

# **(ON)BETROUWBARE MORELE INTUÏTIES**

Bachelorscriptie Wijsbegeerte

Naam: Lisan Gunsing

Studentnummer: 4157427

Inleverdatum: 24-06-18

Hoofdbegeleider: Hanno Sauer

Tweede beoordelaar: Chris Meyns

Faculteit: Geesteswetenschappen

Departement: Filosofie en Religiewetenschap

## **SAMENVATTING**

Jonathan Haidt stelt in zijn sociaal intuïtionistisch model dat morele oordelen primair worden gevormd door snelle, automatische intuïties. Bewuste morele redeneringen spelen geen causale rol in dit proces. Een probleem voor dit model is dat deze morele intuïties niet betrouwbaar zijn. Intuïties maken gebruik van onbetrouwbare heuristieken, zijn gevoelig voor framing en variëren in gevallen waar stabiele oordelen verondersteld worden. Ik stel in deze scriptie dat dit niet hoeft te betekenen dat morele intuïties altijd onbetrouwbaar zijn. Ik beargumenteer dat het flexibele leersysteem zoals ontwikkeld door Peter Railton beter in staat is om te laten zien waar onze intuïties naar verwachting betrouwbaarder zijn en waar minder. Enerzijds is het systeem op evaluatieve wijze in staat intuïties bij te werken, en anderzijds maakt het systeem gebruik van impliciete kennis en competenties om morele oordelen af te stemmen op de werkelijkheid. Acceptatie van dit systeem vereist fundamentele herziening van het model van Haidt. Het betekent echter niet dat alle morele intuïties betrouwbaar zijn en zonder kritische blik aangenomen moeten worden.

# INHOUDSOPGAVE

Inleiding .....	4
1. Sociaal intuïtionisme .....	6
1.1 Haidt's social intuitionist model .....	6
1.2 SIM over rationaliteit en morele redeneringen .....	7
1.3 Morele intuïties en oordeelsvorming in cognitieve wetenschappen .....	8
2. Onbetrouwbare morele intuïties .....	10
2.1 Imperfecte heuristieken en weerbarstige intuïties .....	10
2.2 Variërende intuïties .....	12
2.3 Rationaliteit, confabulatie en morele intuïties.....	14
3. Intelligente intuïties en het flexibele leersysteem .....	16
3.1 Railton's flexibele leersysteem .....	16
3.2 Flexibel leersysteem als betrouwbaarder alternatief voor het SIM.....	18
3.3 Beperkingen van het flexibele leersysteem .....	20
4. Conclusie .....	21
Referenties.....	23

## INLEIDING

Een belangrijk vraagstuk in de morele psychologie is de vraag naar de werking van morele oordeelsvorming. Dit is tegenwoordig een interdisciplinair vraagstuk dat voortbouwt op empirische bronnen uit de cognitieve psychologie en daarnaast op conceptuele bronnen uit de ethiek.<sup>1</sup> Lang hebben rationalistische modellen centraal gestaan waarbij een dominante rol aan ‘de rede’ werd toegekend. Rationalisten stellen over het algemeen dat morele oordelen worden veroorzaakt door bewuste redenering. Als reactie hierop zijn intuïtionistische modellen ontwikkeld waarin de causale rol van de rede wordt betwijfeld. Intuïtionisten stellen over het algemeen dat morele oordelen primair worden veroorzaakt door snelle, automatische intuïties zonder dat bewuste redeneringen een causale rol in het proces spelen. Het basisidee is dat we over het algemeen niet tot de overtuiging komen dat iets moreel goed of fout is door erover te redeneren, maar dat het hebben van een affectieve intuïtie leidt tot het oordeel dat iets goed of fout is.<sup>2</sup> Een van de pioniers op het gebied van morele intuïties is Jonathan Haidt. Zijn indrukwekkende *social intuitionist model* (SIM) is nog steeds invloedrijk.<sup>3</sup> Het SIM stelt niet alleen dat intuïties primair van belang zijn in morele oordeelsvorming, maar stelt ook dat sociale invloeden een belangrijke rol hierin hebben.<sup>4</sup>

In deze scriptie neem ik het SIM als uitgangspunt. Een belangrijke zorg voor het SIM is het probleem dat intuïties niet in staat zouden zijn om consistente en betrouwbare morele oordelen te leveren. Deze scriptie legt de focus op dit probleem. Er zijn verschillende ideeën over het vergroten van de betrouwbaarheid van morele intuïties, maar deze ideeën lopen vaak tegen fundamentele problemen aan.<sup>5</sup> Peter Railton betoogt een alternatief affectief systeem dat stelt dat morele intuïties onderdeel zijn van een flexibel leersysteem dat op evaluatieve wijze in staat is intuïties bij te werken en daarbij gebruik maakt van impliciete kennis en competenties.<sup>6</sup> Ik wil onderzoeken of dit flexibele leersysteem van Railton een betrouwbaar alternatief

---

<sup>1</sup> John Doris, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley, “Moral Psychology: Empirical Approaches,” in *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/win2017/entries/moral-psych-emp/> (geraadpleegd 6 juni 2018).

<sup>2</sup> Renée Jeffery, *Reason and Emotion in International Ethics* (Cambridge University Press, 2014), 39-41; David O. Brink, “Principles and Intuitions in Ethics: Historical and Contemporary Perspectives,” *Ethics* 124, no. 4 (July 2014): 681, <https://doi.org/10.1086/675878>.

<sup>3</sup> Terry Horgan and Mark Timmons, “Morphological Rationalism and the Psychology of Moral Judgment,” *Ethical Theory and Moral Practice* 10 (2017): 281, <https://doi.org/10.1007/s10677-007-9068-4>.

<sup>4</sup> Jonathan Haidt, “The Emotional Dog and its Rational Tail,” *Psychological Review* 108, no. 4 (2001): 814-834, <https://doi.org/10.1037/0033-295X.108.4.814>.

<sup>5</sup> Brink, “Principles and Intuitions in Ethics,” 665-667, 681.

<sup>6</sup> Peter Railton, “The Affective Dog and Its Rational Tale: Intuition and Attunement,” in *Ethics* 124, no. 4 (July 2014): 813-859, <https://doi.org/10.1086/675876>.

is voor het SIM van Haidt. Anders gezegd: zijn intuïties die worden gegenereerd door een flexibel leersysteem betrouwbaardere intuïties dan intuïties die worden gegenereerd door het SIM? Ik zal beargumenteren dat het SIM onbetrouwbare morele oordelen genereert en dat het flexibele leersysteem inderdaad betrouwbaardere morele oordelen biedt, maar dit betekent niet dat het flexibele leersysteem volledig betrouwbaar is. Ik zal in deze scriptie niet ingaan op de normatieve vraag of we op morele intuïties moeten vertrouwen, maar mijn focus ligt op de descriptieve vraag: kunnen we op morele intuïties vertrouwen?

In het eerste hoofdstuk zal ik het SIM van Haidt nader toelichten door te expliciteren wat Haidt onder morele intuïties en morele redenering verstaat. Het empirisch gefundeerde model stelt dat intuïties primair van belang zijn in morele oordeelsvorming, en dat morele redenering hoofdzakelijk post hoc rationalisaties zijn. Het claimt redenering, emotie, intuïtie en sociale beïnvloeding volledig te integreren.<sup>7</sup> Om context te bieden zal ik het SIM plaatsen in een breder empirisch debat rondom morele oordeelsvorming.

In het tweede hoofdstuk zal ik ingaan op de vraag wat betrouwbare morele oordelen zijn en waarom (veelal) rationalisten stellen dat het model van Haidt onbetrouwbare morele oordelen oplevert. Ik focus me daarbij op twee argumenten: ten eerste het empirische argument dat intuïties gebruikmaken van onbetrouwbare vuistregels en ten tweede het argument dat intuïties variabel zijn op verschillende manieren. Ik onderzoek wat de consequenties van deze argumenten voor het SIM zijn. Tot slot ga ik in op aannames die Haidt maakt over rationaliteit en redenering, die de betrouwbaarheidsproblemen versterken en alternatieve benaderingen uitsluit.

In het derde en laatste hoofdstuk ga ik in op het flexibele leersysteem van Railton dat een alternatief affectief systeem beoogt dat stelt dat morele intuïties geïntegreerd zijn in een breed cognitief systeem dat gebruik maakt van impliciete kennis en competenties om morele oordelen correct af te stemmen op de werkelijkheid. Fundamentele herziening van het SIM zou nodig om het systeem te accepteren vanwege Haidt's aannames op het gebied van rationaliteit en redenering. Ik onderzoek of het betrouwbaardere morele oordelen oplevert dan het SIM en stel dat dit het geval is, maar dat betekent niet dat het flexibele leersysteem perfect is.

---

<sup>7</sup> Haidt, "The Emotional Dog," 828.

# 1. SOCIAAL INTUITIONISME

Ik dit eerste hoofdstuk zal het SIM van Jonathan Haidt inleiden en fundeer het model in een breder dominant psychologisch systeem dat het bestaan en belang van intuïties bevestigt. Het SIM zelf is gevormd uit conclusies van veel voorgaande onderzoeken in de psychologie, cognitieve- en neurowetenschappen die betrekking hebben op oordeelsvorming. Het SIM claimt bovendien redenering, emotie, intuïtie en sociale beïnvloeding volledig te integreren, en lijkt daarom op het eerste oog een goede kandidaat voor betrouwbare intuïties.<sup>8</sup>

## 1.1 Haidt's social intuitionist model

De centrale positie die Haidt verdedigt is dat morele oordelen niet worden veroorzaakt door bewuste deliberatie, maar primair door automatische morele intuïties.<sup>9</sup> Haidt fundeert zijn model op twee manieren. Ten eerste betwijfelt hij de oorzakelijkheid van morele redenering in morele oordeelsvorming. Hij ontkent niet dat morele redenering bestaat, maar hij stelt dat niet volgt dat dergelijke redeneringen de oorzaak zijn van morele oordelen.<sup>10</sup> Hier kom ik in paragraaf 1.2 op terug. Ten tweede laat hij middels bestaande en empirisch gefundeerde onderzoeksresultaten zien dat morele oordelen primair het product zijn van onbewuste en automatische intuïtieve processen. Deze intuïties verschijnen plotseling en moeiteloos in het bewustzijn in de vorm van een affectieve evaluatie (dat wil zeggen: een emotionele reactie) van goed of afkeuring. Nadat het oordeel is gevormd volgt er post hoc redenering: er worden argumenten gezocht die een reeds gemaakt oordeel ondersteunen en vervolgens het oordeel aan anderen rechtvaardigen.<sup>11</sup> Sociale beïnvloeding kan hier een rol spelen: mensen kunnen elkaars intuïtie beïnvloeden en door groepsnormen elkaars oordelen overnemen.<sup>12</sup>

De intuïties zijn volgens Haidt onderdeel van onze cognitie en zijn de automatisch output van onderliggende, grotendeels onbewuste processen en onderling verbonden morele concepten.<sup>13</sup> Deze zijn van sociale context afhankelijk en daardoor een aspect van de cultuur.<sup>14</sup>

---

<sup>8</sup> Haidt, "The Emotional Dog," 817.

<sup>9</sup> Haidt, "The Emotional Dog," 817.

<sup>10</sup> Haidt, "The Emotional Dog," 817.

<sup>11</sup> Haidt, "The Emotional Dog," 819; Het onderscheid tussen intuïtie en redenering is niet hetzelfde onderscheid als die tussen emotie en cognitie.

<sup>12</sup> Haidt, "The Emotional Dog," 819.

<sup>13</sup> Haidt, "The Emotional Dog," 825.

<sup>14</sup> Haidt, "The Emotional Dog," 818.

Haidt noemt morele intuïties zowel aangeboren als cultureel, omdat het een evolutionaire aanpassing is die tegelijk vorming vanuit een bepaalde cultuur vereist.<sup>15</sup>

## 1.2 SIM over rationaliteit en morele redeneringen

Het SIM is een anti-rationalistisch model dat stelt dat de rol van ‘de rede’ overschat wordt door rationalisten: de werking van morele oordeelsvorming is niet zo rationeel als rationalisten beweren. Waarom zouden we sociaal intuïtionisme boven rationalisme moeten verkiezen? Haidt noemt vier redenen om te twijfelen aan het causale belang van de rede, en hierin zit vooral de aantrekkingskracht van het model. Ten eerste noemt Haidt empirische onderzoeken die het bestaan bevestigen van automatische processen in menselijke cognitie. Dit heb ik paragraaf 1.1 beschreven. Ten tweede zijn er onderzoeken die laten zien dat we selectief naar bewijs zoeken voor claims waar we al achter staan. Ten derde zijn er aanwijzingen dat mensen de neiging hebben om irrelevante factoren aan te halen wanneer hen gevraagd te beschrijven hoe ze tot een bepaald oordeel kwamen.<sup>16</sup> Tot slot introduceert Haidt bewijs dat er een zwak verband is tussen moreel redeneren en morele handelingen, en dat morele handelingen nauwer verbonden zijn met morele emoties dan met morele oordelen. Dit suggereert dat morele emoties motiverender werken dan morele redematies.<sup>17</sup>

Uit deze stellingen en empirische onderzoeken concludeert Haidt dat de redelijke processen in morele oordeelsvorming gewoonlijk naderhand plaatsvinden in de vorm van een rechtvaardiging van de morele intuïtie en niet causaal van belang zijn. Haidt ontkent niet morele redeneringen niet bestaan, maar hij stelt dat zulke redeneringen niet de oorzaak zijn van morele oordelen. Het is plausibeler dat redeneringen de consequentie zijn van het oordeel. Het zijn met andere woorden post hoc rationalisaties.<sup>18</sup> Haidt vergelijkt de rol van redenering in morele oordeelsvorming met “a lawyer trying to build a case rather than a judge searching for the truth.”<sup>19</sup> De morele intuïtie ontstaat primair en achteraf probeert men daar redenen voor te vinden. De redenen die ze daarvoor geven zijn redenen die in werkelijkheid geen rol speelden in het produceren van het betreffende oordeel; oordelen zijn ongevoelig voor redenen. Dit wordt ‘confabulatie’ genoemd.<sup>20</sup> Mensen kunnen echter ook niet in staat zijn om redenen te kunnen geven voor hun intuïtie, of kunnen de redenen niet uitleggen. Ze hebben onverklaar-

---

<sup>15</sup> Haidt, “The Emotional Dog,” 826-827.

<sup>16</sup> Haidt, “The Emotional Dog,” 819-825; Steve Clarke, “SIM and the city: Rationalism in psychology and philosophy and Haidt’s account of moral judgment,” *Philosophical Psychology* 21, no. 6 (2008): 806, <https://doi.org/10.1080/09515080802513250>.

<sup>17</sup> Haidt, “The Emotional Dog,” 825; Clarke, “SIM and the city,” 807.

<sup>18</sup> Haidt, “The Emotional Dog,” 825; Clarke, “SIM and the city,” 807.

<sup>19</sup> Haidt, “The Emotional Dog,” 814.

<sup>20</sup> Horgan and Timmons, “Morphological Rationalism,” 281-282; Clarke, “SIM and the city,” 807.

bare ingevingen of overtuigingen zonder ondersteunende redenen. Dit wordt ‘moral dumbfounding’ genoemd.<sup>21</sup>

Haidt erkent echter ook dat er zeldzame gevallen zijn waarin het mogelijk is dat bewuste redenering de onbewuste intuïties kan aanpassen. Ten eerste noemt hij het beredeneerde oordeel: mensen (vooral filosofen) kunnen naar een oordeel toe redeneren waarbij zij hun oorspronkelijke intuïtie negeren. Ten tweede noemt hij de persoonlijke reflectie: door na te denken kan iemand een nieuwe intuïtie activeren die tegenstrijdig is met de eerste intuïtie.<sup>22</sup>

### 1.3 Morele intuïties en oordeelsvorming in cognitieve wetenschappen

Er is de laatste jaren veel cognitief onderzoek gedaan naar de werking en vorming van oordelen en morele oordelen. Het bestaan van morele intuïties wordt door veel onderzoeken bevestigd. Intuïties worden gezien als een emotionele reactie en blijken vaak een cruciale rol te spelen in (morele) oordelen, zelfs in grondig overwogen en goed geïnformeerde beoordelingen.<sup>23</sup> Emoties worden daarbij gezien als een cognitief proces.<sup>24</sup> Vrijwel elke fMRI-studie heeft aangetoond dat er in morele oordeelsvorming sprake is van een emotionele reactie.<sup>25</sup> Hoewel er nog wel debat is, wordt de intuïtionistische morele psychologie van oordeelsvorming bevestigd door het tegenwoordig in de psychologie dominante duaal procesmodel van cognitie. De essentie van dit model is dat menselijke cognitie bestaat uit zowel snelle, onbewuste, automatische en intuïtieve cognitie (systeem 1) en langzame, bewuste en deliberatieve cognitie (systeem 2). Deze mentale ‘subsystemen’ zijn geen aanwijsbare entiteiten in de hersenen; ‘systeem’ wordt gebruikt om mentale processen van elkaar te onderscheiden.<sup>26</sup>

Er is consensus dat morele intuïties gebruikmaken van systeem 1. Simpel gezegd legt de rationalist dus de focus op systeem 2 en intuïtionisten op systeem 1. De discussie lijkt er vooral om te draaien welk systeem het meest dominant is.<sup>27</sup> De opvattingen van zowel rationalisten als intuïtionisten bouwen dus op dezelfde psychologische elementen die de morele

---

<sup>21</sup> Daniel Jacobson, “Moral Dumbfounding and Moral Stupefaction,” in *Oxford Studies in Normative Ethics*, vol. 2, ed. Mark Timmons (Oxford: Oxford University Press, 2012), 289.

<sup>22</sup> Haidt, “The Emotional Dog,” 819.

<sup>23</sup> Mark Alfano, *Moral Psychology: An Introduction* (Cambridge: Polity Press, 2016), 97; Jeffery, *Reason and Emotion in International Ethics*, 160-164; Railton, “The Affective Dog,” 816.

<sup>24</sup> Dit betekent niet dat emoties niet ook lichamelijk kunnen zijn.

<sup>25</sup> Jesse J. Prinz and Shaun Nichols, “Moral Emotions,” in *The Moral Psychology Handbook*, ed. John M. Doris (New York: Oxford University Press, 2010), 115-116, <https://doi.org/10.1093/acprof:oso/9780199582143.003.0005>; Jeffery, *Reason and Emotion in International Ethics*, 160-164.

<sup>26</sup> Alfano, *Moral Psychology*, 99-103; Daniel Kahneman, “A perspective on judgment and choice: Mapping bounded rationality,” *American Psychologist* 58 (September 2003), no. 9: 697–720, <https://doi.org/10.1037/0003-066X.58.9.697>; Brink, “Principles and Intuitions in Ethics,” 681-682; Clarke, “SIM and the city,” 805.

<sup>27</sup> Clarke, “SIM and the city,” 805-806; Alfano, *Moral Psychology*, 99-100.



psychologie van mensen vormen. Deze elementen omvatten onder andere: verschillende emoties (bijvoorbeeld afgunst en angst), intuïties, morele oordelen en morele redenering. Het verschil tussen de opvattingen is hoe en in welke mate deze elementen betrokken zijn in het produceren van morele oordelen.<sup>28</sup> De data lijkt dus consistent met verschillende modellen van morele oordeelsvorming.<sup>29</sup>

---

<sup>28</sup> Horgan and Timmons, "Morphological Rationalism," 281.

<sup>29</sup> Prinz and Nichols, "Moral Emotions," 115-116.

## 2. ONBETROUWBARE MORELE INTUÏTIES

Tot nu toe hebben we gezien dat Haidt sterke redenen heeft om het bestaan van morele intuïties te erkennen en het SIM aan te nemen. In dit hoofdstuk zal ik echter middels twee verschillende argumenten laten zien dat morele oordelen die voortkomen uit morele intuïties over het algemeen als onbetrouwbaar kunnen worden beschouwd. Vooral empirisch onderzoek kan vragen oproepen over de betrouwbaarheid of bewijskracht van morele intuïties. De vraag is of hiermee morele kennis wordt ondermijnd. Tot slot plaats ik de twee bezwaren in context van aannames van het SIM met betrekking tot rationaliteit en redenering.

### 2.1 Imperfecte heuristieken en weerbarstige intuïties

Het eerste bezwaar dat ik wil noemen stelt dat intuïties gebruik maken van imperfecte heuristieken. Betrouwbare intuïties lijken een goede, betrouwbare infrastructuur te vereisen. Wanneer de infrastructuur niet betrouwbaar blijkt te zijn, lijkt het lastig te worden om op intuïties te kunnen vertrouwen. Er zijn volgens psychologen en cognitieve wetenschappers duidelijke aanwijzingen dat dit het geval is. Zij hebben ontdekt dat we functioneren door middel van grove vuistregels of heuristieken die ons helpen om op een soepele manier complexe informatie te verwerken en in ingewikkelde situaties beslissingen te maken.<sup>30</sup> Heuristieken worden over het algemeen gezien als onderdeel van systeem 1 cognitie van het duale proces model: ze ontstaan snel, automatisch en intuïtief.<sup>31</sup> Het is essentieel voor onze overleving en efficiënte taakbeheer dat we vaak en over het algemeen gebruik maken van betrouwbare vuistregels. Dergelijke heuristieken zijn geïnternaliseerd en werken betrouwbaar in bekende situaties.<sup>32</sup>

Heuristieken hebben echter de neiging om onze morele intuïties onbewust te beïnvloeden doordat ze in onbekende situaties te ver door generaliseren of juist onvoldoende generaliseren. Dit kan zorgen voor (grote) fouten in de oordeelsvorming. Deze fouten ontstaan meestal doordat heuristieken gebruikmaken van opvallende kenmerken in een situatie die een ruwe indicatie is voor een meer complexe reeks van relevante eigenschappen die moeilijker te ont-

---

<sup>30</sup> Cass Sunstein, "Moral Heuristics," *Behavioral and Brain Sciences* 28 (2005): 532, <https://doi.org/10.1017/S0140525X05400096>; Walter Sinnott-Armstrong, Liane Young, and Fiery Cushman, "Moral Intuitions," in *The Moral Psychology Handbook*, ed. John M. Doris (New York: Oxford University Press, 2010), 250-249-267, <https://doi.org/10.1093/acprof:oso/9780199582143.003.0008>; Brink, "Principles and Intuitions in Ethics," 681.

<sup>31</sup> Kahneman, "A perspective on judgment and choice," 697–720.

<sup>32</sup> Brink, "Principles and Intuitions in Ethics," 681-684.

dekken en onderscheiden zijn.<sup>33</sup> Daardoor lijkt het het geval dat heuristische in deze situaties niet in staat zijn om conclusies te trekken over individuele gevallen en daardoor bij systeem 2 het vermogen ontbreekt om gebalanceerde, toekomstgerichte besluiten te kunnen vormen.<sup>34</sup>

Volgt uit deze imperfecte morele heuristiek dat we onze morele intuïties niet meer kunnen vertrouwen? Dat lijkt wel het geval. Betrouwbare intuïties lijken een goede, betrouwbare infrastructuur te vereisen. Wanneer de infrastructuur niet betrouwbaar blijkt te zijn, wordt het lastig om op morele intuïties te kunnen vertrouwen. Als morele intuïties inderdaad ontstaan met behulp van imperfecte vuistregels, dan hebben we een goede reden om aan te nemen dat we niet op de morele intuïties kunnen en zouden moeten vertrouwen. Aannemend dat het argument dat intuïties gebruik maken van imperfecte vuistregels klopt, zouden we kunnen stellen dat morele intuïties onbetrouwbaar zijn omdat de algemene betrouwbaarheid van intuïties niet te garanderen is. Cass Sunstein nuanceert dit en stelt dat we sommige morele intuïties niet meer kunnen vertrouwen. Meer specifiek denk hij dat het een vergissing is om in nieuwe of exotische gevallen te vertrouwen op onze intuïtie omdat dit de gevallen zijn waarop onze oordeelsvorming niet voor is ontworpen. De methode wordt verdacht wanneer implicaties van testprincipes voor nieuwe gevallen worden vergeleken met onze intuïties over zulke gevallen.<sup>35</sup>

Een oplossing zou kunnen zijn om de heuristiek als uitvoeringsfouten te beschouwen. Door omstandigheden te herkennen waarin een heuristiek zich kan vergissen, en als gevolg hiervan verkeerde resultaten geeft, zou met behulp van principes hieruit kunnen worden bepaald of een heuristiek betrouwbaar is en nuttig is, of dat dat niet het geval is.<sup>36</sup> Zo bezien zou het niet het geval hoeven te zijn dat imperfecte heuristiek een negatieve invloed hebben op de morele methodologie omdat de imperfecte heuristiek herkend kunnen worden en de hieruit volgende morele intuïties herzien. Voor het op deze wijze nagaan van de betrouwbaarheid van morele intuïties zouden echter twee zaken vereist zijn. Ten eerste moeten intuïties omkeerbaar zijn om de imperfecte intuïties te kunnen herzien. Ten tweede is het bestaan van principes nodig om te verklaren waarom een heuristiek betrouwbaar is en wanneer intuïties juist in strijd zijn met die principes. Zo bezien worden intuïties beschouwd als potentieel bewijsmateriaal.

---

<sup>33</sup> Brink, "Principles and Intuitions in Ethics," 681-684; Sinnott-Armstrong, Young, and Cushman, "Moral Intuitions," 250-259.

<sup>34</sup> Railton, "The Affective Dog," 833.

<sup>35</sup> Sunstein, "Moral Heuristics," 531-73.

<sup>36</sup> Brink, "Principles and Intuitions in Ethics," 684.

Er zijn hier echter twee bezwaren tegenin te brengen. Ten eerste volgt uit het anti-rationalisme van het SIM dat het bestaan van principes wordt afgewezen. Om het genoemde alternatief te accepteren is dus een veel rationelere insteek vereist.<sup>37</sup> In die zin is het alternatief geen oplossing voor de betrouwbaarheid van morele intuïties vanuit intuïtionisme, maar een oplossing vanuit rationalisme dat stelt dat we niet zonder kritische redenering primair op morele intuïties kunnen vertrouwen. Hierdoor valt de mogelijkheid weg om op niet-rationele gronden een onderscheid te maken tussen meer en minder betrouwbare intuïties.

Ten tweede stelt het SIM dat herziening van morele intuïties een zeldzaamheid is. Dit impliceert dat morele intuïties in principe ongevoelig zijn voor potentieel vernietigend bewijs. Herziening van overtuigingen is zeldzaam en vooral een zaak van filosofen, aldus het SIM. Anders gezegd is kunnen intuïties gewoonlijk hun fouten niet herstellen. De imperfecties van de heuristieken lijken pas echt problematisch te worden als er sprake is van weerbarstige intuïties die niet open staan voor aanpassing en herziening. Intuïties lijken star en ongevoelig te zijn voor potentieel vernietigend bewijs. Als dit klopt dan missen morele intuïties betrouwbare bewijskracht en kunnen niet worden beschouwd als potentieel bewijsmateriaal waardoor de algemene betrouwbaarheid niet kan worden gegarandeerd. Ik ga er hier uit van alledaagse morele oordeelsvorming waar geen gebruik wordt gemaakt van het beredeneerde oordeel of persoonlijke reflectie. Het erkennen van die twee processen zou een rationelere insteek vereisen waar ik niet op inga in de scriptie.

## **2.2 Variërende intuïties**

Het tweede bezwaar dat ik wil aankaarten is het probleem van variërende intuïties. Hiermee bedoel ik niet dat intuïties open staan voor verandering – waar ik in 2.1 tegen heb gepleit – maar ik bedoel dat intuïties niet consistent en stabiel zijn. Er zijn duidelijke aanwijzingen waaruit geconcludeerd kan worden dat intuïties op verschillende manieren variabel zijn. Deze verschillende manieren vormen in verschillende mate een probleem voor de betrouwbaarheid van morele intuïties. Ten eerste bestaat er variatie tussen verschillende culturen. Het is echter een bekend fenomeen dat niet alle intuïties universeel gedeeld zijn.<sup>38</sup> Dit vormt voor het SIM dan ook geen probleem, integendeel zelfs. Haidt stelt zelf dat cultuur de basis van intuïties vormt. Dat hieruit variërende intuïties voortkomen is meer dan logisch. Echter, er zijn ook aanwijzingen dat niet alleen tussen culturen er variatie bestaat, ook binnen culturen zijn ver-

---

<sup>37</sup> Voor een model dat automatische morele oordelen met het bestaan van morele principes combineert, zie Horgan and Timmons, “Morphological Rationalism,” 279-295.

<sup>38</sup> Brink, “Principles and Intuitions in Ethics,” 691.

schillen.<sup>39</sup> Neem bijvoorbeeld genderspecifieke verschillen in morele intuïties. In de studie ‘Gender and Philosophical Intuition’ hebben Wesley Buckwalter and Stephen Stich ontdekt dat er aanzienlijke variatie is in filosofische intuïtie tussen verschillende genders. Het opvallende aan het onderzoek is dat deze verschillen in intuïtie niet beperkt blijven tot gevallen die als controversieel worden beschouwd, maar dat het ook gevallen betreft waarin verondersteld wordt of zou moeten worden dat intuïties op grote schaal worden gedeeld, zoals zaken waarbij onschuldigen worden gestraft.<sup>40</sup> Dit probleem lijkt wel een bezwaar te zijn tegen het SIM omdat het stelt dat intuïties juist cultuurafhankelijk zijn. Hoe kunnen we dan nog op ze vertrouwen? Het fenomeen van morele onenigheid is echter niet nieuw. Toch is het wel een grote zorg voor de rechtvaardiging van ethiek en moraliteit. Als moraliteit afhangt van overtuigingen waar we over van mening verschillen, hoe kunnen we dan ooit bepalen wat moreel juist is? Het is echter wel de vraag of overeenstemming wel mogelijk en wenselijk is. Moet de onenigheid in intuïties wel opgelost worden? Het antwoord op deze vraag laat ik in het midden.

Het belangrijkste bezwaar lijkt mij de derde manier waarop morele intuïties variabel zijn, namelijk binnen individuen zelf. Er zijn aanwijzingen dat mensen onderhevig zijn aan framing effecten. Intuïties blijken niet consistent binnen een persoon te zijn: er is interne variatie. Dit is zorgwekkend voor de betrouwbaarheid van de methodologie. De framing effecten zijn een bekend onderwerp in empirische wetenschappen. Daniel Kahneman en Amos Tversky onderzochten bijvoorbeeld of er framing effecten ontstaan door de manier waarop beleid wordt beschreven. Ze ontdekten dat tussen de verschillende onderzoeksgroepen verschillende uitkomsten bleken te ontstaan die verband hielden met de beschrijving van het beleid, namelijk of dat in termen van geredde levens was of in termen van verloren levens. De verschillende groepen proefpersonen bleken conflicterende voorkeuren hebben over inhoudelijk dezelfde optie. Hieruit kan geconcludeerd worden dat voorkeur afhankelijk is van hoe de opties gepresenteerd worden. Het is aannemelijk dat wat geldt voor voorkeuren ook kan gelden voor morele oordelen en morele intuïties.<sup>41</sup> Het lijkt er dus op dat mensen inconsistent zijn en onder invloed van externe factoren hun intuïtie vormen.

Ik wil hiermee laten zien dat er aanwijzingen zijn dat de intuïtionistische methodologie van morele oordeelsvorming onbetrouwbaar is. Intuïties blijken te variëren en inconsequent te

---

<sup>39</sup> Brink, “Principles and Intuitions in Ethics,” 691.

<sup>40</sup> Wesley Buckwalter and Stephen Stich, “Gender and Philosophical Intuition,” in *Experimental Philosophy*, vol. 2, ed. J. Knobe and S. Nichols (Oxford: Clarendon, 2013), 338-339.

<sup>41</sup> Brink, “Principles and Intuitions in Ethics,” 688-690.

zijn, maar bovenal zijn intuïties niet stabiel.<sup>42</sup> Intuïtief lijkt het voor een betrouwbare methodologie een vereiste te zijn dat de oordelen die eruit voortkomen stabiel zijn, in ieder geval is dit wenselijk.

### **2.3 Rationaliteit, confabulatie en morele intuïties**

Tot nu toe hebben we gezien dat er sterke aanwijzingen zijn waaruit geconcludeerd kan worden dat morele intuïties onbetrouwbare maatstaven voor morele oordelen zijn. Er is echter een aantal problemen met aannames die Haidt maakt met betrekking tot rationaliteit en morele redenering die oplossingen uitsluiten en bovengenoemde problemen bevestigen.

Ten eerste stelt het SIM dat het intuïtieve proces altijd vooraf gaat aan het rationele proces en is onbereikbaar voor bewuste redelijke processen. Dit volgt uit het feit dat morele redenering post hoc rationalisatie is. Hierdoor zijn alle vormen van redenering volgens het SIM niet in staat rechtvaardiging voor morele intuïties te bieden. Natuurlijk is SIM een anti-rationalistisch model dat de causale rol van morele redenering ontkent. Echter, het lijkt het geval te zijn dat Haidt te snel conclusies heeft getrokken met betrekking tot morele redenering. Het feit dat we soms dumbfounded zijn en soms confabuleren, betekent niet dat alledaagse redenen die worden aangevoerd voor moreel oordeel louter post hoc rationalisaties zijn. Het feit dat een oordeel gemaakt kan worden zonder bewuste redenering betekent niet dat oordelen helemaal niet of zelden gevoelig zijn voor beredeneerde rechtvaardiging voor die overtuiging.<sup>43</sup> De aanname bevestigt daarnaast het probleem van de imperfecte heuristieken. Intuïties zijn over het algemeen onbereikbaar voor aanpassing; verkeerde intuïties worden meestal niet herzien.

Ten tweede lijkt het SIM te stellen dat of iets rationeel is afhangt van het feit of het een causaal en toegankelijk proces is: rationaliteit wordt geïdentificeerd met bewustwording en causaliteit. Ik sluit me hier aan bij de aanname van het SIM zoals die geformuleerd is door Horgan & Timmons: “Unless conscious moral reasoning is part of the process leading to moral judgment, subjects’ reason-giving efforts are not a matter of citing considerations that really did play a causal role in the generation of the judgments in question; rather, reason-giving in these cases is a matter of confabulation.”<sup>44</sup> Anders gezegd: alleen voorafgaande, bewuste rationale redeneringen zijn redeneringen die causale invloed kunnen hebben op morele oor-

---

<sup>42</sup> Hieruit kan niet geconcludeerd worden dat variatie in intuïtieve oordelen niet ook in morele redenering voorkomt. Anders gezegd: variatie hoeft niet specifiek aan intuïties toegeschreven te worden. Ik ga verder niet in op de betrouwbaarheid van morele redenering.

<sup>43</sup> Jacobson, “Moral Dumbfounding,” 293-294; Horgan and Timmons, “Morphological Rationalism,” 289; Brink, “Principles and Intuitions in Ethics,” 686.

<sup>44</sup> Horgan and Timmons, “Morphological Rationalism,” 282.

deelsvorming. Dit volgt uit het feit dat alleen het beredeneerde oordeel en persoonlijke reflectie van invloed kunnen zijn op morele oordeelsvorming. Alle niet-causale processen zijn post hoc rationalisaties. Haidt stelt hiermee impliciet dat causale morele deliberatie altijd een bewust proces is. Haidt zet hier echter een te simplistisch en onrealistisch beeld van rationaliteit en morele redenering neer dat veel andere (realistische) opties uitsluit. Hiermee ontkent Haidt dat onbewuste redelijke processen van causale invloed kunnen zijn. Hij impliceert hij hiermee dat een onbewust intuïtief proces nooit rationeel kan zijn. Een automatische reactie is onbewust en daarom niet rationeel.

Uit het samennemen van beide aannames volgt dat Haidt impliceert dat causale morele redenering is altijd een bewust proces is en morele intuïties niet gevoelig kunnen zijn voor onbewuste morele redeneringen. Hiermee sluit hij uit dat causale redengeving iets anders kan zijn dan een niet-automatisch en toegankelijk proces. Anders gezegd: door de aannames sluit Haidt meer rationele of intelligente invullingen van intuïtie uit. Het model ontkent dat intuïties gebruik kunnen maken van cognitieve informatie zoals impliciete praktische kennis of intelligentie. Haidt stelt dat rationalisten stellen dat moraliteit uitgaat van een set van kennisstructuren.<sup>45</sup> Ik zal echter nu beargumenteren dat dit wel verenigd kan worden met intuïtisme.

---

<sup>45</sup> Haidt, "The Emotional Dog," 816.

### 3. INTELLIGENTE INTUÏTIES EN HET FLEXIBELE LEERSYSTEEM

Tot nu hebben we gezien dat het SIM problemen heeft met het bieden van een betrouwbare basis voor morele oordelen. Is een meer intelligente invulling van intuïties om de betrouwbaarheid ervan te vergroten? In dit derde en laatste hoofdstuk biedt ik een nieuwe invulling van het intuïtieve systeem: het flexibele leersysteem van Peter Railton waar hij een flexibel en meer intelligent alternatief biedt voor morele intuïties. Een fundamentele herziening van het SIM zou nodig zijn om het systeem te accepteren. Ik onderzoek of het betrouwbare morele oordelen oplevert dan het SIM biedt, en stel dat dat deels het geval is. Ik zal eerst het systeem uitleggen en vervolgens per bezwaar nagaan of het flexibele leersysteem een betrouwbare oplossing biedt, en concludeer dat dat deels het geval is.

#### 3.1 Railton's flexibele leersysteem

Railton biedt een alternatief systeem voor het intuïtieve systeem (het affectieve systeem in de terminologie van Railton) zoals dat wordt ingevuld door het duaal proces model. De kritiek die hij hierop heeft komt overeen met de kritiek die ik in hoofdstuk 2 uiteen heb gezet. Railton verzamelt indrukwekkend bewijs uit empirische onderzoeken dat een heel ander en nieuw beeld van het affectieve systeem schetst: er zijn veel aanwijzingen dat het affectieve systeem in plaats van onderdeel van het duale proces model van cognitie een complex leersysteem is dat gebruik maakt van impliciete kennis en competentie.<sup>46</sup> Adam Gjesdal noemt het een nieuw ontdekte verfijning van onze cognitieve processen die naturalistische steun biedt voor het vertrouwen en gebruik van morele intuïties in filosofische argumenten.<sup>47</sup>

Railton beoogt het filosofisch gebruik van morele intuïties te rechtvaardigen door te laten zien dat intuïties een psychologische etiologie hebben in rationale systemen die zijn afgestemd op moreel relevante kenmerken van de werkelijkheid.<sup>48</sup> Anders gezegd zijn morele intuïties geïntegreerd in een breed cognitief systeem dat gebruik maakt van impliciete kennis en competenties om morele oordelen correct af te stemmen op de werkelijkheid.<sup>49</sup> Railton noemt dit het *tacit-competence-based* model van intuïtie: intuïties zijn manifestaties van onderliggende competenties en impliciete kennis, oftewel van onderliggende structuren van re-

---

<sup>46</sup> Railton, "The Affective Dog," 833; Adam Gjesdal, "Moral Learning, Rationality, and the Unreliability of Affect," *Australasian Journal of Philosophy* (2017): 1, <https://doi.org/10.1080/00048402.2017.1366535>.

<sup>47</sup> Gjesdal, "Moral Learning," 1.

<sup>48</sup> Gjesdal, "Moral Learning," 1.

<sup>49</sup> Railton, "The Affective Dog," 814.



gels en informatie.<sup>50</sup> Het afstemmen op de werkelijkheid gebeurt middels een complex flexibel leersysteem dat informatie verwerkt op basis van ervaring en representaties en op de achtergrond gebruik maakt van competenties zoals het herkennen van statistische afhankelijkheden en het afwegen van voordelen, kosten en risico's. Het leert ons om in de loop van de tijd beter te reageren.<sup>51</sup> Het affectieve systeem wordt hier niet gezien als tegenstrijdig aan of verstorend voor cognitieve, evaluatieve, berekende en beslissingscapaciteiten, maar als een integraal onderdeel ervan.<sup>52</sup>

De onderzoeken schetsen een beeld waaruit blijkt dat emoties te begrijpen zijn als reacties op interne en externe stimuli die (evolutionair) zijn ontwikkeld om het organisme systematisch af te stemmen op veranderende fysieke en sociale omstandigheden, en dienen ervoor om het organisme te situeren, het in te lichten over de situatie en richting te geven aan het handelen.<sup>53</sup> Zo heeft onderzoek in de affectieve neurowetenschap ontdekt dat affectieve systemen een complexe set van interactieve subsystemen bevatten.<sup>54</sup> De systemen worden continu bijgewerkt door middel van 'verschil verkleinende' leerprocessen. Deze processen genereren voortdurend verwachtingen en vergelijken deze met werkelijke resultaten en produceren vervolgens een leersignaal dat ervoor zorgt dat de verwachtingen gecorrigeerd worden. Anders gezegd zijn het functies die continu de verwachte waarde bijwerken.<sup>55</sup>

Op deze alternatieve manier beschouwd zijn affectieve reacties plausibele kandidaten voor de belichaming van praktische intelligentie en kennis.<sup>56</sup> Affectieve toestanden kunnen passende representaties van waarde vormen, of kunnen een element van praktische kennis vormen die handelingen 'op de juiste manier' kan sturen om het op redenen te laten reageren.<sup>57</sup> Morele intuïties stemmen via rationele systemen af op moreel relevante kenmerken van de werkelijkheid die superveniëren op morele feiten.<sup>58</sup> Deze morele feiten vormen dus de basis van morele kenmerken, en morele intuïties stemmen af op de kenmerken. Het leerproces voldoet volgens Gjesdal aan geavanceerde principes van rationele probabilistische gevolgtrekkingen.<sup>59</sup> Samengevat pleit Railton voor een nieuwe invulling van morele intuïties dat enerzijds een flexibel systeem is dat in staat is zaken te herzien en te corrigeren. Het systeem

---

<sup>50</sup> Railton, "The Affective Dog," 816-817.

<sup>51</sup> Railton, "The Affective Dog," 833, 846; Gjesdal, "Moral Learning," 1.

<sup>52</sup> Railton, "The Affective Dog," 835-836.

<sup>53</sup> Railton, "The Affective Dog," 834.

<sup>54</sup> Railton, "The Affective Dog," 834.

<sup>55</sup> Railton, "The Affective Dog," 834-835.

<sup>56</sup> Railton, "The Affective Dog," 846-847.

<sup>57</sup> Railton, "The Affective Dog," 840-841.

<sup>58</sup> Gjesdal, "Moral Learning," 1.

<sup>59</sup> Gjesdal, "Moral Learning," 1.

wordt op evaluatieve wijze bijgewerkt en gevormd om beslissingen en acties te sturen. Aan de andere kant maakt het gebruik van impliciete kennis en competentie om af te kunnen stemmen op de werkelijkheid.

### **3.2 Flexibel leersysteem als betrouwbaarder alternatief voor het SIM**

Railton beoogt leer- en representatieprocessen te ontdekken die ons in staat stellen om duidelijk te zien waar onze intuïties naar verwachting betrouwbaarder of minder betrouwbaar zijn.<sup>60</sup> Intuïties zijn gezaghebbend omdat we goede empirische redenen hebben om te geloven dat ze ons op betrouwbare wijze in overeenstemming brengen met de morele realiteit en de betrouwbaarheid verbetert naarmate het leren vooruitgang boekt.<sup>61</sup> Ik zal nu beargumenteren dat dit deels een betrouwbaarder alternatief voor het SIM is.

Ten eerste hebben we gezien dat het SIM het bestaan van onbetrouwbare heuristieken moet erkennen. Dit wordt versterkt door het feit dat intuïties niet bereikbaar zijn voor reflectie: er kan geen onderscheid gemaakt worden tussen betrouwbare en minder betrouwbare heuristieken. De onderzoeken die Railton aanhaalt lijken echter een meer complex en genuanceerd beeld te schetsen van het affectieve systeem. Het idee van een imperfect systeem dat gebruik maakt van imperfecte heuristieken lijkt daarmee achterhaald. Het flexibele leersysteem wijst het hele bestaan van heuristieken af: er wordt geen gebruik gemaakt van vuistregels en heuristieken.<sup>62</sup> Een reden om aan het gebruik van heuristieken te twijfelen is het argument dat intuïties impliciete kennis en competenties omvatten. Dit betekent dat intuïties tot op zekere hoogte als intelligent gezien kunnen worden en bovendien zijn ze in staat om individuele gevallen te beoordelen. Uit het argument van impliciete kennis volgt dat het niet zo is dat intuïtieve proces vooral gevoelsmatig is en vooraf gaat aan het bewuste rationale proces dat vooral cognitief en verstandelijk is. Intuïties zijn onderdeel van cognitieve capaciteiten: ze maken niet alleen gebruik van impliciete kennis en capaciteiten, het zijn er de belichaming van.

Ten tweede hebben we hierop aansluitend gezien dat het SIM impliceert dat morele intuïties over het algemeen weerbarstig en onbuigzaam zijn. Natuurlijk erkent Haidt de mogelijkheid van het aanpassen van intuïties door middel van het beredeneerde oordeel en persoonlijke reflectie. Als we ervanuit gaan deze processen betrouwbaar(der) zijn, dan zou een meer rationale benadering dat stelt dat de rol van deze processen groter is dan het SIM erkent, een

---

<sup>60</sup> Railton, "The Affective Dog," 846.

<sup>61</sup> Railton, "The Affective Dog," 847.

<sup>62</sup> Ik bedoel hiermee heuristieken in de zin van vuistregels. Ik ga hier niet in op andere betekenissen van heuristiek.

meer betrouwbare benadering bieden. Er zou dan meer ruimte zijn voor buigzame intuïties die kunnen veranderen wanneer nodig. Echter biedt dit een oplossing binnen het duale procesmodel van cognitie waarvan we hebben gezien dat het problemen oplevert.<sup>63</sup> Bovendien stel ik in deze scriptie niet de vraag of rationele benaderingen zorgen voor meer betrouwbare morele intuïties dan het SIM, maar ik beargumenteer dat het flexibele leersysteem een meer betrouwbare optie is dan het SIM. Het flexibele leersysteem erkent het bestaan van en vertrouwen in een alternatief affectief systeem. Het is een flexibel leersysteem dat ontworpen is om ons denken en handelen op flexibele wijze te informeren en af te stemmen op de werkelijkheid. De systemen worden continu bijgewerkt en gecorrigeerd doordat ze continu de verwachte waarde vergelijken en bijwerken. Het probleem van weerbarstige en onbuigzame intuïties lijkt er dus niet te zijn: het flexibele leersysteem brengt ons op betrouwbare in overeenstemming met de morele realiteit en de betrouwbaarheid verbetert naarmate het leren vooruitgang boekt.<sup>64</sup>

Tot slot hebben het probleem gezien dat intuïties niet constant blijken te zijn in situaties waar dat wel te verwachten is of wel het geval zou moeten zijn. Daarnaast blijken intuïties bevooroordeeld te zijn op allerlei terreinen. Vormt dit ook een probleem voor het flexibele leersysteem? Het lijkt erop dat dit deels het geval is. Enerzijds werkt het flexibele leersysteem op evaluatieve wijze en brengt ons in overeenstemming met de morele realiteit. Bovendien stelt Railton dat het verschilstelsel goed is in het verminderen van de effecten van aanvallende vertekeningen door ervaring.<sup>65</sup> Anderzijds blijkt dit minder het geval als de ervaringen zelf op dezelfde manier bevooroordeeld.<sup>66</sup> Het lijkt erop dat ook het flexibele leersysteem moeite heeft met vertekeningen en framing. In de volgende paragraaf ga ik verder op de beperkingen van het flexibele leersysteem.

Een geschikter en meer plausibel systeem dat het gebruik maakt van impliciete kennis vereist dat het hele systeem van het SIM herzien moet worden. We hebben in paragraaf 2.3 gezien dat het SIM uitsluit dat in het intuïtieve proces naast de emotionele respons er ook sprake kan zijn van cognitieve informatie en kennis dat meer impliciet en stilzwijgend is. Het affectieve systeem als flexibel leersysteem wordt daarentegen niet gezien als tegenstrijdig aan of verstorend voor cognitieve, evaluatieve, berekende en beslissingscapaciteiten, maar als veeleer een integraal onderdeel ervan. Er is dus geen sprake van een voorafgaande automatische intuïtie die vooral affectief is, maar intuïties zijn de belichaming van cognitieve capaci-

---

<sup>63</sup> Voor meer argumenten tegen het duale proces model van cognitie, zie Railton, "The Affective Dog," 826-833.

<sup>64</sup> Railton, "The Affective Dog," 847.

<sup>65</sup> Railton, "The Affective Dog," 845-846.

<sup>66</sup> Railton, "The Affective Dog," 845-846.

teiten. Dit betekent ook dat de rationaliteit niet met bewustwording wordt geassocieerd: intuïtie en cognitie zijn geïntegreerd.

### 3.3 Beperkingen van het flexibele leersysteem

Betrouwbare intuïties hebben een betrouwbare infrastructuur nodig. Dat wil zeggen: ze komen voort uit een betrouwbaar systeem. Het affectieve systeem functioneert als zowel een statistisch leersysteem als een regulatief systeem. Dit heeft zijn sterke en zwakke punten.<sup>67</sup> De sterke punten hebben we hierboven gezien. Het affectieve leersysteem en intuïtieve denken is sterk maar ook gelimiteerd van betreft zijn capaciteiten. In hun experimenten zoeken Amitai Shenhav en Joshua Greene bijvoorbeeld de verklaring hiervoor net als Railton in evolutie. Zij suggereren dat de evolutionaire logica van spontane morele beoordelingen soms gebruikt wordt in zaken waar die logica helemaal niet toepasselijk is.<sup>68</sup> Zo bezien blijken ons statistische leersysteem en empathische vermogens, net als perceptie, niet perfect te zijn. Ze zouden (deels) onderworpen zijn aan capaciteitslimieten: ze zijn bijvoorbeeld niet-representatief en bevooroordeeld. Als dit klopt dat zal onze statistische empathische reactie dat ook zijn.<sup>69</sup> Dit lijkt een groot probleem te zijn: de kwaliteit van de intelligentie en kennis zal de beperkingen in de capaciteiten van het systeem weerspiegelen net als de ervaringen die het ontvangt.<sup>70</sup>

We moeten dus toch wel voorzichtig zijn. Het is volgens Railton van het allergrootste belang voor epistemologie en moraliteit dat we wezens zijn die kritisch onze intuïtieve antwoorden kritisch kunnen bekijken, hoewel geloofsvorming, beoordeling en toepassing onlosmakelijk vermengd zijn met intuïties. Railton wil niet onkritische acceptatie van intuïties aanmoedigen, maar hij betoogt dat het flexibele leersysteem ons in staat stelt duidelijker te zien waar onze intuïties naar verwachting betrouwbaarder zouden zijn en waar dat naar verwachting minder het geval is.<sup>71</sup>

---

<sup>67</sup> Railton, "The Affective Dog," 845.

<sup>68</sup> Railton, "The Affective Dog," 845.

<sup>69</sup> Railton, "The Affective Dog," 845.

<sup>70</sup> Railton, "The Affective Dog," 847.

<sup>71</sup> Railton, "The Affective Dog," 846.

## 4. CONCLUSIE

Ik heb in deze scriptie onderzocht of intuïties die gegenereerd worden door een flexibel leersysteem betrouwbaardere intuïties zijn dan intuïties die worden gegenereerd door het SIM. De betrouwbaarheid van deze laatste intuïties kan niet worden gegarandeerd omdat de intuïties weerbarstig zijn, gevoelig zijn voor framing en gebruik maken van imperfecte heuristieken. Fundamentele herziening van het SIM biedt de mogelijkheid om de betrouwbaarheid van meer intelligente intuïties te onderzoeken, meer specifiek intuïties die onderdeel zijn van een flexibel leersysteem. Dit leersysteem is beter in staat om te laten zien waar onze intuïties naar verwachting betrouwbaarder zijn en waar minder. Het flexibele leersysteem is in staat ons denken en handelen op flexibele wijze te informeren en af te stemmen op de werkelijkheid. Enerzijds maakt het systeem gebruik van impliciete kennis en competenties om morele oordelen af te stemmen op de werkelijkheid, en ontkent hiermee het gebruik van heuristieken. Anderzijds is het systeem op evaluatieve wijze in staat intuïties bij te werken. De systemen worden continu bijgewerkt en gecorrigeerd doordat ze continu de verwachte waarde vergelijken en bijwerken. Intuïties zijn dus niet weerbarstig maar flexibel.

Toch moeten we voorzichtig zijn met vertekeningen en daarom kunnen intuïties niet onkritisch aangenomen worden. Hoewel ik concludeer dat het flexibele leersysteem een betrouwbaarder alternatief is voor het SIM, hebben beide systemen capaciteitslimieten. Dat het leersysteem meer betrouwbaar is betekent niet dat we er helemaal en altijd op kunnen vertrouwen. Het flexibele leersysteem blijkt gevoelig voor vertekeningen in de ervaring. Niet alle morele intuïties zijn betrouwbaar en kunnen zonder kritische blik aangenomen. Toch is het een beter alternatief dan het SIM. Haidt claimt dat hij redenering, emotie, intuïtie en sociale beïnvloeding volledig integreert, maar dat blijkt niet het geval te zijn. Zijn aannames met betrekking tot morele redenering en intuïtie sluiten uit dat causale redeneving iets anders kan zijn dan een niet-automatisch en toegankelijk proces. Railton blijkt wel in staat om intuïties en cognitie te integreren: praktische intelligentie kan worden gezien als belichaamd in intuïtie.

Een belangrijk verschil tussen Haidt en Railton is dat ze van andere empirische data gebruik maken en daar verschillende conclusies uit trekken. Idealiter zou een oplossing voor de betrouwbaarheid van het SIM geïntegreerd kunnen worden met het SIM of met de data waar het van gebruik maakt. Toch denk ik er ook een verschil is weggelegd voor de interpretatie van de data. Veel van de onderzoeken die Haidt gebruikt zouden wel passen in het leersysteem van Railton maar andersom is dat niet het geval. Het flexibele leersysteem is een

breed cognitief model dat niet gezien moet worden als tegenstrijdig aan cognitieve, evaluatieve, berekende en beslissingscapaciteiten, maar als een integraal onderdeel ervan. Geloofsvorming, beoordeling en toepassing zijn onlosmakelijk vermengd met intuïties. Het negeren van morele intuïties lijkt dus niet mogelijk.

## REFERENTIES

Alfano, Mark. *Moral Psychology: An Introduction*. Cambridge: Polity Press, 2016.

Brink, David O. "Principles and Intuitions in Ethics: Historical and Contemporary Perspectives." *Ethics* 124, no. 4 (July 2014): 665–694. <https://doi.org/10.1086/675878>.

Buckwalter, Wesley, and Stephen Stich. "Gender and Philosophical Intuition." In *Experimental Philosophy*, volume 2, edited by J. Knobe and S. Nichols, 307-346. Oxford: Clarendon, 2013.

Clarke, Steve. "SIM and the city: Rationalism in psychology and philosophy and Haidt's account of moral judgment." *Philosophical Psychology* 21, no. 6 (2008): 799–820. <https://doi.org/10.1080/09515080802513250>.

Doris, John, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley. "Moral Psychology: Empirical Approaches." In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/win2017/entries/moral-psych-emp/> (geraadpleegd 6 juni 2018).

Gjesdal, Adam. "Moral Learning, Rationality, and the Unreliability of Affect." *Australasian Journal of Philosophy* (2017): 1-14. <https://doi.org/10.1080/00048402.2017.1366535>.

Haidt, Jonathan. "The Emotional Dog and its Rational Tail." *Psychological Review* 108, no. 4 (2001): 814-834. <https://doi.org/10.1037/0033-295X.108.4.814>.

Horgan, Terry, and Mark Timmons. "Morphological Rationalism and the Psychology of Moral Judgment." *Ethical Theory and Moral Practice* 10 (2017): 279-295. <https://doi.org/10.1007/s10677-007-9068-4>.

Jacobson, Daniel. "Moral Dumbfounding and Moral Stupefaction." In *Oxford Studies in Normative Ethics*, volume 2, edited by Mark Timmons, 289-316. Oxford: Oxford University Press, 2012.

Jeffery, Renée. *Reason and Emotion in International Ethics*. Cambridge University Press, 2014.

Kahneman, Daniel. "A perspective on judgment and choice: Mapping bounded rationality." *American Psychologist* 58 no. 9 (September 2003): 697 –720. <https://doi.org/10.1037/0003-066X.58.9.697>.

Prinz, Jesse J., and Shaun Nichols. "Moral Emotions." In *The Moral Psychology Handbook*, edited by John M. Doris, 111-147. New York: Oxford University Press, 2010. <https://doi.org/10.1093/acprof:oso/9780199582143.003.0005>.

Railton, Peter. "The Affective Dog and Its Rational Tale: Intuition and Attunement." In *Ethics* 124, no. 4 (July 2014): 813-859. <https://doi.org/10.1086/675876>.

Sinnott-Armstrong, Walter, Liane Young, and Fiery Cushman. "Moral Intuitions." In *The Moral Psychology Handbook*, edited by John M. Doris, 246-272. New York: Oxford University Press, 2010. <https://doi.org/10.1093/acprof:oso/9780199582143.003.0008>.

Sunstein, Cass, "Moral Heuristics," *Behavioral and Brain Sciences* 28 (2005): 531–73. <https://doi.org/10.1017/S0140525X05400096>.