeven en diermodellen. Het is daarbij goed om op te merken dat discussie mogelijk is zowel op het niveau van de beoordeling van dierproeven aan de hand van de vier hoofdcriteria, als op het niveau van de empirische beoordeling hoe een bepaald type dierproeven scoort op een (sub)criterium. In de literatuur zijn tools te vinden die ontwikkeld zijn om deze laatste vorm van beoordeling te faciliteren (zie bijvoorbeeld de paragraaf 'Tools om de validiteit van diermodellen te bepalen' voor het geval van transleerbaarheid). Bij de keuze voor het gebruik van een dergelijke (kwantitatieve) tool moet bedacht worden dat ze wel een schijnobjectiviteit kunnen creëren, doordat gemaakte keuzes sedimenteren in het ontwikkelde scoresysteem (Grimm, Olsson, en Sandøe 2019).

Op basis van de verschillende criteria kan een oordeel geveld worden of een type dierproef hoog of juist lager op de hypothetische lijst van dierproeven die zo snel mogelijk vervangen zouden moeten worden, moet komen. Goed om op te merken is dat er hierbij enerzijds synergismen en anderzijds conflicten kunnen optreden tussen verschillende criteria. Een voorbeeld van het eerste is het werken aan semi-natuurlijke habitatten voor proefdieren uit het oogpunt van transleerbaarheid, waarbij ook het welzijn van proefdieren gebaat is. Een voorbeeld van juist een conflict tussen criteria is een thymectomie versus het gebruik van een naakte muizenlijn. Het gebruik van een naakte muizenlijn brengt een grotere aantasting van integriteit met zich mee omdat alle individuen van die foklijn het fenotype van het ontberen van een thymus zullen vertonen, in tegenstelling tot de enkele muizen die een thymectomie zullen ondergaan. Een thymectomie echter brengt weer inherent (matig) ongerief met zich mee. Per geval zal daarom bekeken moeten worden in hoeverre er sprake is van synergismen dan wel conflicten, en hoe daarmee om te gaan.

Conclusie

Het uiteindelijke doel van dit rapport is een eerste bijdrage te leveren voor het opstellen van criteria om dierproeven en diermodellen in beeld te krijgen waar met voorrang gewerkt zou moeten worden aan proefdiervrije alternatieven. Het is gebleken dat de mate van transleerbaarheid van het onderzoek, de mate van relevantie van het doel van een dierproef, de mate van ongerief dat de proefdieren berokkend wordt, en de mate van aantasting van integriteit voldoende houvast geven als criteria om te prioriteren in het vervangen van dierproeven.



Transleerbaarheid gaat over het bepalen van de reproduceerbaarheid en veralgemeniseerbaarheid van diermodellen die gebruikt worden in dierproeven. Zo wordt bijvoorbeeld in de wetenschappelijke literatuur discussie gevoerd over de vraag wat de veralgemeniseerbaarheid en daarmee transleerbaarheid van de Gedwongen Zwemtest in depressieonderzoek is: kan deze test echt iets zeggen over depressie in de mens of zegt hij louter iets over coping-strategieën van knaagdieren – en misschien mensen – in een stressvolle context? Dierproeven waarbij diermodellen gebruikt worden die een lage mate van transleerbaarheid hebben binnen een bepaald onderzoeksveld, komen daarbij als eerste in aanmerking om vervangen te worden, mede afhankelijk van de beoordeling op andere criteria, zoals de mate van relevantie van het doel van een dierproef.

Bij de relevantie van het doel van een dierproef spelen twee factoren een rol. Ten eerste is dit het type doel dat nagestreefd wordt met het onderzoek, dus onderzoek ten bate van cosmetica, voeding, gezondheid, chemicaliën, of anderszins. Ten tweede is dit de variant binnen een type onderzoek, dus bijvoorbeeld onderzoek ten bate van migraine of kanker binnen gezondheidsonderzoek. In het algemeen is een onderscheiding naar type doel van onderzoek hanteerbaar als criterium ter prioritering in het vervangen van dierproeven. Doelen van onderzoek zijn daarbij in te delen van basaal, naar serieus, naar perifeer, waarbij proefdieronderzoek dat perifere doelen dient als eerste in aanmerking komt ter vervanging. Een voorbeeld van dit laatste is onderzoek ten bate van veiligheidstesten voor chemicaliën ter gebruik in de plasticindustrie en onderzoek naar speendiarree in biggen.

Het derde criterium van de mate van aantasting van het ongerief van de proefdieren bouwt voort op bestaande wetgeving met betrekking tot de ongeriefinschatting en -beoordeling van proefdieronderzoek. Het ongerief van proefdieren gaat over pijn, lijden, angst of blijvende schade die een individueel dier als gevolg van de procedures naar verwachting zal ondervinden. De Europese, en in navolging daarvan de Nederlandse wetgeving, deelt de mate van ongerief in naar minimaal, mild, matig en ernstig. In het prioriteren van het vervangen van dierproeven komt proefdieronderzoek dat ernstig en langdurig is als eerste in aanmerking, gevolgd door proefdieronderzoek met ernstig ongerief. Een voorbeeld van dit laatste zijn de *batch potency*-testen waarbij de werkzaamheid van iedere batch vaccins getest moet worden.

Het criterium van de mate van aantasting van integriteit verwijst naar iedere lichamelijke interventie (smalle interpretatie van integriteit) of iedere interventie die leidt tot een verlies

aan soortspecifieke vermogens (brede interpretatie van integriteit). Dierproeven waarbij een ernstige mate van aantasting van integriteit betrokken is, komen daarbij als eerste in aanmerking om vervangen te worden. Of een interventie een lichte, matige of ernstige aantasting van integriteit met zich meebrengt zal vooral afhangen van de frequentie van integriteitsaantastingen in geval van de smalle interpretatie van integriteit, en daarnaast afhangen van de mate waarin dierlijk vermogens aangetast wordt en het aantal vermogens dat aangetast wordt in geval van de brede interpretatie van integriteit. Analooq aan wat gedaan is voor ongeriefinschatting, zou dit echter verder uitgewerkt moeten worden door bijvoorbeeld een werkgroep.

De gebruikte diersoort, mate van proportionaliteit van dierproeven, en de aantallen gebruikte dieren of procedures worden beoordeeld als onbruikbaar als criteria in het vervangen van dierproeven. De ontwikkelingsstatus van alternatieven voor dierproeven is strategisch wel van belang, maar niet relevant in het licht van het leveren van een extra inspanning bij het vervangen van dierproeven die ethisch gezien het meest problematisch zijn.

Concluderend kan gesteld worden dat de mate van transleerbaarheid van het onderzoek, de mate van relevantie van het doel van een dierproef, en de mate van ongerief en integriteitsaantasting dat de proefdieren berokkend wordt, kunnen dienen als kader voor het stellen van prioriteiten in het vervangen van dierproeven. Op basis van deze criteria is het denkbaar om te komen tot een prioriteitenlijst van typen dierproeven en diermodellen die zo snel als mogelijk vervangen moeten worden. Voorbeelden van typen dierproeven die slecht scoren op in ieder geval één criterium zijn de Gedwongen Zwemtest, onderzoek naar speendiarree bij varkens, batch-testen van Clostridiumvaccins, en het gebruik van genetisch gemodificeerde naakte muizen.

Aanbevelingen

Op basis van de interviews en de literatuurstudie wordt het volgende aanbevolen richting verscheidene stakeholders met betrekking tot dierproeven:

- Proefdierdeskundigen, ethologen, en wetenschappers met aanverwante expertise, vorm een werkgroep waarin een classificatie – bijvoorbeeld in termen van licht, matig en ernstig – van integriteitsaantasting ontwikkeld wordt, analoog aan wat bestaat voor de ongeriefinschatting. Belangrijk is om dit Europa-breed te doen om harmonisatie te bevorderen.



- Onderzoekers die werken met diermodellen (uit academie en industrie) en wetenschapsfilosofen, vorm een werkgroep waarin een classificatie – bijvoorbeeld in termen van laag, gemiddeld en hoog – van transleerbaarheid ontwikkeld wordt. Bestaande tools voor het bepalen van de transleerbaarheid kunnen hierbij een startpunt vormen. Ook hier is het belangrijk om dit Europa-breed te doen.
- Bestuurders en financiers van onderzoek, gebruik de vier uitgewerkte criteria van de mate van transleerbaarheid van het onderzoek, de mate van relevantie van het doel van een dierproef, de mate van ongerief dat de proefdieren berokkend wordt, en de mate van aantasting van integriteit als raamwerk om te sturen op proefdiervrije innovatie waar dit ethisch gezien het meest noodzakelijk is. Idealiter valt dit samen met het steunen van proefdiervrije alternatieven die al in een ver gevorderd stadium zijn (zie ook 'Ontwikkelingsstatus van alternatieven voor proefdieronderzoek'). Instrumenten kunnen zijn: financiering (bijvoorbeeld via NWO), publiek-private samenwerking en publieke communicatie (voor meer voorbeelden, zie Borrás en Edquist 2013).
- Adviseurs van dierproevenbeleid en dierenwelzijnsvertegenwoordigers, in samenspraak met wetenschapsbestuurders en onderzoekers, biotechnici en dierverzorgers werkzaam in specifieke onderzoeksvelden, identificeer op basis van de vier uitgewerkte criteria dierproeven en diermodellen waar dringend gezocht moet worden naar proefdiervrije alternatieven. Dat dit niet eenvoudig is, is wel gebleken bij de totstandkoming van dit rapport. Onderzoekers hebben vaak belang bij het gebruik van bepaalde diermodellen en dierproeven, en het publieke debat is sterk gepolariseerd. Bovendien worden dierproeven in hele verschillende contexten gebruikt, wat het lastig maakt om ergens te starten. Het belangrijkste lijkt daarom om te proberen om innovatie te bevorderen 'van binnenuit', dat wil zeggen vanuit de gebruikerspraktijk van dierproeven. De ontwikkeling van een 'prioriteringstool' kan daarbij behulpzaam zijn, omdat wetenschapsbestuurders en onderzoekers, biotechnici en dierverzorgers binnen een bepaald veld hier zelf mee aan de slag kunnen. Het zal daarbij belangrijk zijn om de waarde van een dergelijke tool Het zal daarbij belangrijk zijn om de waarde van een dergelijke tool voor het voetlicht te brengen; dit kan zowel financieel (dierproeven zijn niet goedkoop), strategisch-maatschappelijk ('*license to do science*'), wetenschappelijk (een slechte vertaalbaarheid betekent ondermaatse wetenschap) en ethisch (zorg voor dieren) zijn.



Bibliografie

Alliance for Human Relevant Science. 2020.

'Accelerating the Growth of Human Relevant Life Sciences in the United Kingdom: A White Paper by the Alliance for Human Relevant Science'.

https://www.humanrelevantscience.org/wp-content/uploads/Accelerating-the-Growth-of-Human-Relevant-Sciences-in-the-UK_2020-final.pdf

Animals (Scientific Procedures) Act 1986. 2013.

Animals Today. 2015.

'Labradors Universiteit Maastricht door oog van de naald'. 2015.

<https://www.animalstoday.nl/labradors-universiteit-maastricht-oog-naald/>

Bailoo, Jeremy D., Eimear Murphy, Maria Boada-Saña, Justin A. Varholick, Sara Hintze, Caroline Baussière, Kerstin C. Hahn, e.a. 2018. 'Effects of Cage Enrichment on Behavior, Welfare and Outcome Variability in Female Mice'. *Frontiers in Behavioral Neuroscience* 12.

<https://www.frontiersin.org/article/10.3389/fnbeh.2018.00232>

Balzer, Philipp, Klaus Peter Rippe, en Peter Schaber. 2000. 'Two Concepts of Dignity for Humans and Non-Human Organisms in the Context of Genetic Engineering'. *Journal of Agricultural and Environmental Ethics* 13 (1): 7–27.

<https://doi.org/10.1023/A:1009536230634>

Bauer, Thomas R., Yu-Chen Gu, Kate E. Creevy, Laura M. Tuschong, Lisa Embree, Steven M. Holland, Robert A. Sokolic, en Dennis D. Hickstein. 2004. 'Leukocyte Adhesion Deficiency in Children and Irish Setter Dogs'. *Pediatric Research* 55 (3): 363–67.

<https://doi.org/10.1203/01.PDR.0000111287.74989.1B>

Belzung, Catherine, en Maël Lemoine. 2011. 'Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression'. *Biology of Mood & Anxiety Disorders* 1 (1): 9.

<https://doi.org/10.1186/2045-5380-1-9>

Berger Eriksen, Tommy, Thomas Fraser, Håkon Gregersen, Mads Kristiansen, Marianne Olufsen, Binoy Rajan, Raul Ramirez, Jareeporn Ruangsri, Øyvind Røn, en Michel Guarjardo Sarmiento. 2011. 'Should fin clipping be used as a method for identification of fish?' Ås: Norecopa.

<https://norecopa.no/media/7290/fin-clipping-fish.pdf>

Bertzbach, Luca D., Daria Vladimirova, Kristina Dietert, Azza Abdelgawad, Achim D. Gruber, Nikolaus Osterrieder, en Jakob Trimpert. 2021. 'SARS-CoV-2 Infection of Chinese Hamsters (*Cricetus griseus*) Reproduces COVID-19 Pneumonia in a Well-Established Small Animal Model'. *Transboundary and Emerging Diseases* 68 (3): 1075–79.

<https://doi.org/10.1111/tbed.13837>

Borrás, Susana, en Charles Edquist. 2013. 'The choice of innovation policy instruments'. *Technological Forecasting and Social Change* 80 (8): 1513–22.

<https://doi.org/10.1016/j.techfore.2013.03.002>

Boven, Epie, en Benjamin Winograd. 2018. *The Nude Mouse in Oncology Research*. Boca Ration, FL: CRC Press.

Bovenkerk, Bernice, Frans W.A. Brom, en Babs J. Van Den Bergh. 2002. 'Brave New Birds: The Use of "Animal Integrity" in Animal Ethics'. *Hastings Center Report* 32 (1): 16–22.

<https://doi.org/10.2307/3528292>

Brom, Frans W.A. 1997. *Onherstelbaar verbeterd: Biotechnologie bij dieren als een moreel probleem*. Assen: Van Gorcum.

Brom, Frans W.A., M.T. Hilhorst, R.H.J. Ter Meulen, en J.M.G. Vorstenbosch. 1996. 'Het toetsen van biotechnologische handelingen bij dieren. Rapport van een commissie van externe deskundigen ten behoeve van de Commissie Biotechnologie bij Dieren'. Utrecht.

<http://docplayer.nl/18577912-Het-toetsen-van-biotechnologische-handelingen-bij-dieren.html>



Brønstad, Aurora, Christian E Newcomer, Thierry Decelle, Jeffrey I Everitt, Javier Guillen, en Kathy Laber. 2016. 'Current Concepts of Harm–Benefit Analysis of Animal Experiments – Report from the AALAS–FELASA Working Group on Harm–Benefit Analysis – Part 1'. *Laboratory Animals* 50 (1_suppl): 1–20.
<https://doi.org/10.1177/0023677216642398>

Broom, D M. 1991. 'Animal welfare: concepts and measurement'. *Journal of Animal Science* 69 (10): 4167–75.

Broom. 1993. 'A usable definition of animal welfare'. *Journal of Agricultural and Environmental Ethics* 6 (Supplement 2): 15–25.

Cameron, Daniel, Michael Clemence, en Kyra Xypolia. 2018. 'Public attitudes to animal research in 2018'. London: Ipsos Mori.

Campbell, Donald T., en Donald W. Fiske. 1959. 'CONVERGENT AND DISCRIMINANT VALIDATION BY THE MULTITRAIT–MULTIMETHOD MATRIX'. *Psychological Bulletin* 56 (2).

Campbell, Joy M., Joe D. Crenshaw, en Javier Polo. 2013. 'The biological stress of early weaned piglets'. *Journal of Animal Science and Biotechnology* 4 (1): 19.
<https://doi.org/10.1186/2049-1891-4-19>

CBD. 2002. 'Jaarverslag 2001'. Utrecht: Commissie Biotechnologie bij Dieren.

CBD. 2008. 'Jaarverslag 2007'. Utrecht: Commissie Biotechnologie bij Dieren.

CBG. 2002. 'Integriteit van dieren: Bouwsteen of Struikelblok? Verslag van de eerste discussiebijeenkomst over biotechnologie bij dieren. Gehouden op 21 november 2002 te Rotterdam'. Utrecht: Centrum voor Bio-ethiek en Gezondheidsrecht.
<https://edepot.wur.nl/118392>

CCD. 2019. 'Ethisch toetsingskader voor proefdiergebruik. Praktische handreiking voor Dierexperimentencommissies'. Centrale Commissie Dierproeven.
<https://www.centralecommissiedierproeven.nl/documenten/formulieren/16/6/6/ethisch-toetsingskader>

CCD. 2021. 'Niet-technische samenvatting 202011508: Opkweek van Europese paling van ei tot glasaal'. Centrale Commissie Dierproeven.
<https://www.centralecommissiedierproeven.nl/onderwerpen/niet-technische-samenvatting/onderzoek-behoud-diersoort/nts-202011508-europese-paling-glasaal-voortplanting-larven-larvale-voeding-kopie>

COGEM. 2018. 'CRISPR & Het Dier: Implicaties van genome editing voor maatschappij en beleid, COGEM signalering CGM/180501-01'. Bilthoven: Commissie Genetische Modificatie.

Commons, Kathryn G., Aram B. Cholanians, Jessica A. Babb, en Daniel G. Ehlinger. 2017. 'The Rodent Forced Swim Test Measures Stress-Coping Strategy, Not Depression-like Behavior'. *ACS Chemical Neuroscience* 8 (5): 955–60.
<https://doi.org/10.1021/acscemneuro.7b00042>

Comstock, Gary. 1992. 'The Moral Irrelevance of Autonomy'. *Between the Species*, 15–27.

Cook, David, Dearg Brown, Robert Alexander, Ruth March, Paul Morgan, Gemma Satterthwaite, en Menelas N. Pangalos. 2014. 'Lessons Learned from the Fate of AstraZeneca's Drug Pipeline: A Five-Dimensional Framework'. *Nature Reviews Drug Discovery* 13 (6): 419–31.
<https://doi.org/10.1038/nrd4309>

Cooper, David E. 1998. 'Intervention, humility and animal integrity'. In *Animal Biotechnology and Ethics*, door Allen Holland en Andrew Johnson, 145–55. Dordrecht: Springer Science + Business Media.



Cooper, Jane, en Maggy Jennings. 2008. 'Advancing animal welfare and the 3Rs in the batch testing of veterinary vaccines'. Horsham, UK: RSPCA.

Cruelty Free Europe. 2022. 'Replace Animal Tests: the RAT list: 10 animal tests that should be replaced'. 2022.

<https://www.crueltyfreeeurope.org/RATlist>

De Kloet, E. R., en M. L. Molendijk. 2016. 'Coping with the Forced Swim Stressor: Towards Understanding an Adaptive Mechanism'. *Neural Plasticity* 2016: 6503162.

<https://doi.org/10.1155/2016/6503162>

De Vries, Rob. 2004. 'Niet alleen ongerief: Waarom DEC's ook zouden moeten toetsen op aantasting van de integriteit van proefdieren'. In *DEC's in discussie: De beoordeling van dierproeven in Nederland*, 61–68. Budel: Damon.

De Vries. 2006. 'Genetic Engineering and the Integrity of Animals'. *Journal of Agricultural and Environmental Ethics* 19 (5): 469–93.

<https://doi.org/10.1007/s10806-006-9004-y>

De Vries. 2009. 'Genetic engineering and our duties to laboratory animals: A theoretical and empirical analysis of the concepts of intrinsic value and animal integrity Rob de'. Nijmegen: Radboud Universiteit Nijmegen.

Degeling, Chris, en Jane Johnson. 2013. 'Evaluating Animal Models: Some Taxonomic Worries'. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 38 (2): 91–106.

<https://doi.org/10.1093/jmp/jht004>

DeGrazia, David, en Tom L. Beauchamp. 2020. 'Principles of Animal Research Ethics'. In *Principles of Animal Research Ethics*, onder redactie van Tom L. Beauchamp en David DeGrazia, 5–41. Oxford: Oxford University Press.

Deisseroth, Karl. 2015. 'Optogenetics: 10 years of microbial opsins in neuroscience'. *Nature neuroscience* 18 (9): 1213–25.

<https://doi.org/10.1038/nn.4091>

Denayer, Tinneke, Thomas Stöhr, en Maarten Van Roy. 2014. 'Animal Models in Translational Medicine: Validation and Prediction'. *New Horizons in Translational Medicine* 2 (1): 5–11.

<https://doi.org/10.1016/j.nhtm.2014.08.001>

Eggel, Matthias, en Hanno Würbel. 2021. 'Internal Consistency and Compatibility of the 3Rs and 3Vs Principles for Project Evaluation of Animal Research'. *Laboratory Animals* 55 (3): 233–43.

<https://doi.org/10.1177/0023677220968583>

Engster, Daniel. 2006. 'Care Ethics and Animal Welfare'. *Journal of Social Philosophy* 37 (4): 521–36.

<https://doi.org/10.1111/j.1467-9833.2006.00355.x>

Engster. 2019. 'Care Ethics, Dependency, and Vulnerability'. *Ethics and Social Welfare* 13 (2): 100–114.

<https://doi.org/10.1080/17496535.2018.1533029>

Ethik-Studienkommission des Eidgenössischen Volkswirtschaftsdepartements zur Gentechnologie im ausserhumanen Bereich. 1995. 'Bericht der Ethik-Studienkommission des Eidgenössischen Volkswirtschaftsdepartements Gentechnologie im ausserhumanen Bereich'.

European Commission. 2020. 'REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL. 2019 report on the statistics on the use of animals for scientific purposes in the Member States of the European Union in 2015-2017'. Brussels: European Commission.

European Parliament. 2021. 'European Parliament Resolution of 16 September 2021 on Plans and Actions to Accelerate a Transition to Innovation without the Use of Animals in

Research, Regulatory Testing and Education (2021/2784(RSP))'.

https://www.europarl.europa.eu/doceo/document/TA-9-2021-0387_EN.html



European Parliament and the Council. 2019. *Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on Cosmetic Products (Recast)*.

<http://data.europa.eu/eli/reg/2009/1223/2019-o8-13/eng>

Europees Parlement en de Raad van de Europese Unie. 2010. 'RICHTLIJN 2010/63/EU VAN HET EUROPEES PARLEMENT EN DE RAAD van 22 september 2010 betreffende de bescherming van dieren die voor wetenschappelijke doeleinden worden gebruikt'. *Publicatieblad van de Europese Unie*, 20 oktober 2010.

Eurostat. 2021. 'Glossary:Basic research'. Eurostat: Statistics Explained. 20 juli 2021.

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Basic_research

Expert Working Group for Project Evaluation and Retrospective Assessment. 2011. 'National Competent Authorities for the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes: Working document on Project Evaluation and Retrospective Assessment'.

https://ec.europa.eu/environment/chemicals/lab_animals/pdf/Consensus_document.pdf

Fenwick, Nicole, Gilly Griffin, en Clément Gauthier. 2009. 'The welfare of animals used in science: How the "Three Rs" ethic guides improvements'. *The Canadian Veterinary Journal* 50 (5): 523–30.

Ferreira, Guilherme S., Désirée H. Veening-Griffioen, Wouter P. C. Boon, Ellen H. M. Moors, Christine C. Gispen-de Wied, Huub Schellekens, en Peter J. K. van Meer. 2019. 'A standardised framework to identify optimal animal models for efficacy assessment in drug development'. *PLoS ONE* 14 (6): e0218014.

<https://doi.org/10.1371/journal.pone.0218014>

Ferreira, Guilherme S., Désirée H. Veening-Griffioen, Wouter P. C. Boon, Ellen H. M. Moors, en Peter J. K. van Meer. 2020. 'Levelling the Translational Gap for Animal to Human Efficacy Data'. *Animals* 10 (7): 1199.

<https://doi.org/10.3390/ani10071199>

Fraser, D. 1995. 'Science, Values and Animal Welfare: Exploring the "Inextricable Connection"'. *Animal Welfare* 4 (2): 103–17.

Fraser, D. 2008. 'Animal Welfare and the Intensification of Animal Production'. In *The Ethics of Intensification: Agricultural Development and Cultural Change*, door Paul B. Thompson, 2008:167–89. The International Library of Environmental, Agricultural and Food Ethics 16. Dordrecht: Springer Science + Business Media B.V.

Fraser, D., D. M. Weary, E. A. Pajor, en B. N. Milligan. 1997a. 'A scientific conception of animal welfare that reflects ethical concerns'. *Animal Welfare* 6 (3): 187–205.

Fraser, D., D. Weary, E. Pajor, en B. Milligan. 1997b. 'A Scientific Conception of Animal Welfare that Reflects Ethical Concerns'. *Ethics and Animal Welfare*, januari.

<http://animalstudiesrepository.org/ethawel/1>

Freriks, A.A., B.M.J. Van der Meulen, H. Van den Belt, H. Ten Holt, en J. Verstappen. 2005. 'Noodzakelijk kwaad: Evaluatie Wet op de dierproeven'.

<https://edepot.wur.nl/35442>

Garner, Joseph P. 2014. 'The Significance of Meaning: Why Do over 90% of Behavioral Neuroscience Results Fail to Translate to Humans, and What Can We Do to Fix It?' *ILAR Journal* 55 (3): 438–56.

<https://doi.org/10.1093/ilar/ilu047>

Garner, Joseph P., Brianna N. Gaskill, Elin M. Weber, Jamie Ahloy-Dallaire, en Kathleen R. Pritchett-Corning. 2017. 'Introducing Therioepistemology: The Study of How Knowledge Is Gained from Animal Research'. *Lab Animal* 46 (4): 103–13.

<https://doi.org/10.1038/labana.1224>

Geyer, Mark, en Athina Markou. 1995. 'Animal Models of Psychiatric Disorders'. In *Psychopharmacology: The Fourth Generation of Progress*, onder redactie van Floyd E. Bloom en David J. Kupfer, 787–98. New York: Raven Press, Ltd.



Gjerris, Mickey, en C. Gamborg. 2010. 'Is there more to life than welfare? How the concept of animal integrity can contribute to discussions of animal ethics'. In *Global food security: ethical and legal challenges. EurSafe 2010, Bilbao, Spain, 16-18 September, 2010, 372-77*. Wageningen: Wageningen Academic Publishers.

Gjerris, Mickey, en Peter Sandøe. 2006. 'Farm animal cloning: The role of the concept of animal integrity in debating and regulating the technology'. In *Ethics and the Politics of Food, 6th EurSafe Congress, 22-24 June 2006, Oslo, Norway*, door Matthias Kaiser en Marianne Lien, 320-24. Wageningen: Wageningen Academic Publishers.

Greek, Ray, en Lisa A. Kramer. 2019. 'The Scientific Problems with Using Non-human Animals to Predict Human Response to Drugs and Disease'. In *Animal Experimentation: Working Towards a Paradigm Change*, onder redactie van Kathrin Herrmann en Kimberley Jayne, 391-416. Leiden: Brill.
<http://www.jstor.org/stable/10.1163/j.ctvjhzqof.8>

Green, T. C., en D. J. Mellor. 2011. 'Extending ideas about animal welfare assessment to include "quality of life" and related concepts'. *New Zealand Veterinary Journal* 59 (6): 263-71.
<https://doi.org/10.1080/00480169.2011.610283>

Grimm, Herwig. 2014. 'Ethics in Laboratory Animal Science'. In *Comparative Medicine: Anatomy and Physiology*, onder redactie van Erika Jensen-Jarolim, 281-300. Vienna: Springer.
https://doi.org/10.1007/978-3-7091-1559-6_15

Grimm, Herwig, I Anna S Olsson, en Peter Sandøe. 2019. 'Harm-benefit analysis – what is the added value? A review of alternative strategies for weighing harms and benefits as part of the assessment of animal research'. *Laboratory Animals* 53 (1): 17-27.
<https://doi.org/10.1177/0023677218783004>

Gross, Alexandra N., S. Helene Richter, A. Katarina J. Engel, en Hanno Würbel. 2012. 'Cage-Induced Stereotypies, Perseveration and the Effects of Environmental Enrichment in Laboratory Mice'. *Behavioural Brain Research* 234 (1): 61-68.
<https://doi.org/10.1016/j.bbr.2012.06.007>

Hau, Jann, en Steven J. Schapiro. 2011. *Handbook of Laboratory Animal Science, Volume II: Animal Models*. Boca Ration, FL: CRC Press.

Hauskeller, Michael. 2005. 'Telos: The Revival of an Aristotelian Concept in Present Day Ethics'. *Inquiry* 48 (1): 62-75.
<https://doi.org/10.1080/00201740510015356>

Hauskeller. 2007. 'The Reification of Life'. *Genomics, Society and Policy* 3 (2): 70.
<https://doi.org/10.1186/1746-5354-3-2-70>

Hauskeller. 2016. *Biotechnology and the Integrity of Life: Taking Public Fears Seriously*. London: Routledge.

Haverkamp, Beatrijs, Bernice Bovenkerk, en Marcel F Verweij. 2018. 'A Practice-Oriented Review of Health Concepts'. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 43 (4): 381-401.
<https://doi.org/10.1093/jmp/jhy011>

Hay, Michael, David W. Thomas, John L. Craighead, Celia Economides, en Jesse Rosenthal. 2014. 'Clinical Development Success Rates for Investigational Drugs'. *Nature Biotechnology* 32 (1): 40-51.
<https://doi.org/10.1038/nbt.2786>

Heeger, F.R., en Frans W.A. Brom. 2001. 'Beyond feeling well: Our direct duties towards animals'. In *EurSafe 2011. 'Food Safety, Food Quality and food Ethics'. The Third Congress of the European Society for Agricultural and Food Ethics. 3-5 October 2001. Florence, Italy, 243-46*. Milan: A&O.

Held, J.R. 1980. 'Muhlbock memorial lecture: Considerations in the provision and characterization of animal models'. In *Animal Quality and Models in Biomedical Research, 7th ICLAS Symposium Utrecht 1979*. Stuttgart, Germany: Gustav Fisher Verlag.

Henderson, Valerie C., Jonathan Kimmelman, Dean Fergusson, Jeremy M. Grimshaw, en Dan G. Hackam. 2013. 'Threats to Validity in the Design and Conduct of Preclinical Efficacy Studies: A Systematic Review of Guidelines for In Vivo Animal Experiments'. *PLOS Medicine* 10 (7): e1001489.
<https://doi.org/10.1371/journal.pmed.1001489>



Holdrege, Craig. 2002. 'Seeing the Integrity and Intrinsic Value of Animals: Developing Appreciative Modes of Understanding'. In *Genetic Engineering and the Intrinsic Value and Integrity of Animals and Plants. Proceedings of a Workshop at the Royal Botanical Garden, Edinburgh, UK, 18-21 September 2002*. Llanystumdwy, UK: International Forum for Genetic Engineering.

Holmes, Philip V. 2003. 'Rodent Models of Depression: Reexamining Validity without Anthropomorphic Inference'. *Critical Reviews in Neurobiology* 15 (2): 143-74.
<https://doi.org/10.1615/critrevneurobiol.v15.i2.30>

Home Office. 2021. 'Annual Statistics of Scientific Procedures on Living Animals Great Britain 2020. Presented to Parliament pursuant to section 21(7) and 21A(1) of the Animals (Scientific Procedures) Act 1986 Ordered by the House of Commons to be printed 15 July 2021'.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1002895/annual-statistics-scientific-procedures-living-animals-2020.pdf

Hooijmans, Carlijn R., Marlies Leenaars, en Merel Ritskes-Hoitinga. 2010. 'A Gold Standard Publication Checklist to Improve the Quality of Animal Studies, to Fully Integrate the Three Rs, and to Make Systematic Reviews More Feasible'. *Alternatives to Laboratory Animals* 38 (2): 167-82.
<https://doi.org/10.1177/026119291003800208>

Hursthouse, Rosalind. 2011. 'Virtue Ethics and the Treatment of Animals'. In *The Oxford Handbook of Animal Ethics*, onder redactie van Tom L. Beauchamp en Raymond G. Frey, 119-43. Oxford: Oxford University Press.

Insel, Thomas R. 2007. 'From Animal Models to Model Animals'. *Biological Psychiatry* 62 (12): 1337-39.
<https://doi.org/10.1016/j.biopsych.2007.10.001>

Janssens, Monique. 2020. *Dierenzaken: Blinde vlek van bedrijven*. Management Boek.

Kagan, Shelly. 2016. 'What's Wrong with Speciesism? (Society for Applied Philosophy Annual Lecture 2015)'. *Journal of Applied Philosophy* 33 (1): 1-21.
<https://doi.org/10.1111/japp.12164>

Kagen, sShelly. 2018. 'For Hierarchy in Animal Ethics'. *Journal of Practical Ethics* 6 (1): 1-18.

Kittay, Eva Feder. 2011. 'The Ethics of Care, Dependence, and Disability'. *Ratio Juris* 24 (1): 49-58.
<https://doi.org/10.1111/j.1467-9337.2010.00473.x>

Knight, Andrew. 2019. 'Critically Evaluating Animal Research'. In *Animal Experimentation: Working Towards a Paradigm Change*, onder redactie van Kathrin Herrmann en Kimberley Jayne, 321-40. Leiden: Brill.
<http://www.jstor.org/stable/10.1163/j.ctvjhzqof.8>

Knight, Sarah, en Louise Barnett. 2008. 'Justifying Attitudes toward Animal Use: A Qualitative Study of People's Views and Beliefs'. *Anthrozoös* 21 (1): 31-42.
<https://doi.org/10.2752/089279308X274047>

Kola, Ismail, en John Landis. 2004. 'Can the Pharmaceutical Industry Reduce Attrition Rates?' *Nature Reviews Drug Discovery* 3 (8): 711-16.
<https://doi.org/10.1038/nrd1470>

Korte, S. Mechiel, Berend Olivier, en Jaap M. Koolhaas. 2007. 'A new animal welfare concept based on allostasis'. *Physiology & Behavior* 92 (3): 422-28.
<https://doi.org/10.1016/j.physbeh.2006.10.018>

Kramer, K., en F. L. B. Meijboom. 2021. 'Using Breeding Technologies to Improve Farm Animal Welfare: What Is the Ethical Relevance of Telos?' *Journal of Agricultural and Environmental Ethics* 34 (1): 2.
<https://doi.org/10.1007/s10806-021-09843-6>



Kristensen, M., A. A. Nierenberg, en S. D. Østergaard. 2018. 'Face and Predictive Validity of the ClockΔ19 Mouse as an Animal Model for Bipolar Disorder: A Systematic Review'. *Molecular Psychiatry* 23 (1): 70–80.
<https://doi.org/10.1038/mp.2017.192>

LaFollette, Hugh, en Niall Shanks. 1996. *Brute Science: Dilemmas of Animal Experimentation*. London: Routledge.

Lammerts van Bueren, E. T., P. C. Struik, M. Tiemens-Hulscher, en E. Jacobsen. 2003. 'Concepts of Intrinsic Value and Integrity of Plants in Organic Plant Breeding and Propagation'. *Crop Science* 43 (6): 1922–29.
<https://doi.org/10.2135/cropsci2003.1922>

Landi, Margaret, Jeffrey Everitt, en B Berridge. 2021. 'Bioethical, Reproducibility, and Translational Challenges of Animal Models'. *ILAR Journal*, nr. ilaa027 (maart).
<https://doi.org/10.1093/ilar/ilaa027>

Leenaars, Cathalijn H. C., Carien Kouwenaar, Frans R. Stafleu, André Bleich, Merel Ritskes-Hoitinga, Rob B. M. De Vries, en Franck L. B. Meijboom. 2019. 'Animal to Human Translation: A Systematic Scoping Review of Reported Concordance Rates'. *Journal of Translational Medicine* 17 (1): 1–22.
<https://doi.org/10.1186/s12967-019-1976-2>

Lijmbach, Susanne. 1996. 'De actualiteit van Heideggers dierfilosofie'. In *Heidegger en de wereld van het dier*, onder redactie van Bert Blans en Susanne Lijmbach, 63–75. Assen: Van Gorcum & Comp.

Lund, Thomas Bøker, Jesper Lassen, en Peter Sandøe. 2012. 'Public Attitude Formation Regarding Animal Research'. *Anthrozoös* 25 (4): 475–90.
<https://doi.org/10.2752/175303712X13479798785896>

Lund, Thomas Bøker, Morten Raun Mørkbak, Jesper Lassen, en Peter Sandøe. 2014. 'Painful Dilemmas: A Study of the Way the Public's Assessment of Animal Research Balances Costs to Animals against Human Benefits'. *Public Understanding of Science* 23 (4): 428–44.
<https://doi.org/10.1177/0963662512451402>

Lund, Vonne. 2006. 'Natural living—a precondition for animal welfare in organic farming'. *Livestock Science* 100 (2–3): 71–83.
<https://doi.org/10.1016/j.livprodsci.2005.08.005>

Luisetto, Siro, Ornella Rossetto, Cesare Montecucco, en Flaminia Pavone. 2003. 'Toxicity of Botulinum Neurotoxins in Central Nervous System of Mice'. *Toxicon: Official Journal of the International Society on Toxinology* 41 (4): 475–81.
[https://doi.org/10.1016/s0041-0101\(02\)00370-7](https://doi.org/10.1016/s0041-0101(02)00370-7)

Macnaghten, Phil. 2004. 'Animals in Their Nature: A Case Study on Public Attitudes to Animals, Genetic Modification and "Nature"'. *Sociology* 38 (3): 533–51.
<https://doi.org/10.1177/0038038504043217>

Mak, Isabella WY, Nathan Evaniew, en Michelle Ghert. 2014. 'Lost in translation: animal models and clinical trials in cancer treatment'. *American Journal of Translational Research* 6 (2): 114–18.

Mason, G., R. Clubb, N. Latham, en S. Vickery. 2007. 'Why and how should we use environmental enrichment to tackle stereotypic behaviour?' *Applied Animal Behaviour Science* 102 (3–4): 163–88.
<https://doi.org/10.1016/j.applanim.2006.05.041>

Meijboom, Franck L. B., Elzbieta Kostrzewa, en Cathalijn H. C. Leenaars. 2020. 'Joining forces: the need to combine science and ethics to address problems of validity and translation in neuropsychiatry research using animal models'. *Philosophy, Ethics, and Humanities in Medicine* 15 (1): 1.
<https://doi.org/10.1186/s13010-019-0085-4>

Molendijk, Marc L., en E. Ronald De Kloet. 2015. 'Immobility in the Forced Swim Test Is Adaptive and Does Not Reflect Depression'. *Psychoneuroendocrinology* 62 (december): 389–91.
<https://doi.org/10.1016/j.psyneuen.2015.08.028>



Molendijk, Marc L., en E. Ronald De Kloet. 2019. 'Coping with the Forced Swim Stressor: Current State-of-the-Art'. *Behavioural Brain Research* 364 (mei): 1–10.

<https://doi.org/10.1016/j.bbr.2019.02.005>

Molendijk, Marc L., en E. Ronald De Kloet. 2021. 'Forced Swim Stressor: Trends in Usage and Mechanistic Consideration'. *European Journal of Neuroscience* n/a (n/a).

<https://doi.org/10.1111/ejn.15139>

Morrice, Jessica R., Cheryl Y. Gregory-Evans, en Christopher A. Shaw. 2018. 'Animal models of amyotrophic lateral sclerosis: a comparison of model validity'. *Neural Regeneration Research* 13 (12): 2050–54.

<https://doi.org/10.4103/1673-5374.241445>

NCad. 2020. 'Zorgvuldige weging van methoden voor identificatie en genetische karakterisatie'. Den Haag: Nationaal Comité advies dierproevenbeleid.

Nicholson, Pamela, Julia Furrer, Michael Hässig, Christian Strauss, Manfred Heller, Sophie Braga-Lagache, en Joachim Frey. 2019. 'Production of Neutralizing Antibodies against the Secreted Clostridium Chauvoei Toxin A (CctA) upon Blackleg Vaccination'. *Anaerobe* 56 (april): 78–87.

<https://doi.org/10.1016/j.anaerobe.2019.02.011>

Niemi, Steven M. 2021. 'Harm-Benefit Analyses Can Be Harmful'. *ILAR Journal* 60 (3): 341–46.

<https://doi.org/10.1093/ilar/ilaa016>

Nordgren, Anders. 2010. *For Our Children: The Ethics of Animal Experimentation in the Age of Genetic Engineering*. Amsterdam: Rodopi.

NOS. 2015. 'Universiteit Maastricht stopt dierproeven met honden'. 2 juli 2015.

<https://nos.nl/artikel/2044710-universiteit-maastricht-stopt-dierproeven-met-honden>

Novak, Janja, Jeremy D. Bailoo, Luca Melotti, en Hanno Würbel. 2016. 'Effect of Cage-Induced Stereotypies on Measures of Affective State and Recurrent Perseveration in CD-1 and C57BL/6 Mice'. Onder redactie van Kathleen R. Pritchett-Corning. *PLOS ONE* 11 (5): e0153203.

<https://doi.org/10.1371/journal.pone.0153203>

Nussbaum, Martha. 2006. *Een waardig bestaan. Over dierenrechten*. Vertaald door Rogier Van Kappel. Amsterdam: Ambo.

NVWA. 2022. 'Zo doende 2020. Jaaroverzicht dierproeven en proefdieren van de Nederlandse Voedsel- en Warenautoriteit'. Utrecht: Nederlandse Voedsel- en Warenautoriteit van het Ministerie van Landbouw, Natuur en Voedselkwaliteit.

Olsson, I Anna S, Christine J Nicol, Steven M Niemi, en Peter Sandøe. 2019. 'From Unpleasant to Unbearable—Why and How to Implement an Upper Limit to Pain and Other Forms of Suffering in Research with Animals'. *ILAR Journal* 60 (3): 404–14.

<https://doi.org/10.1093/ilar/ilzo18>

Olsson, I. Anna S., en Peter Sandøe. 2021. 'Chapter 1 - Animal Research Ethics'. In *Handbook of Laboratory Animal Science: Essential Principles and Practices*, onder redactie van Jann Hau en Steven J. Schapiro. Boca Raton, FL: CRC Press.

Olsson, I Anna S, Orsolya Varga, en Peter Sandøe. 2015. 'A Matter of Importance: Considering Benefit in Animal Ethics Review'. *ALTEX - Proceedings* 4 (1): 33–36.

Ormandy, Elisabeth H., en Catherine A. Schuppli. 2014. 'Public Attitudes toward Animal Research: A Review'. *Animals* 4 (3): 391–408.

<https://doi.org/10.3390/ani4030391>

Ortiz, Sara Elizabeth Gavrell. 2004. 'Beyond Welfare: Animal Integrity, Animal Dignity, and Genetic Engineering'. *Ethics and the Environment* 9 (1): 94–120.

PETA. 2020. 'The Research Modernisation Deal'. London: People for the Ethical Treatment of Animals Foundation.

PETA. 2022. 'Victories! PETA Is Ending Near-Drowning Experiments on Animals'. 2022.

<https://www.peta.org/features/peta-ends-near-drowning-tests-small-animals/>

Porsolt, R. D., M. Le Pichon, en M. Jalfre. 1977. 'Depression: A New Animal Model Sensitive to Antidepressant Treatments'. *Nature* 266 (5604): 730–32.

<https://doi.org/10.1038/266730a0>



Pound, Pandora, en Christine J. Nicol. 2018. 'Retrospective Harm Benefit Analysis of Pre-Clinical Animal Research for Six Treatment Interventions'. *PLOS ONE* 13 (3): e0193758.

<https://doi.org/10.1371/journal.pone.0193758>

Pound, Pandora, en Merel Ritskes-Hoitinga. 2018. 'Is It Possible to Overcome Issues of External Validity in Preclinical Animal Research? Why Most Animal Models Are Bound to Fail'. *Journal of Translational Medicine* 16 (304): 1–8.

<https://doi.org/10.1186/s12967-018-1678-1>

Procaccini, Claudio, Veronica De Rosa, Valentina Pucino, Luigi Formisano, en Giuseppe Matarese. 2015. 'Animal models of Multiple Sclerosis'. *European Journal of Pharmacology* 759 (juli): 182–91.

<https://doi.org/10.1016/j.ejphar.2015.03.042>

Ram, Rebecca. 2019. 'Extrapolation of Animal Research Data to Humans: An Analysis of the Evidence'. In *Animal Experimentation: Working Towards a Paradigm Change*, onder redactie van Kathrin Herrmann en Kimberley Jayne, 341–75. Leiden: Brill. <http://www.jstor.org/stable/10.1163/j.ctvjhzqof.8>.

RDA. 2016. 'Fokkerij en Voortplantingstechnieken: Anima(l) sana in corpore sano'. Raad voor Dierenaangelegenheden. <https://www.rda.nl/publicaties/zienswijzen/2016/02/29/zienswijze-fokkerij-en-voortplantingstechnieken-uitgebreid-verslag-pdf>.

RDA. 2018a. 'Denkkader Dierenwelzijn'. Raad voor Dierenaangelegenheden.

<https://www.rda.nl/publicaties/publicaties/2018/12/21-rda-denkkader/denkkader>

RDA. 2018b. 'Dierproeven ten behoeve van de Veehouderij. Over het belang van dierenwelzijn, duurzaamheid en een ketenbenadering'. Raad voor Dierenaangelegenheden.

<https://www.rda.nl/publicaties/zienswijzen/2018/10/01/rda-zienswijze-dierproeven-ten-behoeve-van-de-veehouderij>

Rhouma, Mohamed, John Morris Fairbrother, Francis Beaudry, en Ann Letellier. 2017. 'Post Weaning Diarrhea in Pigs: Risk Factors and Non-Colistin-Based Control Strategies'. *Acta Veterinaria Scandinavica* 59 (1): 31.

<https://doi.org/10.1186/s13028-017-0299-7>

Röcklinsberg, H., C. Gamborg, en M. Gjerris. 2014. 'A Case for Integrity: Gains from Including More than Animal Welfare in Animal Ethics Committee Deliberations'. *Laboratory Animals* 48 (1): 61–71.

<https://doi.org/10.1177/0023677213514220>

Rogers Brambell, F.W., D.S. Barbour, M.B. Barnett, T.K. Ewer, A. Hobson, H. Pitchforth, W.R. Smith, W.H. Thorpe, en F.J.W. Winship. 1965. 'Report of the Technical Committee to Enquire into the Welfare of Animals kept under Intensive Livestock Husbandry Systems. Presented to Parliament by the Secretary of State for Scotland and the Minister of Agriculture, Fisheries and Food by Command of Her Majesty'. London: Her Majesty's Stationery Office.

Rolston, Holmes. 2002. 'What do we mean by the intrinsic value and integrity of plants and animals?' In *Genetic Engineering and the Intrinsic Value and Integrity of Animals and Plants. Proceedings of a Workshop at the Royal Botanical Garden, Edinburgh, UK, 18-21 September 2002*, 5–10. Llanystumdwy, UK: International Forum for Genetic Engineering.

Russell, William Moy Stratton, en Rex Leonard Burch. 1959. *The Principles of Humane Experimental Technique*. London: Methuen.

Rust, Aleksander, Ciara Doran, Rosalyn Hart, Thomas Binz, Paul Stickings, Dorothea Sesardic, Andrew A. Peden, en Bazbek Davletov. 2017. 'A Cell Line for Detection of Botulinum Neurotoxin Type B'. *Frontiers in Pharmacology* 8.

<https://www.frontiersin.org/article/10.3389/fphar.2017.00796>

Rutgers, Bart, en Robert Heeger. 1999. 'Inherent Worth and Respect for Animal Integrity'. In *Recognizing the Intrinsic Value of Animals: Beyond Animal Welfare*, onder redactie van Marcel Dol, Martje Fentener van Vlissingen, Soemini Kasanmoentalib, Thijs Visser, en Hub Zwart, 41–52. Assen: Uitgeverij Van Gorcum.

Sams-Dodd, Frank. 1999. 'Phencyclidine in the Social Interaction Test: An Animal Model of Schizophrenia with Face and Predictive Validity'. *Reviews in the Neurosciences* 10 (1): 59–90.

<https://doi.org/10.1515/REVNEURO.1999.10.1.59>

Sams-Dodd, Frank. 2006. 'Strategies to Optimize the Validity of Disease Models in the Drug Discovery Process'. *Drug Discovery Today* 11 (7–8): 355–63.



Sandgren, Eric P., Robert Streiffer, Jennifer Dykema, Nadia Assad, en Jackson Moberg. 2020. 'Attitudes toward Animals, and How Species and Purpose Affect Animal Research Justifiability, among Undergraduate Students and Faculty'. *PLOS ONE* 15 (5): e0233204. <https://doi.org/10.1371/journal.pone.0233204>

Sandøe, Peter, Nuno H Franco, Thomas B Lund, Daniel M Weary, en I Anna S Olsson. 2015. 'Harms to Animals – Can We Agree on How Best to Limit Them?' *ALTEX - Proceedings* 4 (1): 28–32.

Schmidt, Kirsten. 2008. 'Würde oder Integrität - Verlangt die gentechnische Veränderung von Tieren neue tierethische Konzepte?' *ALTEX* 25 (4): 313–20.

Schurgers, R.G. 2005. 'Ethiek in dierexperimentencommissies'. P-UB-2005-02. Utrechtse Wetenschapswinkels, voor maatschappijgericht onderzoek. Utrecht: Universiteit Utrecht. <https://dspace.library.uu.nl/handle/1874/45043>.

Scruton, Roger. 1998. *Animal Rights and Wrongs*. London, UK: Demos.

Sewell, Fiona, Ian Waterson, David Jones, Mark David Tricklebank, en Ian Ragan. 2021. 'Preclinical Screening for Antidepressant Activity – Shifting Focus Away from the Forced Swim Test to the Use of Translational Biomarkers'. *Regulatory Toxicology and Pharmacology* 125 (oktober): 105002. <https://doi.org/10.1016/j.yrtph.2021.105002>

Singer, Peter. 1993. *Practical Ethics*. 2nd dr. Cambridge: Cambridge University Press.

Singh, Ashutosh, Rahul Soloman Singh, Phulen Sarma, Gitika Batra, Rupa Joshi, Hardeep Kaur, Amit Raj Sharma, Ajay Prakash, en Bikash Medhi. 2020. 'A Comprehensive Review of Animal Models for Coronaviruses: SARS-CoV-2, SARS-CoV, and MERS-CoV'. *Virologica Sinica* 35 (3): 290–304. <https://doi.org/10.1007/s12250-020-00252-z>

Singh, K.P., N.S. Parihar, en B.N. Tripathi. 1992. 'Pathology and pathogenesis of Clostridium chauvoei infection in guinea-pigs.' *Indian Journal of Animal Sciences* 62: 611–15.

Sjoberg, Espen A. 2017. 'Logical fallacies in animal model research'. *Behavioral and Brain Functions* 13 (1): 3. <https://doi.org/10.1186/s12993-017-0121-8>

Slattery, David A., en John F. Cryan. 2012. 'Using the Rat Forced Swim Test to Assess Antidepressant-like Activity in Rodents'. *Nature Protocols* 7 (6): 1009–14. <https://doi.org/10.1038/nprot.2012.044>

Smith, Jane A., en Kenneth M. Boyd, red. 1991. *Lives in the Balance: The Ethics of Using Animals in Biomedical Research. The Report of a Working Party of the Institute of Medical Ethics*. Oxford: Oxford University Press.

Steel, Daniel. 2010. 'A New Approach to Argument by Analogy: Extrapolation and Chain Graphs'. *Philosophy of Science* 77 (5): 1058–69. <https://doi.org/10.1086/656543>

Swearingen, James R. 2018. 'Choosing the Right Animal Model for Infectious Disease Research'. *Animal Models and Experimental Medicine* 1 (2): 100–108. <https://doi.org/10.1002/ame2.12020>

Tannenbaum, J. 1991. 'Ethics and animal welfare: the inextricable connection'. *Journal of the American Veterinary Medical Association* 198 (8): 1360–76.

Thomas, David W., Justin Burns, John Audette, Adam Carroll, Corey Dow-Hygelund, en Michael Hay. 2016. 'Clinical Development Success Rates 2006-2015'. <https://www.bio.org/sites/default/files/legacy/bioorg/docs/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf>

Treit, Dallas, Elif Engin, en Kris McEown. 2010. 'Animal Models of Anxiety and Anxiolytic Drug Action'. In *Behavioral Neurobiology of Anxiety and Its Treatment*, onder redactie van Murray B. Stein en Thomas Steckler, 121–60. Current Topics in Behavioral Neurosciences. Berlin, Heidelberg: Springer. https://doi.org/10.1007/7854_2009_17



Tricklebank, Mark, en Joseph Garner. 2012. 'The Possibilities and Limitations of Animal Models for Psychiatric Disorders'. In *Drug Discovery for Psychiatric Disorders*, onder redactie van Z. Rankovic, M. Bingham, E.J. Nestler, en R. Hargreaves, 28:534–56. London: The Royal Society of Chemistry.

<https://doi.org/10.1039/9781849734943-00534>

Trouw. 1995. 'Onderzoekers laten mens voor op muis groeien'. *Trouw*, 26 oktober 1995, sec. voorpagina.

<https://www.trouw.nl/gs-b6fe515d>

Trunnell, Emily R., en Constança Carvalho. 2021. 'The Forced Swim Test Has Poor Accuracy for Identifying Novel Antidepressants'. *Drug Discovery Today* 26 (12): 2898–2904.

<https://doi.org/10.1016/j.drudis.2021.08.003>

Tweede Kamer der Staten-Generaal. 2021. 'Wijziging van de Wet dieren in verband met de uitvoering van de herziene Europese diergezondheidswetgeving. Amendement van het lid Vesterling. Ontvangen 21 april 2021'.

https://www.parlementairemonitor.nl/9353000/1/j4nvgs5kjqz7kof_j9vwijsepjmj1eyo/vli6ae95lkqr/f=/kst353989.pdf

Van den Heuvel, Steven C., Patrick Nullens, en Angela Roothaan, red. 2017. *Theological Ethics and Moral Value Phenomena: The Experience of Values*. 1st edition. London: Routledge.

Van der Staay, F. Josef, Saskia S. Arndt, en Rebecca E. Nordquist. 2009. 'Evaluation of animal models of neurobehavioral disorders'. *Behavioral and Brain Functions*: BBF 5 (februari): 11.

<https://doi.org/10.1186/1744-9081-5-11>

Van der Worp, H. Bart, David W. Howells, Emily S. Sena, Michelle J. Porritt, Sarah Rewell, Victoria O'Collins, en Malcolm R. Macleod. 2010. 'Can Animal Models of Disease Reliably Inform Human Studies?' *PLOS Medicine* 7 (3): e1000245.

<https://doi.org/10.1371/journal.pmed.1000245>

VanDeVeer, Donald. 1995. 'Interspecific Justice'. In *People, Penguins, and Plastic Trees: Basic Issues in Environmental Ethics*, onder redactie van Christine Pierce en Donald VanDeVeer, 85–99. Belmont, CA: Wadsworth Publishing Company.

Verhoog, Henk. 1992. 'The Concept of Intrinsic Value and Transgenic Animals'. *Journal of Agricultural and Environmental Ethics* 5 (2): 147–60.

<https://doi.org/10.1007/bfo1966357>

Verhoog, Henk. 2003. 'Naturalness and the genetic modification of animals'. *Trends in Biotechnology* 21 (7): 294–97.

[https://doi.org/10.1016/S0167-7799\(03\)00142-2](https://doi.org/10.1016/S0167-7799(03)00142-2)

Verhoog, Henk. 2004. 'Utilisme en de integriteit van het dier'. In *DEC's in discussie: De beoordeling van dierproeven in Nederland*, 51–60. Budel: Damon.

Verhoog, Henk. 2007. 'The tension between common sense and scientific perception of animals: recent developments in research on animal integrity'. *NJAS: Wageningen Journal of Life Sciences* 54 (4): 361–73.

[https://doi.org/10.1016/S1573-5214\(07\)80009-1](https://doi.org/10.1016/S1573-5214(07)80009-1)

Vieira de Castro, Ana Catarina, en I. Anna S. Olsson. 2015. 'Does the Goal Justify the Methods? Harm and Benefit in Neuroscience Research Using Animals'. In *Ethical Issues in Behavioral Neuroscience*, onder redactie van Grace Lee, Judy Illes, en Frauke Ohl, 47–78. Current Topics in Behavioral Neurosciences. Berlin, Heidelberg: Springer.

https://doi.org/10.1007/7854_2014_319

Visser, M.B.H., en Henk Verhoog. 1999. 'De aard van het beestje: De morele relevantie van natuurlijkheid'. NWO Ethiek en Beleid.

Voelkl, Bernhard, Hanno Würbel, Martin Krzywinski, en Naomi Altman. 2021. 'The Standardization Fallacy'. *Nature Methods* 18 (1): 5–7.

<https://doi.org/10.1038/s41592-020-01036-9>

Vorstenbosch, J. 1993. 'The concept of integrity. Its significance for the ethical discussion on biotechnology and animals'. *Livestock Production Science, Biotechnology and Animal Welfare*, 36 (1): 109–12.

[https://doi.org/10.1016/0301-6226\(93\)90144-7](https://doi.org/10.1016/0301-6226(93)90144-7)



VWA. 2004. 'Zo doende 2003. Jaaroverzicht van de Voedsel en Waren Autoriteit over dierproeven en proefdieren'. Den Haag: Voedsel en Waren Autoriteit.

Wehling, Martin. 2009. 'Assessing the Translatability of Drug Projects: What Needs to Be Scored to Predict Success?' *Nature Reviews. Drug Discovery* 8 (7): 541–46.
<https://doi.org/10.1038/nrd2898>

Wendler, Alexandra, en Martin Wehling. 2012. 'Translatability Scoring in Drug Development: Eight Case Studies'. *Journal of Translational Medicine* 10 (1): 39.
<https://doi.org/10.1186/1479-5876-10-39>

Wendler. 2017. 'Translatability Score Revisited: Differentiation for Distinct Disease Areas'. *Journal of Translational Medicine* 15 (1): 226. <https://doi.org/10.1186/s12967-017-1329-y>.

Wensley, Madie R, Mike D Tokach, Jason C Woodworth, Robert D Goodband, Jordan T Gebhardt, Joel M DeRouchev, en Denny McKilligan. 2021. 'Maintaining continuity of nutrient intake after weaning. II. Review of post-weaning strategies'. *Translational Animal Science* 5 (1): txabo22.
<https://doi.org/10.1093/tas/txabo22>

Wet op de dierproeven. 2021.
<https://wetten.overheid.nl/BWBR0003081/2021-07-01>

Willner, Paul. 1984. 'The Validity of Animal Models of Depression'. *Psychopharmacology* 83 (1): 1–16.
<https://doi.org/10.1007/BF00427414>

Willner, Paul. 1991. 'Methods for Assessing the Validity of Animal Models of Human Psychopathology'. In *Animal Models in Psychiatry*, I, door Alan A. Boulton, Glen B. Baker, en Mathew T. Martin-Iverson, 18:1–24. New Jersey: Humana Press.
<https://doi.org/10.1385/0-89603-198-5:1>

Willner, Paul, en P. J. Mitchell. 2002. 'The Validity of Animal Models of Predisposition to Depression'. *Behavioural Pharmacology* 13 (3): 169–88.

Willner, Paul, Richard Muscat, en Mariusz Papp. 1992. 'Chronic Mild Stress-Induced Anhedonia: A Realistic Animal Model of Depression'. *Neuroscience & Biobehavioral Reviews* 16 (4): 525–34.
[https://doi.org/10.1016/S0149-7634\(05\)80194-0](https://doi.org/10.1016/S0149-7634(05)80194-0)

Wong, Chi Heem, Kien Wei Siah, en Andrew W Lo. 2019. 'Estimation of clinical trial success rates and related parameters'. *Biostatistics* 20 (2): 273–86.
<https://doi.org/10.1093/biostatistics/kxx069>

Würbel, Hanno. 2017. 'More than 3Rs: The Importance of Scientific Validity for Harm-Benefit Analysis of Animal Research'. *Lab Animal* 46 (4): 164–66.
<https://doi.org/10.1038/lablan.1220>

Zeiss, Caroline J. 2015. 'Improving the Predictive Value of Interventional Animal Models Data'. *Drug Discovery Today* 20 (4): 475–82.
<https://doi.org/10.1016/j.drudis.2014.10.015>

Zhai, Chongkai, Mingda Wang, Hea-Jong Chung, Mehedi Hassan, Seungkoo Lee, Hyeon-Jin Kim, en Seong-Tshool Hong. 2021. 'Roborovski hamster (*Phodopus roborovskii*) strain SH101 as a systemic infection model of SARS-CoV-2'. *Virulence* 12 (1): 2430–42.
<https://doi.org/10.1080/21505594.2021.1972201>



Appendix A

Table 1. Lijst van geïnterviewde personen

Persoon	Organisatie	Details
Bas Blaauboer	Centrale Commissie Dierproeven, Universiteit Utrecht	Interview op 29-09-2021
Anoniem	Radboudumc	Interview op 30-09-2021
Jan Langermans	Biomedical Primate Research Centre, Universiteit Utrecht	Interview op 30-09-2021
Anoniem	Universiteit Maastricht	Interview op 04-10-21
Nelleke Verhave	Instantie voor Dierenwelzijn Leids Universitair Medisch Centrum/ Nederlandse Vereniging van Dierexperimentencommissies	Interview op 05-10-21
Harry Emmen	Charles River Laboratories	Interview op 8-10-21
Erik Baltussen	Charles River Laboratories	Interview op 8-10-21
Lennert Schrader	Nederlandse Voedsel en Waren Autoriteit	Duo interview op 14-10-21
Anoniem	Nederlandse Voedsel en Waren Autoriteit	Duo interview op 14-10-21
Saskia Aan	Proefdiervrij	Groepsinterview op 7-10-21
Janneke Hogervorst	People for the Ethical Treatment of Animals (PETA)	Groepsinterview op 7-10-21
Elly von Jessen	Dierenbescherming	Groepsinterview op 7-10-21
Anoniem	Animal Rights	Groepsinterview op 7-10-21



Bijlage 4

Teksten over fundamenteel onderzoek uit het Werkdocument over projectevaluatie en beoordeling achteraf

Teksten uit: Nationale instanties die bevoegd zijn voor de tenuitvoerlegging van Richtlijn 2010/63/EU betreffende de bescherming van dieren die voor wetenschappelijke doeleinden worden gebruikt. Werkdocument over projectevaluatie en beoordeling achteraf. Brussel, 18 en 19 september 2013⁵⁷.

Doelstellingen van het project bij fundamenteel onderzoek (p.11 en 12):

- de huidige stand van de kennis waarop het project beoogt voort te bouwen;
- de manier waarop het project zal bijdragen aan kennisbevordering.

Voor zover van toepassing moet ook melding worden gemaakt van streefdoelen die door middel van eerdere projecten werden behaald: de aanvrager moet een beknopt overzicht geven van de geboekte vorderingen en duidelijk vermelden welke specifieke doelstellingen middels dit project worden nagestreefd.

Op bepaalde gebieden van het fundamenteel onderzoek kan kennisverruiming een volwaardige en geschikte doelstelling zijn. Niettemin moet waar mogelijk een koppeling worden gemaakt met de verspreiding van resultaten (rekening houdend met de intellectuele eigendom) en mogelijke baten op langere termijn.

Beoordeling van de baten bij fundamenteel onderzoek (p.24):

- Beter begrip van het vraagstuk dat aan de orde is (kennisvermeerdering – erkenning van het belang om een informatie- en kenniskloof te dichten).
- Waar mogelijk dient „kennisvermeerdering” als belangrijkste baat in verband te worden gebracht met een tastbaarder strategisch doel, zelfs wanneer de kans bestaat dat ruimere baten pas veel later worden gerealiseerd en minder goed te voorspellen zijn. De baten moeten het *nice-to-know*-principe overstijgen.

⁵⁷ Werkdocument over projectevaluatie en beoordeling achteraf, dat is samengesteld door de werkgroep van Nationale instanties die bevoegd zijn voor de tenuitvoerlegging van Richtlijn 2010/63/EU. https://ec.europa.eu/environment/chemicals/lab_animals/pdf/guidance/project_evaluation/nl.pdf



- Omvang van de verbetering (voor mens, dier en milieu) in aantallen; grootte en kwaliteit (behoefte aan gefundeerde oordeelsvorming: weesgeneesmiddelen kunnen aan een klein aantal patiënten worden toegediend, maar verstrekkende gevolgen hebben op individueel niveau), en belasting die het probleem betekent voor de maatschappij (zowel wat fundamenteel als wat toegepast onderzoek betreft).
- Erkenning van hypothese-gestuurd fundamenteel onderzoek: uit de evaluatie moet blijken dat de hypothese wetenschappelijk verantwoord en realistisch is.
- Voor fundamenteel onderzoek is de verspreiding van zowel positieve als negatieve informatie van het allergrootste belang om ervoor te zorgen dat de baten worden gerealiseerd.

Belangrijke overwegingen bij de beoordeling van baten (p.25):

1. Houd rekening met de onmiddellijke/korte-termijn-baten (producten, gegevens, resultaten).
2. Houd rekening met de baten op langere termijn (producten, specifieke lange-termijn-baten).
3. Houd rekening met de ruimere effecten (kennisvermeerdering, translationeel potentieel).

Contactgegevens

Nationaal Comité advies dierproevenbeleid

Postbus 93118 | 2595 AL Den Haag

Mail: ncad@rvo.nl | website: ncadierproevenbeleid.nl

April 2023 | Publicatie-nr. 23402600