

Theory of mind en hoe wij anderen begrijpen

Student: Linda Lamers

Studentnummer: S4687477

Naam begeleider: dr. Leon de Bruin

Aantal woorden (excl. bibliografie): 14.176

Datum: 31 juli 2018

Scriptie ter verkrijging van de graad “Master of arts” in de filosofie

Radboud Universiteit Nijmegen

Hierbij verklaar en verzeker ik, Linda Lamers, dat deze scriptie zelfstandig door mij is opgesteld, dat geen andere bronnen en hulpmiddelen zijn gebruikt dan die door mij zijn vermeld en dat de passages in het werk waarvan de woordelijke inhoud of betekenis uit andere werken – ook elektronische media – is genomen door bronvermelding als ontlening kenbaar gemaakt worden.

Plaats: Nijmegen datum: 31 juli 2018

Samenvatting

In het debat over hoe het kan dat we de ander in alledaagse situaties begrijpen, keert Theorie-Theorie steeds terug als een schijnbaar onvermijdelijke verklaring. Ondanks de logische problemen waar Theorie Theorie tegen aanloopt, passen we in alledaagse sociale interactie vaak theorieën toe om verklaringen voor het gedrag van de ander te vinden. In mijn scriptie heb ik de dominante theorieën Theorie Theorie en Simulatie Theorie geanalyseerd om te zien op welke manier ze een verklaring geven over hoe we anderen begrijpen. Daarnaast heb ik een recente theorie, de Interactie Theorie, tegenover de Theorie Theorie en Simulatie Theorie uiteengezet en daarmee aangegeven dat er nog veel kritische punten blijven bestaan bij de Interactie theorie. Theoretisering blijft een belangrijke rol blijft spelen bij alledaagse interactie.

Inhoudsopgave

SAMENVATTING.....	2
1. INLEIDING	4
2. THEORIE THEORIE ALS THEORIE VAN MINDREADING	6
2.1 MODULAIRE THEORIE THEORIE.....	8
2.2 DE KIND-ALS-WETENSCHAPPER THEORIE THEORIE.....	9
2.3 FALSE-BELIEF TEST ALS BEWIJS VOOR TT	12
2.4 ANDER EMPIRISCH ONDERZOEK.....	14
2.5 PROBLEMEN MET DE TT	16
2.5.1 <i>Eliminativisme tegenover TT</i>	18
3. SIMULATIE THEORIE ALS THEORIE VAN MINDREADING	21
3.1 SIMULATIE THEORIE	21
3.2 IMPLICIETE EN EXPLICIETE ST	22
3.3 SPIEGELNEURONEN KRITIEK	24
3.4 PROBLEMEN MET DE ST	26
4. TT EN ST	29
4.1 HET FRAME-PROBLEEM EN DE TT EN ST	29
4.2 HYBRIDE THEORIE TT/ST	30
5. INTERACTIE THEORIE ALS ALTERNATIEF	33
5.1 IT ALS THEORIE VOOR HET BEGRIJPEN VAN DE ANDER	33
5.2 IT BIJ BABY'S EN KINDEREN.....	34
5.3 NARRATIVE PRACTICE HYPOTHESIS.....	35
5.4 IT ALS REACTIE OP ST EN TT	36
6. KRITIEK OP IT	39
7. CONCLUSIE	42
8. BIBLIOGRAFIE	44

1. Inleiding

In het dierenrijk bestaan er tal van sociale dieren. Mieren, bijen en wespen staan bekend om sociaal gedrag, ook wel eusocialiteit genoemd. Eusociaal gedrag wordt beschreven als het leven in een groep van verschillende generaties, waarbij er coöperatief voor elkaars kroos wordt gezorgd, maar ook het reproductief verdelen van arbeid. De insecten voorzien zichzelf van directe reproductie in de functie waarin zij zich bevinden, dus als werkers, beveiligers of als verzorgers van de nakomelingen van de koningin.¹ De kolonies hebben sociale hiërarchieën en complexe vormen van communicatie.

Homo sapiens zijn een bijzondere sociale soort; een opvallende eigenschap bij deze soort is de vaardigheid om gedachten van de ander te kunnen lezen en te voorspellen. Deze term staat ook bekend als *mindreading*.² Mensen zijn in staat om zichzelf en anderen een reeks van mentale toestanden toe te wijzen, zoals overtuigingen, verlangens en verwachtingen, maar ook emoties zoals boosheid of blijheid.³

Hoe kunnen we anderen begrijpen? Hoe is het mogelijk om de gedachten van de ander te kunnen voorspellen? Hoe weet ik wat de ander denkt? Welke mechanismen zorgen ervoor dat ik de ander kan begrijpen? Dit zijn belangrijke vragen in het filosofisch debat over de Theory of Mind (ToM). Twee van de belangrijkste en dominantste theorieën binnen de cognitiefilosofie zijn de Theorie Theorie (TT) en de Simulatie Theorie (ST). De TT stelt dat we overtuigingen of verlangens van anderen kunnen afleiden door het toepassen van een psychologische theorie. De later ontwikkelde ST daarentegen stelt dat we geen theorie hoeven toe te passen om de ander te kunnen begrijpen. In plaats daarvan, simuleren we wat de ander zou doen, we verplaatsen onszelf in de ‘schoenen van de ander’. Meer recent is een andere theorie, de Interactie Theorie (IT), ontwikkeld die TT en ST niet nodig vindt om elkaar in het alledaagse leven te begrijpen. Deze theorie suggereert dat we anderen begrijpen door middel van directe perceptie, we zien gelijk wat de ander bedoelt, zonder een onderliggende theorie of simulatie.

¹ Bryan, N. Danforth, “Social insects: are ants just wingless bees?” *Current Biology* 23:22 (2013): 1011.

² De Engelse term *mindreading* wordt gebruikt, omdat er mijns inziens geen Nederlandse variant van deze term bestaat die de filosofische lading dekt.

³ Alvin I. Goldman: *Simulating minds: the philosophy, psychology, and neuroscience of mindreading*, (Oxford: Oxford University Press, 2006), 3.

In mijn scriptie wil ik aantonen dat alhoewel TT op de achtergrond is komen te staan door kritische argumenten, het als verklaring blijft terugkeren over hoe we anderen kunnen begrijpen. Hiermee is de TT een aannemelijke theorie over hoe we het gedrag van de ander kunnen voorspellen, verklaren en begrijpen.

2. Theorie Theorie als theorie van mindreading

Sociale cognitie gaat over alle processen waardoor individuen van het zelfde soort met elkaar kunnen interacteren. Die interactie is belangrijk voor het overleven van individuen maar ook voor de gehele soort. Aan de hand van gezichtsuitdrukkingen en lichaamshoudingen kunnen we vaak zien wat iemand voelt. Door het volgen van iemands ogen zien we waar hij naar staart en weten we waar zijn aandacht is op gefocust. Maar ook door het kijken naar lichaamsbewegingen kunnen we vaak bepalen wat diegene wil doen.⁴

Mindreading kan worden gezien als een vorm van sociale cognitie. De vaardigheid om mindreading toe te kunnen passen gaat over het toeschrijven van mentale toestanden aan een individu om zijn gedrag te kunnen begrijpen en te anticiperen op zijn gedrag in de toekomst.⁵ Bijvoorbeeld, het regent buiten en mijn buurman wil naar buiten gaan om boodschappen te doen. Hoe weet ik wat hij gaat doen? Met behulp van een psychologische theorie ga ik er van uit dat hij niet nat wil worden, dus schrijf ik hem de mentale toestand toe dat hij het *verlangen* heeft om niet nat te willen worden. Omdat ik hem het verlangen ‘niet nat willen worden’ toeschrijf, kan ik hieruit concluderen dat als hij er voor kiest om naar buiten te gaan, hij hoogstwaarschijnlijk een paraplu meeneemt. Hier wordt de volkspychologie benadert vanuit een theorie. Het niet nat willen is ook wel een stereotypische theorie die voor mensen in het algemeen geldt, deze stereotypische theorie wordt aangeleerd van jongs af aan. Een manier om de ander te kunnen begrijpen, kan via een volkspychologische theorie waardoor we in staat zijn om verschillende overtuigingen, zoals verlangens en andere mentale toestanden aan anderen toe te schrijven. Mindreading staat ook wel bekend als volkspychologie, en de TT kan een manier zijn om te verklaren hoe volkspychologie via een theorie werkt. Er zijn veel versies van de TT die op verschillende manieren geïnterpreteerd kan worden. De meest bekende TT versies gaan over dat we aangeleerde of aangeboren theorieën in ons brein hebben, waardoor we het gedrag van de ander kunnen begrijpen of voorspellen. Conceptuele veranderingen bevorderen de ontwikkeling van de mind van kinderen, zodat zij er steeds beter in

⁴ Chris D. Frith en Uta Frith, “Social cognition in humans,” *Current Biology* 17 (2007): 724.

⁵ Shannon Spaulding, “Mirror neurons are not evidence for the Simulation Theory,” *Synthese* 189 (2012): 515.

worden om mentale toestanden van anderen te begrijpen.⁶ In mijn scriptie zal ik deze twee versies van de TT verder uitwerken.

Om de TT vanuit een meer abstracter betekenis te bekijken, is het belangrijk om te zien waar theorieën en theoretische concepten in het algemeen voor worden gebruikt en waarom het ons kan helpen om bepaalde dingen te verklaren. Theoretische concepten zijn gepostuleerde abstracte entiteiten, die ons kunnen helpen om een causale verklarende analyse te geven voor het bestaan van bepaalde fenomenen. Neem het fenomeen zwaartekracht. Om te bepalen wat zwaartekracht is, kan het beste naar de functie ervan worden gekeken. Zwaartekracht kan de beweging van lichamen ten opzichte van elkaar uitleggen, zoals waarom een appel van de boom naar beneden valt, maar kan niet worden gedefinieerd als twee lichamen die zich bewegen in verhouding tot de ander. Zwaartekracht is juist gepostuleerd om zulke fenomenen, zoals de beweging van lichamen ten opzichte van elkaar, uit te kunnen leggen. Door het toepassen van een theorie is het ook mogelijk om het gedrag te kunnen voorspellen van hemellichamen, zoals de maan of de zon.⁷

Daarnaast is het geen vereiste dat theoretische concepten geobserveerd moeten worden. In de biologie bestaan er theorieën over entiteiten, zoals virussen en bacteriën die we niet met het blote oog kunnen zien. Door biologische theorieën weten we bijvoorbeeld dat we met antibiotica de meeste bacteriële infecties kunnen bestrijden.⁸ Mensen kunnen het gedrag van anderen begrijpen door een algemene theorie over de mind toe te passen, zoals ik eerder in het voorbeeld verklaarde dat het niet nat willen worden door regen als een algemene theorie gezien kan worden. Volgens de TT vloeit het toewijzen van bepaalde mentale toestanden aan anderen voort uit een theoretische beredenering waarbij de bekende causale wetten vanzelfsprekend zijn.⁹ Een causale wet die bijvoorbeeld vanzelfsprekend lijkt te zijn is dat mensen niet nat willen worden, en dus als zij naar buiten gaan, ze een paraplu of andere bescherming voor de regen meenemen. Sinds de jaren tachtig ontstonden er theorieën die als bewijs of argument

⁶ Caitlin E.V. Mahy, Louis J. Moses en Jennifer H. Pfeifer, "How and where: Theory-of mind in the brain," *Developmental Cognitive Neuroscience* 9 (2014): 74.

⁷ Alison Gopnik en Henry M. Wellman, "Why the child's Theory of Mind really is a theory," *Mind & Language* 7:1&2 (1992): 146.

⁸ Gopnik en Wellman, "Child's Theory of Mind," 147.

⁹ Vittorio Gallese en Alvin Goldman, "Mirror neurons and the simulation theory of mind-reading," *Trends in Cognitive Science* 2:12 (1998): 496.

dienen voor het bestaan van een TT. Ik zal hieronder twee theorieën uitwerken die nogal onder de aandacht zijn gebracht in de literatuur om te illustreren hoe TT werkt.

2.1 Modulaire Theorie Theorie

De modulaire theorie beweert dat een ToM bestaat uit theoretische modulen die aangeboren, ingekapseld en een domein-specifiek deel van de cognitieve architectuur zijn. Leslie en Scholl¹⁰ stellen bijvoorbeeld dat het voor het verklaren van ToM essentieel is om de architectuur van de cognitieve modules die ToM realiseren te ontwaren. De essentie achter de architectuur van modulen is een reeks van beperkingen op de stroom van informatie. Deze beperkingen functioneren als zogenaamde ‘informatiefilters’, wat betekent dat bepaalde informatie in de module niet toegankelijk is buiten de module, of dat bepaalde informatie buiten de module juist niet toegankelijk is binnen de module. Een voorbeeld is de Müller-Lyer illusie, waarbij het lijkt alsof de ene lijn groter is dan de andere lijn. Ondanks dat je de lijnen kan opmeten met een liniaal om te zien dat ze even lang zijn, blijft je waarneming hetzelfde en blijf je zien dat ze verschillende lengten hebben. Dit kan worden verklaard door te stellen dat visuele informatie ten dele is ingekapseld, het proces die de module gebruikt om je waarnemingen samen te stellen is afgeschermd van je overtuigingen over deze waarnemingen. Een cognitieve functie is architectonisch modulair tot het niveau dat informatie op die manier is ingekapseld.¹¹

Een belangrijke aanname binnen de modulaire theorie is dat de basis van ToM aangeboren moet zijn. Deze aangeboren basis maakt deel uit van onze genetische achtergrond.¹² De ToM wordt volgens de modulaire theorie gezien als een vaardigheid die is ontstaan vanuit een aangeboren, ingekapseld en domein specifiek deel van de cognitieve architectuur. Vanuit evolutionaire biologie moet het domein specifiek zijn en kan het niet domein algemeen zijn, omdat het de overlevingskans van de mens en primaten kleiner maakt. Om een beter begrip te hebben van domein specifieke cognitieve mechanismen, gebruiken Cosmides en Tooby¹³ het voorbeeld van

¹⁰ Brian J. Scholl en Alan M. Leslie, “Modularity, development and ‘Theory of Mind’,” *Mind & Language* 14:1 (1999): 131.

¹¹ Scholl en Leslie, “Modularity,” 133.

¹² Ibid., 134.

¹³ Leda Cosmides en John Tooby, *Mapping the mind: Domain specificity in cognition and culture*, red. Lawrence A. Hirschfeld en Susan A. Gelman (Cambridge: Cambridge University Press, 1994), 89.

meerkatten, die vaak in onderzoeken worden gebruikt omdat zij dezelfde genetische voorouders als mensen hebben. Meerkatten kunnen verschillende waarschuwingskreten uitslaan als zij bedreigd worden door roofdieren zoals luipaarden, slangen en adelaars. De verschillende waarschuwingskreten zijn belangrijk voor het overleven van de meerkattensoort. De luipaard, slang en adelaar vormen de drie belangrijkste roofdieren voor meerkatten. Bij bedreiging van een luipaard moet de meerkat in een boom kunnen klimmen, bij een adelaar moet hij omhoog kunnen kijken en zichzelf in het gras kunnen verstoppen en bij slangen kan de meerkat op zijn achterste poten staan om goed in het gras te kunnen kijken. Meerkatten hebben domein specifieke cognitieve mechanismen ontwikkeld die reageren op verschillende waarschuwingskreten van hun soortgenoten voor de drie verschillende roofdieren. Als er één algemene waarschuwingskreet bestaat dan zou het niet erg effectief zijn, want dan zouden de meerkatten die de kreet horen van hun soortgenoten niet weten welke roofdier er in de buurt is en hoe ze daarop moeten reageren, dat zou de overlevingskans van meerkatten aanzienlijk verminderen.

Onze voorouders in het Pleistoceen tijdperk¹⁴ moesten goed zijn in het oplossen van problemen om zichzelf te kunnen voorplanten en te overleven, zoals het zoeken van de juiste partner, omgaan met bedreigingen, het ontwijken van roofdieren en het voorkomen van incest. Deze problemen kunnen niet door hetzelfde mechanisme worden opgelost, en daardoor kan volgens Cosmides en Tooby de menselijke mind bestaan uit verschillende functionele cognitieve aangepaste specialisaties. Op empirisch en theoretisch niveau kunnen twee cognitieve mechanismen niet hetzelfde zijn, net zoals dat twee lichamelijke organen niet hetzelfde kunnen zijn.

Een argument tegen de modulaire theorie is de kind-als-wetenschapper theorie van Gopnik. Cognitieve modulen lijken statisch te zijn en ze ontwikkelen zich niet verder. De kind-als-wetenschapper theorie daarentegen, richt zich juist op de rol van theoretische ontwikkeling.

2.2 De Kind-als-wetenschapper Theorie Theorie

Gopnik¹⁵ ontwikkelde de kind-als-wetenschapper theorie om te verklaren hoe een ToM wordt verworven bij jonge kinderen. Ze maakt hierbij een vergelijking tussen de

¹⁴ Cosmides en Tooby, *Mapping the mind*, 90.

¹⁵ Alison Gopnik, "Developing the idea of intentionality: children's Theories of Mind," *Canadian Journal of Philosophy* 20:1 (1990): 105.

ontwikkeling van een wetenschappelijke theorie en het concept van intentionaliteit of overtuiging als onderdeel van een theorie die kinderen ontwikkelen. Een kind loopt door een vijf stappenplan¹⁶ als er een verandering van een theorie plaatsvindt. In de eerste fase verzamelt het kind tegenbewijs voor de huidige theorie, in de tweede fase ontkent hij het tegenbewijs voor de huidige theorie, in de derde fase maakt hij ad hoc hypothesen, in de vierde fase formuleert hij een nieuwe theorie en in de laatste fase observeert en experimenteert hij met alternatieve theorieën. Gopnik¹⁷ en haar collega's ontwikkelde een test om te zien of jonge kinderen fysieke causale relaties kunnen infereren vanuit statistisch bewijs over correlaties. De kinderen krijgen een doosje te zien die muziek afspeelt als er bepaalde blokken op staan, dus niet bij ieder willekeurige blok. Gopnik liet verschillende patronen zien afhankelijk van de statistische gegevens tussen de blokken en het effect ervan. Vervolgens werd er aan de kinderen gevraagd om het muziekdoosje aan en uit te zetten. Jonge kinderen van twee, drie en vier jaar oud waren in staat om een patroon bij de correlatie tussen de blokken en het aanzetten van het muziekdoosje te vinden door te infereren welke van de causale structuren correct was, zodat ze zelf, door te experimenteren met de blokken, er achter kwamen hoe het doosje aan en uit gaat.

Een wetenschapper doorloopt volgens Gopnik precies dezelfde stappen bij een nieuwe of verandering van een huidige theorie. Deze analogie lijkt volgens Bishop en Downes niet helemaal te kloppen, want de vijf stappen die kinderen doorlopen is niet iets wat een typische wetenschapper doorloopt. Wetenschappers die nieuwe theorieën ontwikkelen zijn vaak nieuw in hun vakgebied, dus het is niet gebruikelijk dat ze eerst bewijs gaan verzamelen om tegen een bestaande theorie in te gaan.¹⁸ Daarnaast passen wetenschappers theorieën toe die ze van hun voorgangers hebben geleerd, en de meerderheid van de wetenschappers ontdekken geen nieuwe theorieën.

De vergelijking tussen een kind en wetenschapper wordt erg letterlijk genomen. Is het gerechtvaardigd om de analogie te interpreteren die op deze manier wordt gemaakt? Ik zal hieronder laten zien waarom de analogie breder geïnterpreteerd moet

¹⁶ Michael A. Bishop en Stephen M. Downes, "The theory theory thrice over: the child as scientist, superscientist or social institution?" *Stud. Hist. Phil. Sci.* 33 (2002): 124.

¹⁷ Alison Gopnik, "Scientific thinking in young children: theoretical advances, empirical research, and policy implications," *Science* 337 (2012): 1625.

¹⁸ Bishop en Downes. "The child as scientist," 125.

worden. Wetenschappelijke verklaringen¹⁹ zijn vergelijkbaar met alledaagse verklaringen omdat ze een theoretisch raamwerk voor een fenomeen geven, verder kijken dan het originele fenomeen, een reeks aan fenomenen integreren, laten zien hoe het originele fenomeen vanuit een raamwerk volgt en daar een bepaald gevoel of begrip bij geven. Hierbij valt te denken aan het eerder genoemde fenomeen zwaartekracht. Een ander belangrijk kenmerk van wetenschappelijke verklaringen is dat ze getest kunnen worden. Creationisme is bijvoorbeeld niet wetenschappelijk omdat het niet getest kan worden.

Kinderen zijn in staat om verklaringen voor de fysische wereld te geven, en de structuur waarmee ze dat doen lijkt op dezelfde structuur als die van een wetenschapper.²⁰ Er is een onderzoek²¹ gedaan waarbij er aan kinderen uit verschillende klassen is gevraagd waar de zon zich bevindt als het nacht is. Tijdens het onderzoek geven, naarmate de kinderen ouder worden, een steeds betere verklaring op de vraag. Een meisje uit groep drie antwoordde op de vraag dat de zon verdwijnt in de ruimte. Als er door wordt gevraagd zegt ze dat dit plaatsvindt op het moment wanneer het donker wordt, de zon dan in de ruimte verdwijnt.

Als dezelfde vraag gesteld wordt aan een iets ouder meisje uit groep 5, waarbij er een uitgebreider antwoord wordt gegeven. Het meisje zegt dat de zon verdwijnt naar China, de onderkant van de aarde, want als het bij ons dag is, is het bij China nacht en andersom zo. Als de onderzoeker vraagt hoe dit kan zegt het meisje dat de maan en zon iedere twaalf uur van plaats veranderen. Bij deze voorbeelden zijn kinderen in staat om een volledige verklaring te geven voor fenomenen. Het kind is daarnaast ook in staat om het fenomeen in een conceptueel raamwerk te plaatsen. Als dezelfde vraag aan een wetenschapper wordt gesteld beantwoordt hij de vraag met een andere inhoud, maar de manier waarop een kind de vraag probeert te verklaren is vergelijkbaar met de manier waarop de wetenschapper de vraag beantwoordt.²² Hoe ouder een kind wordt, hoe beter hij of zij in staat is om vragen over de wereld en anderen te beantwoorden, omdat hij steeds meer theorieën leert over de wereld maar ook over gedragingen van anderen.

¹⁹ William F. Brewer, Clark A. Chinn en Ala Samarapungavan, "Explanation in scientists and childrenen," *Minds and Machines* 8 (1998): 121.

²⁰ Brewer et al. "Explanation in scientists," 128.

²¹ Ibid., 127.

²² Ibid., 134.

In de ontwikkelingspsychologie wordt er vaak gesuggereerd dat TT een goede theorie is die wordt gestaafd door bestaande onderzoeken. Er bestaat bewijs dat kinderen beter in staat zijn de mentale toestanden van anderen te begrijpen doordat de situaties toenemen waarin zij zich bevinden zodat zij steeds meer ervaringen krijgen bij interactie met de ander.²³ Ik zal later terugkomen op empirische onderzoeken die als bewijs kunnen dienen voor de kind-als-wetenschapper theorie.

2.3 False-belief test als bewijs voor TT

Een ander argument wat gezien wordt als empirisch bewijs voor een ToM is de *false-belief* test ontwikkeld door o.a. Baron-Cohen, Wimmer en Perner.²⁴ Het doel van de test is om te laten zien op welk moment kinderen voor het eerst kunnen begrijpen dat de ander een verkeerde overtuiging heeft. De test laat zien dat kinderen onder de vier jaar nog geen begrip hebben van een verkeerde overtuiging. Voor sociale cognitie vormt het een belangrijk punt op het moment dat kinderen in staat zijn om te begrijpen dat een ander een verkeerde overtuiging kan hebben. In de klassieke false-belief test horen kinderen een verhaaltje bij een toneelspel. De pop Sally verstoppt haar steen in een doos en verlaat de ruimte, vervolgens komt de pop Anne en ze verplaatst de steen in een andere doos. Sally komt terug en aan de kinderen wordt gevraagd waar Sally denkt dat haar steen ligt. De meeste 3-jarigen voorspellen dat Sally in de doos zal kijken waar Anne hem in heeft gelegd, dit laat zien dat de kinderen niet het vermogen hebben om te begrijpen dat Sally een verkeerde overtuiging kan hebben van de plaats van de steen. De test is in verschillende culturen en landen gebruikt. Onderzoekers concluderen hieruit dat een belangrijke verandering plaatsvindt bij kinderen van vier jaar, waarbij kinderen kunnen begrijpen dat anderen een verkeerde overtuiging kunnen hebben. Toch wordt de false-belief test recent vaker bekritiseerd. Er bestaan onderzoeken²⁵ die concluderen dat een false-belief begrip al aanwezig is bij kinderen onder de twee jaar, en bij peuters tussen de twee en drie jaar. Een ander kritiekpunt²⁶ is de moeilijkheidsgraad van de test. Ook als kinderen een begrip kunnen hebben van een verkeerde overtuiging (false

²³ Mahy, Moses en Pfeifer, "Theory-of mind in the brain," 75.

²⁴ Rose M. Scott en Renée Baillargeon, "Early false-belief understanding," *Trends in Cognitive Sciences* 21:4 (2017): 237.

²⁵ Scott en Baillargeon, "Early false-belief understanding," 238.

²⁶ Paul Bloom en Tim P. German, "Two reasons to abandon the false belief task as a test of theory of mind," *Cognition* 77 (2000): 27.

belief), blijft de test nog steeds moeilijk. De test blijft moeilijk omdat ten eerste kinderen er bewust van moeten zijn dat Sally niet gezien kon hebben toen de steen werd verplaatst, ten tweede de kinderen moeten onthouden waar de steen eerst lag en waar het op het moment van de test ligt en ten derde moeten zij de vraag goed begrijpen, dus waar Sally gaat kijken en niet waar ze zou moeten kijken. Een ander probleem is dat TT aanhangers juist proberen een ToM via onbewuste processen weer te geven, maar de klassieke test lijkt eerder op een expliciet proces. Er zou dus een duidelijk onderscheid gemaakt moeten worden tussen impliciete en expliciete ToM processen, en hoe deze op verschillende manieren getest kunnen worden. Het onderscheid tussen expliciet en impliciet blijft ook onduidelijk of dat de false-belief test voor een modulaire versie van de TT of de kind-als-wetenschapper theorie bewijs kan vormen. Het zou wel als bewijs kunnen vormen voor TT, maar daarmee blijft het lastig te bepalen voor welke versie.

Daarnaast kan er te snel de conclusie uit worden getrokken, dat als kinderen er niet in slagen om de false-belief test correct uit te voeren, zij minder goed over mentale toestanden van anderen kunnen nadenken. De resultaten van de klassieke tekst kunnen daarom niet te letterlijk worden geïnterpreteerd en moeten voorzichtig worden beoordeeld.

Er bestaan ook nog andere methoden om te zien of kinderen in staat zijn om mentale toestanden aan anderen toe te schrijven. Recent bestaat er een andere test²⁷ waarbij 2-jarigen een mooi speeltje observeren die op een hoog schap wordt gezet. Op het moment dat er gevraagd wordt waar het speeltje ligt, gaven kinderen vaker een naam aan het speeltje en gebaarden naar de plaats waar het speeltje lag als een van de ouders niet aanwezig was. Als het speeltje werd verplaatst en een van de ouders was er wel bij, deden ze dat minder. Dit suggereert dat kinderen hun gedrag aanpassen afhankelijk van wat anderen mensen weten, en dat ze een onbewuste waardering hebben voor de omstandigheden waarin overtuigingen worden gevormd. Aan de andere kant betekent het niet dat de false-belief test niet nuttig kan zijn, het kan juist helpen om de redeneringen van kinderen bij verschillende representaties, zoals overtuigingen in het verhaal van Sally en Anne, te analyseren. Daarnaast kan het ook worden gebruikt om cognitieve en linguïstische beperkingen bij oudere kinderen en volwassenen te diagnosticeren.

²⁷ Bloom en German, "Two reasons to abandon the false belief task," 28-9.

2.4 Ander empirisch onderzoek

In het begin gaf ik aan dat het belangrijk is om naar empirische studies te kijken over de ToM, omdat zij mogelijk als bewijs kunnen dienen voor een TT als vorm van mindreading. Er bestaan veel wetenschappelijke studies over welk gedeelte of gedeelten van het brein gerelateerd kan worden aan sociale cognitie. Nadat de false-belief test werd gepubliceerd, zijn er verschillende onderzoeken gedaan waarbij er gebruik werd gemaakt van neuroimaging, om te zien welke gedeelten van het brein een rol spelen bij sociale cognitie. In eerder onderzoek werd er aangetoond dat er drie hersengebieden geactiveerd worden bij een ToM, namelijk de temporele pole, de sulcus temporalis superior en de cortex cingularis anterior.²⁸ In eerste instantie ging men er vanuit dat de eerste twee gebieden, de temporele pole en de sulcus temporalis superior, geactiveerd werden bij het testen van ToM, omdat ToM over verschillende gedragingen in bepaalde situaties en bewegende lichamen gaat, het cingularis gebied werd dan als verantwoordelijke gezien voor ToM.

De relevantie van de prefrontale cortex voor ToM wordt met name onderstreept door onderzoeken waarbij patiënten letsel hadden opgelopen in het prefrontale gebied. Deze patiënten hadden moeite met het toeschrijven van verkeerde overtuigingen van de eerste en tweede orde en met het herkennen van eenvoudig bedrog. Dit kwam vooral voor wanneer linkse prefrontale cortexen waren beschadigd. Recent onderzoek toont juist aan dat er bij patiënten waarbij de cortex cingularis anterior flink is beschadigd geen afwijkende ToM toonden. Beschadigingen in het linker gedeelte van de temporal parietal junction (TPJ) tonen wel specifieke ToM afwijkingen.

Zoals eerder genoemd, worden er bij deze studies gebruik gemaakt van functionele neuroimaging waarbij neurale mechanismen worden geïdentificeerd die onderliggend kunnen zijn aan sociale cognitie. In recente onderzoeken²⁹ worden er vaak twee richtingen gecreëerd. Bij de ene richting wordt er gekeken naar relevante processen die met perspectief te maken hebben op het gebied van visuospatiele cognitie, typisch bij deze richting is dat er gevraagd wordt naar de visuele ervaring die ontstaat bij het bekijken van iets vanuit het perspectief van de ander. De andere richting focust

²⁸ Josef Perner et al., "Thinking of mental and other representations: The roles of left and right temporo-parietal junction," *Social Neuroscience* 1:3-4 (2006): 245-6.

²⁹ Matthias Schurz et al. "Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies," *Frontiers in Human Neuroscience* 7:712 (2013): 2.

zich op relevante processen die met het perspectief te maken hebben in een sociale context. Om te zien welke gedeelte van het brein wordt geactiveerd bij een false-belief overtuiging worden er vaak korte verhalen, net zoals het verhaal van Sally en Anne, verteld aan volwassenen, onderzoek laat dan ook een consistent netwerk van geactiveerde brein gebieden zien, zoals de medial prefrontol cortex (mPFC), bilaterale temporal poles, the precuneus, en bilaterale temporal parietal junction (TPJ).³⁰

Het mPFC gebied wordt gerelateerd aan de verwerking van sociale en emotionele relevante informatie over andere mensen in de korte verhalen die verteld worden, maar wordt niet specifiek gerelateerd aan de verwerking van een overtuiging. Tegelijkertijd wordt het TPJ gebied verbonden aan de representatie van mentale perspectieven. Het rechter gebied van de TPJ wordt specifiek verbonden aan de representatie van overtuigingen. In het bijzonder werd het linker gebied van de TPJ geactiveerd bij false-belief verhalen. Dit suggereert dat het linker gebied van de TPJ verantwoordelijk is voor een werking die gebruikelijk is wanneer er gepraat wordt over een false belief, dus het verwerken van het perspectief vanuit de ander. Het linker gebied van de TPJ zou mogelijk een goede kandidaat zijn om perspectief en verandering in perspectief weer te geven. Hier is nog maar beperkt neuroimaging onderzoek in gedaan en blijft het lastig om concrete conclusies te maken.

In een ander onderzoek waarbij naar de TPJ wordt gekeken door middel van neuroimaging, zou het ook mogelijk als bewijs kunnen dienen voor de kind-als-wetenschapper theorie, maar in mindere mate voor de modulaire theorie. Reden hiervoor is dat bij dit onderzoek bleek dat de TPJ tijdens de ontwikkeling van ToM steeds selectiever wordt, dat wil zeggen dat dit gebied er steeds beter in wordt om belangrijke informatie van minder belangrijke informatie te kunnen onderscheiden. Dit is dan problematisch voor de modulaire theorie, omdat de modulaire theorie beweert dat er al vaste, aangeboren modules in het brein zitten, die dus niet meer verder ontwikkelen.

Verder kan de TPJ³¹ bij volwassenen een rol spelen bij het loslaten van het eigen, interne, zelf-perspectief om het externe perspectief van de ander te kunnen begrijpen. Dit vormt een belangrijk proces bij ToM. Als de TPJ wordt beschadigd door

³⁰ Matthias Schurz et al. "Common brain areas engaged in false belief reasoning," 2.

³¹ Mahy, Moses en Pfeifer, "Theory-of mind in the brain," 75.

een aanval of elektrische simulatie, dan resulteert dat in een gehallucineerd verkeerd beeld van het lichaam en de omgeving. Dit zou erop kunnen wijzen dat TPJ een rol speelt bij interne en externe representaties. Als TPJ dus inderdaad betrokken zou zijn bij het updaten van iemands interne verwachtingen versus externe realiteit, dan zou het mogelijk een passende hypothese kunnen zijn dat dit gedeelte van het brein een rol kan spelen in het aansturen van conceptueel begrip. Dit kan worden vergeleken met kinderen die informatie moeten integreren vanuit hun externe omgeving met interne toestanden om hypothesen te updaten. TT zou dus ondersteund kunnen worden door evidentie van neurale structuren of netwerken die domein specifieke conceptuele structuren ondersteunen, maar ook door leeftijd gerelateerde neurale veranderingen die geassocieerd worden met conceptuele veranderingen. De belangrijkste hypothese van TT is dat conceptuele verandering als respons op ervaringen de ontwikkeling van ToM kan voortzetten.

Dit blijft vooralsnog moeilijk te beoordelen, omdat als neurale evidentie TT wil ondersteunen, we moeten bepalen of verschillende neurale netwerken ook daadwerkelijk onderliggend zijn aan impliciete versus expliciete ToM, en of neurale veranderingen plaatsvinden bij de transitie van een impliciet naar een expliciet begrip van ToM.³² Maar ook hoe precies conceptuele verandering in het brein weerspiegeld wordt, want de TPJ zou een belangrijke rol kunnen spelen, maar blijft lastig om het empirisch te bewijzen. In het volgende deel licht ik een belangrijke positie tegen de TT toe, die ook nog steeds niet empirisch bewezen kan worden, maar nogal radicale consequenties met betrekking tot de TT kan hebben als deze positie gelijk heeft.

2.5 Problemen met de TT

Het blijft onduidelijk of we een ToM ontwikkelen zoals de kind-als-wetenschapper theorie suggereert, of dat we er mee aangeboren zijn zoals de modulaire theorie voorspelt. Daarnaast zijn er aantal andere logische problemen waarmee de TT worstelt.

Als ik een bepaalde theorie over iemands gedrag aanhou, bijvoorbeeld omdat ik een aantal keer heb meegemaakt dat als iemand met een leeg glas op de tafel tikt, hij of zij nog een glaasje drinken wil, dan zou ik deze theoretische generalisatie op iedereen kunnen toepassen. Hoe weet ik zo zeker dat ik deze theorie op iedereen kan toepassen?

³² Mahy, Moses en Pfeifer, "Theory-of mind in the brain," 76.

Als een vriendin met een leeg glaasje op de tafel tikt, maar dit doet uit verveling, of ritmisch met de muziek die aanstaat mee tikt, moet ik kunnen bepalen welke theoretische generalisatie hier van toepassing is. Dit lijkt op een lang proces en cognitief gezien minder efficiënt, want dan zouden we over oneindig veel theorieën moeten beschikken. Daarnaast zou er ook een onderscheid gemaakt moeten worden op wie welke theorie van toepassing kan zijn, bijvoorbeeld bij mensen met een sociale beperking, zoals autisme of psychopathie, is het ingewikkelder om algemene theoretische generalisaties toe te passen, omdat zij zich niet volgens de ‘algemene regels’ gedragen.

Een ander veel besproken kritiekpunt³³ op de TT kan worden gegeven vanuit de false-belief test. Een kritiekpunt is dat volgens de false-belief test kinderen pas over een ToM kunnen beschikken als zij vier jaar oud zijn, en dus geen begrip van anderen kunnen hebben tot ongeveer het vierde levensjaar. Dit zou betekenen dat baby’s en peuters niet over een ToM kunnen beschikken. Zahavi geeft hierop een aantal kritische argumenten³⁴ (zie ook het laatste hoofdstuk van deze scriptie). Hij zegt dat de TT beweert dat als we anderen willen begrijpen die ook een mind hebben, dat het een vereiste is dat het via een proces van mindreading moet gaan. Daarnaast zou volgens de TT het slagen voor de false-belief test noodzakelijk moeten zijn om over een ToM te kunnen beschikken. Kinderen beschikken niet over een ToM omdat zij niet tot hun vierde levensjaar de test kunnen halen. Dit argument³⁵ houdt mijns inziens niet helemaal stand, omdat de false-belief test in de huidige literatuur niet meer wordt gezien als iets noodzakelijks. Ten eerste vormt het begrijpen van mentale toestanden van anderen al als bewijs voor het beschikken over een ToM, Het begrijpen dat anderen verkeerde overtuigingen kunnen hebben staat los van dat het noodzakelijk moet zijn om over een ToM te kunnen beschikken.

Een ander probleem is dat om toegang te krijgen tot de mind van de ander, dit altijd een inferentieel proces moet zijn, dat wil zeggen het moet ergens uit af te leiden kunnen zijn, dus uit het lichamelijke gedrag. De manier waarop dit wordt uitgelegd leidt tot een problematisch verband tussen verborgen en extern, tussen ervaring en gedrag.

³³ Dan Zahavi, *Subjectivity and selfhood investigating the first-person perspective*, (Cambridge, Massachusetts: MIT Press, 2005), 197.

³⁴ Zahavi, *Subjectivity and selfhood*, 206.

³⁵ Mitchell Herschbach, “Folk psychological and phenomenological accounts of social perception,” *Philosophical Explorations* 11:3 (2008), 225.

Volgens Zahavi³⁶ is het ook iets wat tegen zijn intuïtie ingaat, om anderen te moeten begrijpen door mentale toestanden die niet observeerbaar, verborgen en theoretische gepostuleerde entiteiten zijn. Hij vindt wel dat in het volwassen leven we af en toe de mentale toestanden van een ander inferentieel toeschrijven, maar dat dit niet de basis kan zijn voor een soepele en onmiddellijke interpersoonlijke interactie. Zo toont Zahavi³⁷ ook aan dat baby's al snel het onderscheid kunnen maken tussen en objecten en andere mensen. Er zou dan ook niets inferentieels of theoretisch zijn bij deze vroege sociale interactie, het is bij baby's eerder een vorm van intersubjectiviteit ingeworteld door intuïtie, waardoor ze een begrip kunnen hebben van de gebaren van anderen.

2.5.1 Eliminativisme tegenover TT

Het eliminativisme heeft een radicale positie in de cognitiefilosofie omdat het beweert dat propositionele attitudes niet bestaan, en dus extreme gevolgen voor TT kan hebben. Het toeschrijven van mentale toestanden aan anderen om hun gedrag uit te kunnen leggen zijn volgens eliminativisten een mythe. De mythe is vergelijkbaar met het verhaal van de Oudgriekse Homerus.³⁸ Homerus schakelde de goden in om te voorspellen wie een veldslag zou winnen. Volgens eliminativisten bestaan er geen dingen zoals hoop, angst, gedachten of overtuigingen. De theorie zou uiteindelijk worden overgenomen door de neurowetenschappen. Churchland³⁹, een belangrijke aanhanger van het eliminativisme, stelt dat volkspychologie veel dingen niet kan verklaren. Een voorbeeld hierbij is dat belangrijke mentale fenomenen mysterieus binnen een volkspychologisch raamwerk blijven, zoals de dynamiek achter geestelijke ziekten, creatieve verbeelding, en verschil van intelligentie tussen individuen. Maar ook wordt de psychologische functie van slaap genegeerd, terwijl we een derde van ons leven aan slaap besteden. Daarnaast is er ook een variatie van perceptuele illusies, en het geheugen die belangrijke informatie naar boven kan halen. Churchland vindt het karakter van een leerproces het grootste mysterie, vooral waarbij er sprake is van grootschalig conceptuele verandering en vooral als dit in een pre-linguïstieke of non-linguïstieke vorm plaatsvindt zoals bij baby's. Volkspychologie zou dan niet in staat

³⁶ Zahavi, Subjectivity and selfhood, 207-8.

³⁷ Ibid., 211.

³⁸ Stephen Stich en Ian Ravenscroft, "What is folk psychology?" *Cognition* 50 (1994): 447-8.

³⁹ Paul M. Churchland, "Eliminative materialism and the propositional attitudes," *The Journal of Philosophy* 78:2 (1981): 73.

zijn om verklaringen te geven voor deze mysterieuze mentale fenomenen. Als het eliminativisme waar is, dan zou dat volgens eliminativisten een van de grootste intellectuele catastrofes kunnen zijn.

Om een beeld te krijgen waar het eliminativisme van overtuigd is, staan hieronder vier stellingen⁴⁰ weergegeven:

1. Een overtuiging, een verlangen of andere intentionele uitdrukkingen behoren tot de theoretische termen van een common sense ToM. Deze theorie wordt vaak volkspychologie genoemd.
2. Volkspychologie is een foute theorie. Beweringen die deze theorie maakt over toestanden en processen die een verklaring geven voor gedrag zijn fout.
3. Een volwassen wetenschap die uitlegt hoe de mind werkt en hoe het gedrag produceert wat we observeren kan niet verwijzen naar common sense intentionele toestanden en processen ingeroepen door volkspychologie. Overtuigingen en verlangens kunnen geen deel uit maken van een ontologie van een volwassen wetenschappelijke psychologie.
4. Intentionele toestanden van een common sense psychologie bestaan niet.

Hier botsen we al gelijk met het eliminativisme, want een argument tegen de overtuiging van het eliminativisme is dat volkspychologie geen beweringen maakt of proposities uitdrukt, dus het is niet iets wat waar of niet waar is. De tweede stelling, dat volkspychologie een verkeerde theorie is, kan dus niet kloppen, want als volkspychologie geen beweringen maakt, kan het geen foute beweringen maken.

In de cognitieve wetenschappen⁴¹ gaat de meest gebruikelijke uitleg-strategie, die wordt gebruikt als mensen hun cognitieve vaardigheden gebruiken, over geïnternaliseerde informatie of kennis structuren. Begin jaren negentig werd een andere verklaring gegeven, onze cognitieve vaardigheden zouden ook uitgelegd kunnen worden zonder gebruik te hoeven maken van geïnternaliseerde kennis structuren die lijken te heersen in hedendaagse cognitieve wetenschappen. Om het gedrag van anderen te voorspellen zouden we gebruik kunnen maken van een eigen besluitvormingssysteem om het besluitvormingssysteem van een ander te simuleren. We hebben dan geen theorie meer nodig om de ander te kunnen begrijpen. Deze theorie staat bekend als de

⁴⁰ Stich en Ravenscroft, "Folk psychology," 448-9.

⁴¹ Ibid., 464.

ST. In het volgende hoofdstuk zal ik de ST uitwerken en zien welke problemen er bij deze theorie kunnen ontstaan. De ST lijkt in eerste instantie op een betere theorie dan de TT voor het verklaren van hoe we anderen kunnen begrijpen. Zo is het bijvoorbeeld niet gevoelig voor het argument van eliminativisme, omdat het niet stelt dat mindreading gebeurt op basis van een theorie. Echter, ik zal laten zien dat de ST uiteindelijk niet zonder TT blijkt te kunnen.

3. Simulatie Theorie als theorie van mindreading

3.1 Simulatie Theorie

ST ontwikkelde zich als reactie tegenover de TT. ST suggereert dat individuen hun eigen mentale mechanismen gebruiken om mentale toestanden van anderen te kunnen voorspellen. Het is niet nodig om gebruik te maken van een theorie. Om iets over de mind van een ander individu te kunnen zeggen, verplaats je je in de mentale schoenen van de ander om zo te kunnen bepalen wat die ander ervaart in een bepaalde situatie.⁴² Volkspychologie wordt dan verklaard via een simulatie. Als het buiten regent en ik wil begrijpen wat mijn buurman gaat doen hoef ik daar geen theorie op toe te passen, maar ik vraag mezelf af ‘wat zou ik doen als ik zijn schoenen sta’? Een logisch antwoord hierop is: vanuit vele ervaringen waarin mensen lieten zien dat zij niet graag in de regen willen lopen, kan ik dit ook op de buurman toepassen, dus als hij toch naar buiten wil gaan neemt hij een paraplu mee.

Eind jaren negentig beweerden aanhangers van de ST dat de ontdekking van spiegelneuronen bij apen als bewijs zouden kunnen vormen voor een ST in de mind. Spiegelneuronen vormen een corticaal systeem waarbij observatie en uitvoering van een doel-gerelateerde handeling aan elkaar worden gekoppeld. Een functie van zo’n koppelsysteem kan zijn dat een organisme geactiveerd wordt die bepaalde mentale toestanden van de geobserveerde handeling kan herkennen. Die functie kan dan onderdeel zijn van mindreading. Een voordeel van de ST ten opzichte van TT zou kunnen zijn dat het cognitief economisch is; er hoeft niet eerst een bepaalde theorie te worden gezocht om de ander te kunnen begrijpen. Als je in een situatie simuleert wat de ander zou doen, kun je de uitkomst hiervan toepassen op anderen die een zelfde soort gedrag uitten. Daarnaast kan het goed omgaan met het eliminativisme die tegen een volkspychologie theorie ingaat.

Bij de ST verschillen de aanhangers over de precieze mechanismen van het simulatie proces.⁴³ Goldman is er van overtuigd dat voordat een mind reader een gepretendeerde mentale toestand kan toeschrijven aan een individu, hij eerst bij zichzelf naar binnen moet kijken om de mentale toestand door zijn cognitief systeem te

⁴² Mahy, Moses en Pfeifer, “Theory-of mind in the brain,” 71.

⁴³ Shaun Gallagher, “Simulation trouble,” *Social Neuroscience* 2:3-4 (2007): 355.

genereren om vervolgens de type en het inhoud te kunnen bepalen. Gordon daartegenover wijst introspectie en analogische inferenties af. Deze verschillen staan ook wel bekend als impliciete en expliciete ST, die ik hieronder verder zal uitwerken.

3.2 Impliciete en expliciete ST

Aanhangers van expliciete ST, zoals Goldman, gaan er vanuit dat een simulatie een bewust of introspectief proces is waarbij je jezelf in de situatie van de ander plaatst en een simulatie model gebruikt die gegenereerd wordt om de mentale toestanden van de ander te kunnen voorspellen.⁴⁴ Simulatie is expliciet in zoverre er gebruik wordt gemaakt van een bewuste introspectie waarbij verbeelding conceptuele propositionele attitudes, zoals overtuigingen of verlangens, kan manipuleren.

Impliciete ST⁴⁵ ziet simulatie als een sub-persoonlijk proces. Deze interpretatie werd aannemelijk door het neurowetenschappelijk bewijs waarbij de activatie van sub-persoonlijke spiegelneuronen een grote rol spelen. Het fenomenologisch argument wat vaak gebruikt wordt tegen een expliciete vorm van ST gaat over dat als we daadwerkelijk simuleren, we hier bewust van moeten zijn. Er bestaat geen bewijs voor dat als ik met iemand communiceer of hem of haar wil begrijpen ik er bewust van ben dat ik simuleer. Impliciete versies van ST kunnen beter omgaan met dit fenomenologisch argument, want als simulatie iets sub-persoonlijk is, dan is dat niet iets waarvan we bewust zijn. Een strenge versie van impliciete ST lijkt dan op een argument tegen expliciete ST, want als ons begrip van anderen bemiddeld wordt door een impliciete en automatisch simulatie proces, dan hebben we de expliciete versie van ST niet meer nodig. Impliciete simulatie is belangwekkend voor sociale neurowetenschappen, vooral als spiegelneuronen het bewijs kunnen vormen voor een ST.

Het debat over de relatie tussen spiegelneuronen en ST was afgelopen decennia een veelbesproken onderwerp. De ontdekking van spiegelneuronen zou ST kunnen ondersteunen als vorm van sociale cognitie zodat TT meer op de achtergrond komt te staan. Spiegelneuronen waren ontdekt in het premotor gedeelte van een makaak aap. Een functie van spiegelneuronen zou kunnen zijn dat het stimuleert om door imitatie te leren. Wanneer er nieuwe motorische vaardigheden worden ontwikkeld, wordt er vaak

⁴⁴ Gallagher, "Simulation trouble," 355.

⁴⁵ Gallagher, "Simulation trouble," 356.

in de eerste oefenfase geprobeerd om bewegingen van de persoon die wordt geobserveerd na te doen. Dat zien we vaak terug bij baby's: als een moeder haar tong uitsteekt naar haar baby, doet de baby haar moeder na door ook haar tong uit te steken. Spiegelneuronen vuren af bij het uitvoeren en het observeren van handelingen. Met handelingen bedoel ik iemand die een beweging maakt, bijvoorbeeld iemand die een glas water van de tafel oppakt.

Bij spiegelneuronen resoneert het motor systeem van iemand als hij een ander tegenkomt. Dat betekent dat het motor systeem wordt geactiveerd wanneer hij waarneemt dat een ander een intentionele handeling uitvoert. Spiegelneuronen in de pre-motor cortex en in het Broca gedeelte van het menselijk brein worden geactiveerd wanneer het subject zich betreft in specifieke instrumentele handelingen en wanneer het subject iemand observeert die zich met dezelfde handelingen bezighoudt. Deze subpersoonlijke mechanismen constitueren een simulatie van intenties van de ander.

Zoals eerder genoemd, zijn spiegelneuronen als eerst ontdekt bij aapjes door middel van *single cell recording*, dat betekent dat er specifieke neuronen worden gemeten.⁴⁶ De onderzoeken lieten zien dat de spiegelneuronen aanwezig waren in de hersenen van het aapje in hersengebied F5 en PF. Bij mensen wordt dit niet gemeten door middel van single cell recording, maar door een functional magnetic resonance imaging (fMRI) scan.

Daar staat tegenover dat er geen doorslaggevend bewijs is omdat spiegelneuronen bij apen en mensen niet op dezelfde manier gemeten kunnen worden. Een voordeel hierbij is wel dat ondanks onderzoeken bij apen beperkt worden tot bepaalde motor gebieden in het brein, dat bij mensen het gehele brein kan worden gescand door middel van fMRI scans, en dat het menselijke brein meerdere spiegelneuronensystemen kan weergeven, zoals emotie- en pijn- spiegelneuronensystemen.

Spiegelneuronen worden ook gezien als ondersteuning van *action understanding*.⁴⁷ Handelingen (actions) zijn conceptueel gerelateerd aan intenties. Een handeling is iets wat een individu (agent) intentioneel doet. Om te kunnen herkennen dat de uitvoering van bepaald gedrag (zwaaien met je hand) een handeling is (iemand

⁴⁶ Shannon Spaulding, "Mirror neurons and social cognition," *Mind & Language* 28:2 (2013): 236.

⁴⁷ Spaulding, "Mirror neurons," 238.

willen begroeten), is het belangrijk om de intentie (groeten) van de agent te begrijpen. Een intentie wordt vaak gezien als een mentale toestand. Spiegelneuronen vuren af als reactie op een bepaald gedrag, bijvoorbeeld een hand die iets pakt. Toch is er een metafysisch en epistemologisch probleem tussen de activatie van spiegelneuronen en de mentale toestand die er aan toegeschreven kan worden voor *action understanding*. Als iemand gaapt, en de ander gaat ook gapen door de activatie van spiegelneuronen, betekent het niet dat het ‘gaap-reflex’ van de ander begrijpt dat ik verveeld ben. Het is nodig om een bepaalde mentale toestand, zoals verveeld zijn, te kunnen herkennen om er begrip van te hebben, en een gaap reflex kan die mentale toestand niet herkennen. Een automatische resonantie van spiegelneuronen betekent niet dat mentale toestanden herkend kunnen worden. Zou het mogelijk zijn dat de activiteit van spiegelneuronen niet noodzakelijk is om de intentie van de ander te kunnen begrijpen? In het volgende deel zal ik hier op in gaan en een aantal kritiekpunten uitleggen waarom spiegelneuronen als bewijs voor ST lastig te beoordelen zijn.

3.3 Spiegelneuronen kritiek

Ik zal in dit deel met name de kritiek uitwerken waarbij er vanuit wordt gegaan dat spiegelneuronen action understanding ondersteunen. Zoals ik het deel hierboven met de vraag afsloot of dat de activiteit van spiegelneuronen noodzakelijk zijn om intenties van anderen te kunnen begrijpen, zal ik hier beargumenteren dat dat niet nodig is. Volgens een aantal onderzoekers in de neurowetenschappen, zoals Rizzolatti en Craighero⁴⁸, kan het niet zo zijn dat het spiegelneuronensysteem het enige mechanisme is dat action understanding ondersteunt. Een alternatief voor spiegelneuronen om action understanding te ondersteunen zou de *sulcus temporalis superior (STS)* kunnen zijn, want bij makaak aapjes is ontdekt dat cellen in het STS gebied reageren op uitgebreide handelingen op een manier die meer verfijnd is in vergelijking met spiegelneuronen. De neuronen in het STS gebied reageren op handelingen zoals naar of van iets weglopen, het hoofd omdraaien, arm bewegingen, hand-object interactie waarbij er een selectiviteit plaats vindt voor specifieke handelingen zoals aanreiken, manipuleren en oppakken. Deze cellen bezitten geen motorische vaardigheden, in dat ze niet afvuren tijdens het

⁴⁸ Gregory Hickok, “Eight problems for the mirror neuron theory of action understanding in monkeys and humans,” *Journal of Cognitive Neuroscience* 21:7 (2008): 1233.

uitvoeren van een handeling. De neuronen in het STS⁴⁹ reageren ook op passieve waarneming van handelingen, maar worden in het algemeen niet gezien als onderdeel van spiegelneuronen, omdat ze dus geen motorische vaardigheden bezitten. Het gebied waarin de inferior pariëtale cortex zich bevindt in de hersenen, bevatten spiegelneuronen (PF) en projecteert op F5, ontvangt input vanuit de STS. Dit lijkt op een ideale schakeling om handelingen te representeren (STS) en hun interactie (PF) te coördineren met het motor systeem (F5).

De hype rondom spiegelneuronen⁵⁰ ontstond doordat spiegelneuronen complexe menselijke vaardigheden en stoornissen zoals autisme zouden kunnen uitleggen. Het lastige hierbij is dat de spiegelneuronen die bij de aapjes zijn ontdekt geen hogere-order cognitieve processen kunnen bevatten, net zoals dat mensen die in het bezit zijn van hogere-order cognitieve processen niet uitsluitend over spiegelneuronen beschikken. Volgens Hickok wordt er een onjuist beeld over de functie van spiegelneuronen geschetst, dit komt doordat er in veel onderzoeken menselijk gedrag wordt onderzocht en dat gelinkt wordt aan het spiegelneuronensysteem, terwijl bij apen juist wordt aangetoond dat dat gedrag niet door spiegelneuronen wordt ondersteund. Het is lastig een verband tussen bepaald menselijk gedrag en spiegelneuronen te maken, want het probleem bij spiegelneuronen is dat het systeem gegeneraliseerd is op mensen zonder een systematische bevestiging, en dus ook met de toepassing dat spiegelneuronen action understanding zou ondersteunen. Als een onderzoek naar mensen wordt gedaan met de aanname dat spiegelneuronen action understanding ondersteunen, en dat een homolog en functioneel systeem ook bestaat in mensen, dan is het makkelijk om te zeggen dat er een logische inferentie kan worden gemaakt waarbij het spiegelneuronensysteem hogere-order functies ondersteunen zoals taal en empathie. Deze inferentie is niet meer correct als een van de aannames over spiegelneuronen niet meer klopt. Het is dus belangrijk om eerst de aannames te valideren voordat er gevolgtrekkingen kunnen worden gemaakt over menselijk gedrag. Daarna is het belangrijk te bepalen of onderzoeken kunnen bewijzen dat spiegelneuronen ook in mensen bestaan.

Bij mensen gedraagt het spiegelneuronensysteem zich anders in vergelijking met apen omdat het bij mensen wel imitatie ondersteunt: het wordt geactiveerd bij perceptie

⁴⁹ Pascal Molenberghs et al. "The role of the superior temporal sulcus and the mirror neuron system in imitation," *Human Brain Mapping* 31 (2010): 1317.

⁵⁰ Hickok, "Eight problems for the mirror neuron theory," 1234.

en uitvoering van doelloze handelingen. Op basis van deze observatie concludeerden sommige onderzoekers dat het spiegelneuronensysteem bij mensen zijn ontwikkeld om niet alleen action understanding te ondersteunen, maar ook om imitatie te ondersteunen. De aanname die door de onderzoekers⁵¹ wordt gemaakt is dat tijdens de evolutie oude eigenschappen van spiegelneuronen bewaard zijn gebleven. Kan het dan ook zo zijn dat de spiegelneuronensysteem die bij mensen is ontwikkeld nu imitatie ondersteunen maar geen action understanding meer ondersteunen? Misschien hebben mensen een verfijnde semantische systeem ontwikkeld, onderscheiden van het motor systeem, waarbij het spiegelneuronensysteem bevrijdt werd om imitatie te ondersteunen. Deze mogelijkheden of opties worden niet overwogen in mainstream theorieën over spiegelneuronen. In plaats daarvan wordt data over apen gebruikt en op mensen geprojecteerd, zonder daar echt empirisch validatie te hebben over de aannamen.

In het volgende deel zal ik nog een aantal andere kritische punten uitwerken die een probleem vormen voor de ST, maar ook voor mindreading meer in het algemeen. Daarna zal ik een hybride theorie van ST in combinatie met TT uitwerken, om te zien of dat een betere mogelijkheid voor mindreading vormt.

3.4 Problemen met de ST

Naast het eerder besproken probleem van hoe spiegelneuronen als bewijs voor ST kan vormen, zijn er nog een andere problemen waar de ST tegen aan loopt. Dennett⁵² is bijvoorbeeld erg kritisch over het feit dat als we de ander willen begrijpen, we alleen maar bezig zijn met simuleren. Als ik wil voorspellen of verklaren hoe een hangbrug zich in bepaalde weeromstandigheden gedraagt, dus bijvoorbeeld wat er met de hangbrug gebeurt als de wind waait, lukt dat niet alleen met simuleren. Het antwoord hangt af van of ik over complexe fysische en bouwkundige kennis van een hangbrug beschik. Kennis van een geïmiteerd object is nodig om te kunnen simuleren, maar die kennis moet wel in iets van een theorie zitten. Het blijft dus belangrijk om over een TT te beschikken om daarna vervolgens te kunnen simuleren.

Zoals eerder genoemd, speelt volgens de expliciete vorm van ST introspectie een belangrijke rol om de mentale toestanden van de ander te kunnen voorspellen. Bij de eerste stap in dit proces moet diegene die mentale toestanden toeschrijft aan de ander in

⁵¹ Ibid., 1234.

⁵² Daniel Dennett, *The intentional stance*, (Massachusetts: MIT Press, 1998), 100-1.

zichzelf mentale toestanden creëren om zo met de (verwachte) mentale toestanden van de ander te kunnen matchen. Dit lijkt volgens Gallagher⁵³ nogal problematisch omdat degene die simuleert al een idee moet hebben van wat er in de ander omgaat. Er moet ergens kennis vandaan komen en Gallagher vraagt zich terecht af waarom dat niet het gene is wat we proberen uit te leggen? Een hybride theorie, (zie volgende hoofdstuk), probeert dat op te lossen door TT en ST te combineren en daarbij algemene regels over hoe mensen denken en zich gedragen in bepaalde situaties toe te passen in de vorm van een volkspychologische theorie. De simulationist kan vervolgens volkspychologie gebruiken om 'alsof' mentale toestanden te creëren die nodig zijn voor het simulatie proces. Hierbij wordt er vaak naar het proces van spiegelneuronen verwezen, maar het probleem hierbij is dat de processen 'vertaald' moet worden naar een conceptueel begrip van propositionele attitudes. Het is niet genoeg om het over een bepaalde mentale toestand zoals overtuiging of verlangen te hebben, want er moet eerst worden gekeken welke mentale toestand er precies gebruikt wordt.

Daarnaast bestaat er het klassieke argument⁵⁴ van inferentie door analogie. Het kan niet zo zijn dat als ik van iemand een bepaald gedrag waarneem, en ik vervolgens datzelfde soort gedrag bij iemand anders waarneem, en vervolgens dat simuleer om te begrijpen wat de mentale toestand van diegene is, het dan klopt. ST negeert hierin de diversiteit van gedragingen die aan de verschillende soorten mentale toestanden kunnen worden gekoppeld. Er bestaan zoveel verschillende beweegredenen, overtuigingen, verlangens en gedragingen in de wereld dat het onwaarschijnlijk lijkt dat je je eigen simulatie proces die bestaat uit beperkte ervaringen kan toepassen op de ander.

Ten slotte kan er een argument⁵⁵ worden gebruikt uit de ontwikkelingspsychologie. Peuters van ongeveer 18 maanden zijn al in staat om mislukte intenties van anderen te kunnen begrijpen, en deze zelfs te verbeteren. Er bestaat geen bewijs dat we daadwerkelijk simuleren, want als dat expliciet gebeurt dan zouden we er enigszins bewust moeten zijn van de verschillende stappen die we doorlopen. Wanneer we onze eigen ervaring gebruiken over hoe we anderen begrijpen, gebruiken we geen bewuste simulatie routine om de ander te begrijpen. In het volgende hoofdstuk zal ik nog een

⁵³ Gallagher, "Simulating minds," 355.

⁵⁴ Ibid., 355.

⁵⁵ Ibid., 355.

ander belangrijk argument beschrijven dat beide voor de TT en ST problematisch kan zijn.

4. TT en ST

4.1 Het frame-probleem en de TT en ST

Een ander argument waar de ST en TT beide tegen aan lopen is het zogeheten frame probleem.⁵⁶ Het frame probleem stelt dat om de mentale toestanden van iemand te kunnen theoretiseren of simuleren, het belangrijk is om iets van zijn gedrag of situatie af te weten. Aanhangers van de ST zeggen dat theoretische generalisaties niet toegepast kunnen worden omdat er simpelweg geen manier is om uit te leggen hoe een goede generalisatie snel gevonden kan worden die noodzakelijk is om een begrip van gedrag te hebben. Generalisaties over gedrag kunnen niet werken tenzij je al weet welke generalisaties je op de belangrijke eigenschappen van de situatie kan toepassen. Hetzelfde probleem ontstaat ook voor de ST, want als je moet uitzoeken welke theoretische generalisaties relevant zijn voor een situatie, is het ook complex om te zien welke mentale toestanden je moet simuleren om de situatie te kunnen begrijpen.

Hier uit volgt dat het in eerste instantie problematisch is om te theoretiseren of simuleren zonder eerst iets af te weten van de situatie van de ander. De kennis gaat over het begrijpen van de situatie van de ander, en als ik uitleg probeer te vinden van het getoonde gedrag door simulatie, moet ik eerst weten wat dat gedrag precies is. De betekenis van gedragingen en situaties zijn niet overduidelijk zichtbaar, en meerdere betekenissen van een bepaald gedrag zijn mogelijk. In iedere situatie, volgens Wilkinson⁵⁷, moet onze vaardigheid begrepen worden om situaties en gedragingen te verbinden aan betekenisvolle patronen dat mogelijkheden projecteert op toekomstige handelingen en rationele verklaringen van gedrag in het verleden. De vaardigheid om deze verbindingen te zien tussen situaties en gedragingen met toekomstige mogelijkheden en door het verleden is de vaardigheid om karakteristieke eigenschappen van de ander die belangrijk zijn te voorspellen. Als je je afvraagt waar iemand nu is, kan er oneindig veel worden toegeschreven aan waar diegene is, maar we zijn in staat om onszelf te beperken in het toeschrijven van een of een paar beschrijvingen van situaties en gedragingen waardoor we een voorspelling kunnen doen die nauwkeurig is om te zien wat de motieven en overtuigingen zijn van het gedrag van de ander. Bijvoorbeeld,

⁵⁶ Michelle Maiese, *Embodiment, emotion, and cognition*, (London: Palgrave Macmillan, 2011), 154.

⁵⁷ William S. Wilkerson, "Simulation, theory, and the frame problem: The interpretive moment," *Philosophical Psychology* 14:2 (2001): 149.

je wilt graag met je beste vriend uiteten en je belt hem maar hij neemt zijn telefoon niet op. Je vraagt je af waarom je hem niet kan bereiken en je bedenkt je dan dat hij iedere vrijdagavond gaat uiteten met zijn collega's. Door iets van zijn situatie af te weten, wordt het minder ingewikkeld om te bedenken waarom hij niet bereikbaar is, en hoeven er aan hem niet allerlei mentale toestanden worden toegeschreven.

Voordat een simulatie proces of theoretische beredenering begint, moet er eerst een karakterbeschrijving gevonden worden, deze karakterbeschrijving bepaalt welke voorspelling of verklaring gegenereerd wordt voor het gedrag van de ander. Wilkinson blijft er bij dat ST of TT als verklaring voor een volkpsychologische vaardigheid om anderen mensen te begrijpen niet genoeg is, er moet iets voorafgaand zijn waarna dan pas iets van een TT of ST kan worden toegepast.

4.2 Hybride theorie TT/ST

TT en ST hoeven niet met elkaar in competitie te zijn. Het is ook mogelijk om simulatie en het theoretische gedeelte te combineren als verklaring van mindreading.

In de jaren negentig was Goldman, o.a. door de ontdekking van spiegelneuronen, een aanhanger van de ST. Recent staat hij echter voor een 'samenwerking' tussen de ST en TT en stelt daarom een hybride ST/TT theorie voor. Hij geeft twee voorbeelden⁵⁸ om te laten zien hoe de ST en TT samen kunnen gaan. Het eerste voorbeeld is dat TT kan worden gebruikt om een doen-alsof toestand te selecteren. Op het moment dat er een beslissing of voorspelling wordt gemaakt, gebruikt de persoon die mentale toestanden toeschrijft aan het target (de ander), een theoretische beredenering om de eerdere toestanden van het target te infereren. De pretend-toestanden worden dan in het mechanisme geplaatst die beslissingen maakt. Bij de eerste stap wordt er gebruik gemaakt van de TT, en bij de vervolg stappen wordt er gesimuleerd.

Bij het tweede voorbeeld⁵⁹ gaat het om een andere vorm van mindreading, namelijk om de voorspelling van een beslissing, er wordt hierbij gekeken waarom iemand zich op een bepaalde manier uit. De mentale toestanden moeten als het ware berekend worden om te zien waarom het target zich op een bepaalde manier gedraagt. Hoe gaat dit in zijn werking? De mindreader moet zich eerst verplaatsen in voorafgaand

⁵⁸ Goldman, *Simulating minds*, 44.

⁵⁹ Goldman, *Simulating minds*, 44-5.

bewijs, dus hij kijkt naar hoe het target zich in een eerdere gelijkwaardige situatie gedroeg. Vele mentale toestanden moet passen in een tweede patroon, waarbij een gevraagde mentale toestand de oorzaak is van gevolgen die bekend zijn. Vervolgens denkt de mindreader weer terug aan waargenomen gedrag en gezichtsuitdrukkingen. Een probleem hierbij is dat een psychologisch mechanisme niet achteruit kan gaan, dus bij de toepassing van een ST denken we juist vooruit.

Hoe zou een ST bij deze vorm van mindreading toegepast kunnen worden en samen kunnen gaan met TT? Bijvoorbeeld door middel van het gebruik van een *generate-and-test* strategie. De mindreader begint met een gevolg (waar hij al kennis over heeft) van de gezochte mentale toestand. Dan genereert hij een of meer hypothesen over eerdere mentale toestanden of een combinatie van toestanden die verantwoordelijk kunnen zijn voor het gevolg. Vervolgens test hij de hypothesen door net te doen alsof hij zich in die toestanden bevindt, dan worden de hypothesen in een geschikt psychologisch mechanisme gestopt en kijkt hij of de uitkomst gekoppeld kan worden met wat hij heeft geobserveerd. Als dit samen gaat, schrijft hij de mentale toestanden of combinatie van toestanden toe aan het target. De generate-and-test strategie past simulatie toe op een cruciaal moment maar is ook afhankelijk van een theorie.

Hoe ziet een generate-and-test strategie er precies uit in alledaagse interactie? Een voorbeeld hierbij is: ik ben op een feestje van een goede vriend en hij kijkt naar mij en dan weer terug naar de lege glazen die op tafel staan van de gasten en vervolgens weer naar mij en glimlacht. Ik vraag me af wat hij precies bedoelt met zijn gedrag en probeer hem te begrijpen. Ik kan me van een vorig feestje herinneren dat hij graag wilde dat ik hem hielp met het voorzien van drankjes voor de gasten. Ik bedenk me dat hij ook glimlachte toen ik drankjes inschonk, dus als ik me in hem verplaatst zou het kunnen zijn dat hij met zijn blik naar de lege glazen en glimlach bedoelt of ik hem wil helpen met het voorzien van drankjes voor de gasten. Als ik dan vervolgens de lege glazen inschenk en hij zijn duim opsteekt, klopt het dat ik juiste mentale toestanden aan hem heb toegeschreven door eerst te genereren waarom hij een bepaald gedrag liet zien, en vervolgens te testen of het klopt wat ik dacht dat hij bedoelde.

Doordat de generate-and-test strategie samenwerking vereist tussen TT en ST, ziet een hybride theorie er aannemelijk uit. Toch blijft een aantal problemen nog steeds bestaan, ook als TT en ST samenwerken, zoals het frame-probleem maar ook hoe we

toegang tot de mind van de ander kunnen hebben, want dat lijkt volgens de TT en ST iets verborgen te zijn.

Zou het ook mogelijk zijn om via een andere vorm, dan TT en ST, de ander te kunnen begrijpen, zoals iets wat voorafgaand aan TT of ST gaat? In het volgende hoofdstuk wil ik een nieuwe theorie, de Interactie-Theorie, bespreken om te zien of dat een goede mogelijkheid vormt om de ander te kunnen begrijpen maar ook of het zonder een bepaalde ToM mogelijk is. Daarnaast wil ik bekijken hoe de Interactie-Theorie (IT) met problemen binnen de TT en ST omgaat, en in hoeverre de IT deze problemen kan oplossen en welke blijven bestaan.

5. Interactie Theorie als alternatief

Zoals we in de voorgaande hoofdstukken zagen, zijn de TT en ST lang dominant geweest in het debat over hoe we over mindreading beschikken. In recente ontwikkelingen zijn er andere theorieën ontwikkeld over hoe we de ander kunnen begrijpen. De theorie, IT, die ik in dit hoofdstuk ga bespreken gaat over het elkaar begrijpen door middel van directe perceptie, zoals ik kort in het laatste deel van het vorige hoofdstuk heb besproken. Dit is een radicale theorie die de TT, ST en mindreading in het algemeen afwijst. Belangrijk hierbij is om te zien of IT beter met de argumenten kan omgaan die tegen TT en ST ingaan. Ik zal als eerst de IT toelichten en vervolgens zien hoe IT zich vormde als reactie op TT en ST, met name de kritiek waarmee TT en ST worstelen en ten slotte wil ik laten zien dat IT een minder goede theorie is om elkaar te begrijpen.

5.1 IT als theorie voor het begrijpen van de ander

Voordat een kind vier jaar is, zijn de menselijke interactie vaardigheden en intersubjectief begrip al gevormd in bepaalde belichaamde oefeningen.⁶⁰ De oefeningen zijn emotioneel, perceptueel, non-conceptueel en sensomotorisch en gaan over het verwerken van waargenomen intenties, imitatie en de processen die onder primaire intersubjectiviteit vallen. Deze belichaamde oefeningen constitueren onze primaire toegang voor het begrijpen van anderen.

In de meeste intersubjectieve situaties hebben we een direct perceptueel begrip van de intenties van de ander, dat komt omdat hen intenties expliciet uitgedrukt worden in belichaamde handelingen, uitgedrukt gedrag en emoties. Als iemand een bepaalde frons op zijn gezicht heeft, kunnen we direct waarnemen dat diegene boos is, zonder eerst te moeten bedenken wat de frons betekent. Volgens Gallagher is een overtuiging of verlangen te postuleren of te infereren onnodig om hier een begrip van te kunnen hebben die ergens weggestopt zit in de mind van de ander. Wat we een verlangen of overtuiging van de ander noemen, kan direct uitgedrukt worden in gedrag en handelingen. Het direct begrip is mogelijk gemaakt door complexe neuronale processen dat wordt beschreven als een spiegelneuronensysteem. In plaats van deze neuronale

⁶⁰ Shaun Gallagher en Daniel Hutto, "Understanding other through primary interaction and narrative practice," in *The Shared Mind: Perspectives on Intersubjectivity*, red. J. Zlatev, T. Racine, C. Sinha en E. Itkonen (Amsterdam: John Benjamins, 2008), 20.

resonantie proces te interpreteren als een impliciete simulatie, die op een functioneel niveau over cognitieve processen gaan en perceptie van handeling, beargumenteerd Gallagher⁶¹ dat ze in feite een vorm van *enactive social perception* weergeven.

5.2 IT bij baby's en kinderen

Bij pas geboren baby's is er al een perceptueel begrip van anderen aanwezig. Zij kunnen ook herkennen dat de ander van dezelfde soort als zij zijn, baby's kunnen dus een verschil maken tussen objecten en mensen. Door het feit dat ze menselijke gezichten kunnen imiteren, overtuigt dit ervan dat baby's de omgeving al kunnen onderscheiden tussen menselijke handelingen en dingen die geen menselijke handelingen kunnen uitvoeren, dus het verschil tussen mensen en objecten. Baby's en kleine kinderen zijn in staat zonder een TT of ST om belichaamde bewegingen als een intentionele beweging te kunnen zien. Diverse onderzoeken, zoals die van Baldwin⁶², tonen aan dat baby's van tien maanden de ogen van een persoon volgen en verschillende bewegingen van het hoofd, de mond en algemene lichaamsbewegingen kan waarnemen als betekenisvolle doelgerichte bewegingen.

Daarnaast hangt primaire subjectiviteit ook af van de coördinatie door de gebaren en uitdrukkingen van de baby en die van de verzorger waarmee hij of zij interacteert. Het is bij het waarnemen van emoties in de beweging van anderen in eerste instantie niet nodig om een theoretische stand of een simulatie van een innerlijke stand te creëren. Deze vorm van een waarneming-gebaseerd begrip kan in feite geen vorm van mindreading worden genoemd, want bij het waarnemen van handelingen en uitgedrukte bewegingen van de ander zie je al de betekenis, er is geen inferentie nodig om verborgen mentale toestanden van de ander te kunnen begrijpen.

Vaardigheden die nodig zijn voor primaire intersubjectiviteit, suggereren dat voordat we in een positie zijn om ons af te vragen wat de ander verlangt, we altijd al een specifieke perceptuele begrip hebben van wat zij voelen, ook al zijn de intenties van diegene niet sympathiek. Binnen primaire intersubjectiviteit is een gebruikelijke lichamelijke intentionaliteit die gedeeld wordt door het waarnemende subject en de waargenomen ander. Dus voordat we in een positie zijn om te theoretiseren, simuleren,

⁶¹ Ibid., 20.

⁶² Ibid., 21.

uit te leggen, of de mentale toestanden van anderen te voorspellen, verkeren we al in een positie om te interacteren met anderen, en anderen ook te begrijpen door het waarnemen van uitdrukkingen, gebaren, intenties en emoties, maar ook hoe zij zich gedragen naar zichzelf toe en naar anderen.⁶³

Toch ontkent de IT niet dat het soms nodig is om een volkspychologie toe te passen, maar volgens de IT gebeurt dit vooral als er iets mis gaat bij de interactie. Bij een complexe interactie vinden IT-aanhangers dat een volkspychologie geleerd wordt door rollen en regels van onze wereld, zoals je leert hoe je je in bepaalde situaties moet gedragen, en ook hoe de ander zich in een bepaalde situatie moet gedragen. IT vindt dat het leerproces niet gebeurt volgens theoretische proposities, zoals de TT stelt, maar gebeurt juist door volkspychologische verhalen. Deze theorie staat bekend als de Narrative Practice Hypothesis⁶⁴, die ik hieronder verder zal bespreken.

5.3 Narrative Practice Hypothesis

De Narrative Practice Hypothesis (NPH) stelt dat een verhalende vaardigheid een goed alternatief kan zijn voor de TT of de ST en dat het meerdere vormen van mindreading kan uitleggen. Volgens Hutto⁶⁵ bereiken kinderen een vorm van volkspychologisch begrip door ze te betrekken in het vertellen van verhaaltjes. Kinderen worden vaardiger in het begrijpen van motivaties van het gedrag van anderen door betrokken te zijn bij verhalen (*narratives*) die verteld worden door de verzorgers van het kind. Bij volkspychologische verhalen wordt er vaak de nadruk op mentale termen zoals wensen, overtuigen en weten gelegd, en ook over hoe de karakters zich in een verhaal voelen en wat zij willen. Bij dit proces leren kinderen hoe de verschillende mentale toestanden zich gedragen in relatie tot elkaar. Kinderen leren hierbij ook welke attitudes belangrijk zijn voor de hoofdpersonen in het verhaal.

Een volkspychologisch verhaal is bijvoorbeeld Roodkapje, bij het vertellen van dit verhaal wordt uitgelegd waarom de verschillende personages zich op een bepaalde manier gedragen. De NPH⁶⁶ benadrukt dat volkspychologische verhalen niet de basis van interpersoonlijke interactie is. Ze worden alleen gebruikt als de interactie afwijkt van de normale interactie, dus als de een zich gedraagt op een manier waardoor de

⁶³ Ibid., 22.

⁶⁴ Ibid., 27.

⁶⁵ Ibid., 27-8.

⁶⁶ Ibid., 29.

ander moeilijk begrijpt waarom diegene zich zo gedraagt. Volkspychologie via de NPH methode is alleen nodig in bijzondere gevallen waarbij we het verhaal van de ander niet kennen, of als we heel verbaasd zijn door de handelingen van de ander. In het volgende deel zal ik de kritiek vanuit de IT op mindreading analyseren om te zien of dat een goede mogelijkheid vormt voor mindreading.

5.4 IT als reactie op ST en TT

De ST en TT stellen dat als we de ander willen begrijpen, we dat doen via een simulatie of theoretisering om toegang te krijgen tot de mentale toestanden, omdat we daar geen directe toegang toe hebben. We moeten mindreaden via overtuigingen, verlangens en intenties van de ander, en via deze manier van mindreading hebben we een primair begrip van het gedrag van de ander. De TT en ST karakteriseren sociale cognitie als een proces om uit te kunnen leggen of te voorspellen wat de ander gaat doen of heeft gedaan. TT stelt dat we het gedrag van de ander kunnen uitleggen door een aangeboren of ontwikkelde theorie die uitlegt waarom iemand zich op een bepaalde manier gedraagt. De theorie wordt geplaatst in termen van mentale toestanden zoals overtuigingen en verlangens, en die mentale toestanden veroorzaken of motiveren gedrag. De ST daartegenover vindt een theorie overbodig, want we kunnen onze mind gebruiken als model waarmee we de mentale toestanden van de ander kunnen simuleren. We plaatsten onszelf in de overtuigingen of verlangens van de ander en daarmee doen we alsof we in hun situatie zijn. Hieronder zal ik kort de problematiek weergeven die bij de TT en ST een rol spelen, en ook analyseren in hoeverre de IT met de problemen kan omgaan.

Er bestaan verschillende tegenargumenten voor de processen die TT en ST als impliciet beschrijven.⁶⁷ Bij TT is er geen bewijs dat er ook daadwerkelijke zulke impliciete processen zijn. TT stelt false-belief experimenten als empirisch bewijs voor TT, maar de vraag hierbij is of dat er expliciete of impliciete ToM processen worden getest. Het is moeilijk om de vraag te beantwoorden, want bij de false-belief test wordt er aan de proefpersonen gevraagd om de betekenis te begrijpen vanuit een derde persoon, wat dus op een expliciet proces kan lijken.

⁶⁷ Ibid., 18.

Het idee dat spiegelneuronen als bewijs voor een impliciete ST⁶⁸ kunnen dienen is ook problematisch. Het probleem is dat simulatie wordt gebruikt als een eerste-persoon model om de mentale toestanden van een derde-persoon te begrijpen. Hierbij worden er als-dit of alsof mentale toestanden gevormd vanuit het perspectief van de ander. De activatie van spiegelneuronen kunnen niet over een eerste of derde persoon gaan, en wanneer er een alsof mentale toestand wordt gevormd, kunnen de neuronen niet doen alsof ze zich afvuren, zij vuren zich af of niet. Als spiegelneuronen worden afgevuurd, zijn zij neutraal met betrekking tot de ander.

IT lijkt dus beter om te kunnen gaan met de spiegelneuronentheorie, die voor de ST nogal problematisch lijken te zijn, omdat de IT vanuit een eerste persoon model de ander probeert te begrijpen. Een ander probleem is dat we geen toegang tot de mind van de ander hebben, iets waarmee de IT beter lijkt om te kunnen gaan. Zahavi⁶⁹ stelt dat de TT beweert dat een begrip hebben van anderen hetzelfde is als een begrip hebben van bomen en wolken. Andere mensen zijn dan ingewikkelde objecten in onze omgeving waarvan we het gedrag proberen te voorspellen en te verklaren, maar het innerlijke blijft altijd verborgen. Hij vraagt zich of er geen andere manier is om elkaar te begrijpen, zoals anderen ook zien als subjecten net zoals dat wij een subject zijn en dat het de epistemologische situatie anders maakt. Volgens Zahavi zou de IT hierop een oplossing voor kunnen geven, want volgens de IT hebben we direct toegang tot gevoelens, verlangens en overtuigingen van anderen door het gedrag wat door de ander wordt uitgedrukt, want dat is betekenisvol en expressief een geeft toegang tot de mind van de ander.

De IT suggereert dat er iets eerder, dan de TT en ST, ontstaan is om anderen te begrijpen en wat aan de basis ligt van alledaagse interactie. Daarnaast stelt IT dat een algemene volkpsychologie niet voldoende is om verschillende soorten gedragingen in het alledaagse leven uit te leggen, dit probleem werd eerder bij het frame-probleem benadert. IT lijkt hier ook geen antwoord op te kunnen geven, want hoe bepaalt IT welke karakterbeschrijving toegepast kan worden in een alledaagse interactie? Volgens de IT begrijpen we de ander door middel van directe waarneming, maar is het niet belangrijk om eerst iets van de situatie van de ander af te weten voordat we het gedrag

⁶⁸ Gallagher en Hutto, "Primary interaction," 19.

⁶⁹ Zahavi, *Subjectivity and selfhood*, 206-7.

van de ander begrijpen? Bijvoorbeeld, als er iemand bij het station boos wegloopt van het perron, zie ik natuurlijk als eerst een boos gezicht en begrijp ik dat diegene boos is. Maar waarom is diegene boos? Dan zou ik eerst iets van zijn situatie af moeten weten om te kunnen bepalen waarom hij boos is, bijvoorbeeld omdat hij de trein heeft gemist, hij er te laat achter komt dat hij zijn telefoon is vergeten, of zijn baas die hem net opbelt om te zeggen dat hij ontslagen is.

In het volgende hoofdstuk zal ik nog een aantal andere kritische argumenten tegen de IT bespreken.

6. Kritiek op IT

Een vaak voorkomend argument tegen ToM vanuit de IT is dat het vooronderstelt dat andere minds niet observeerbaar zijn. Dat betekent dat de mentale toestanden van anderen via perceptie niet toegankelijk zijn. Bohl en Gangopadhyay⁷⁰ laten echter zien dat de onobserveerbaarheidsassumptie (OA) niet toegepast kan worden op de ToM. De ToM wordt vaak geïnterpreteerd als een Cartesiaans lichaam-geest dualisme, waarbij de indruk kan worden gewekt dat de kritiek van OA gericht is op een metafysische dualistische achtergrond van ToM. Toch is het onjuist om ToM ervan te beschuldigen dat het over een lichaam-geest dualisme gaat. De metafysische achtergrond van ToM neigt juist naar een functionalisme en ToM zou losgekoppeld moeten worden van de metafysica van de mind. De reden hiervoor is dat het doel van de ToM is om uit te leggen hoe mensen mentale toestanden aan bepaalde entiteiten toeschrijven waarbij het niet uitmaakt of de entiteiten minds hebben of niet op metafysisch gebied. Hieruit kan worden geconcludeerd dat een kritiekpunt op de OA niet gelezen kan worden als een kritiek vanuit een metafysische claim.

Een ander kritiekpunt⁷¹ op de OA wordt vaak als een epistemologische claim gezien. Zoals eerder benoemd, wordt de ToM er vaak van beschuldigd van een Cartesiaanse asymmetrie tussen lichaam en geest uit te gaan, waarbij de mind een innerlijk rijk is, dat verborgen is voor anderen en alleen toegankelijk is voor diegene waarbij de mind toebehoort. Het probleem ontstaat als de OA als een Cartesiaans epistemologische these wordt gezien, want de ToM heeft het juist niet over de Cartesiaanse asymmetrie. Dit bezwaar kan deels gerepareerd worden door sommige versies van ST, zoals die van Goldman waarbij introspectie toegang biedt tot de eigen mentale toestanden, maar het kan niet worden toegepast op de TT of alle versies van ST. De versie van ST volgens Gordon hoeft juist geen onobserveerbare entiteiten te postuleren en zou dan niet gezien hoeven te worden als een vorm van mindreading volgens de ToM, omdat basis interactie met de ander rust op directe waarneming zoals het zien van gezichtsuitdrukkingen en gebaren zonder dat daar in eerste instantie

⁷⁰ Vivian Bohl en Nivedita Gangopadhyay, "Theory of mind and the unobservability of other minds," *Philosophical Explorations* 17:2 (2014): 204-5.

⁷¹ Bohl en Gangopadhyay, "Theory of mind," 206.

mentaliseren voor nodig is.⁷²

TT weerlegt juist de Cartesiaanse asymmetrie door te stellen dat we allemaal dezelfde cognitieve mechanismen gebruiken om mentale toestanden toe te schrijven aan onszelf maar ook aan anderen, dus onze overtuigingen over de mentale toestanden van anderen hebben dezelfde epistemologische status als onze overtuigingen over onze mentale toestanden. Het is hierbij dus onnodig om versies van ToM te interpreteren als een Cartesiaanse epistemologische asymmetrie.

Bovendien is het doel van ToM⁷³ niet om een epistemologische weergave te schetsen over hoe we aan kennis komen over de mentale toestanden van de ander, maar om de cognitieve processen uit te leggen over hoe mensen mentale toestanden aan anderen toeschrijven. Dus ook al is OA een relevant probleem voor epistemologie, het is in zichzelf geen epistemologische claim. Dus de discussie of mindreading afhankelijk is van inferentiële processen of een directe perceptie gaat over de cognitieve processen die er aan onderliggend zijn in plaats van de epistemologische status van de toeschrijving van mentale toestanden.

Ten slotte kan er ook nog gesteld worden dat TT wel degelijk over observeerbare mentale toestanden kan gaan. Een veel voorkomende fout volgens Lavelle⁷⁴ is dat bestaande literatuur over directe perceptie beweert dat als er als er iets theoretisch is, het niet over iets kan gaan wat geobserveerd kan worden. Als X een theoretische entiteit is, dan moet X dus iets zijn wat niet observeerbaar is. Het ziektesymptoom gordelroos is echter juist een voorbeeld van een theoretische term die naar iets verwijst wat wel geobserveerd kan worden. Gordelroos is een theoretische term, dat betekent dat als je de betekenis ervan wilt begrijpen, je het in een bredere context van virussen, oorzaken en gevolgen moet begrijpen. Het is onmogelijk om de term gordelroos te begrijpen zonder dat je kennis van virussen hebt. Gordelroos is ook een term die iets conceptueels zegt over onze observatie: we zien rode vlekken niet als een serie van rode vlekken maar als een deel van het gordelroos virus. Theoretische entiteiten hoeven dus niet onzichtbaar te zijn, maar ze zijn alleen niet zichtbaar voor diegene die de theorie niet kennen.

⁷² Ibid., 208.

⁷³ Ibid., 207.

⁷⁴ Jane Suilin Lavelle, "Theory-Theory and the direct perception of mental states," *Rev. Phil. Psych.* 3 (2012): 215-27.

De IT stelt dus dat we een volkspychologie toe passen op momenten dat het misgaat, maar maken we niet vaker gebruik van een volkspychologie als theorie dan we denken? De reden waarom ik begrijp dat het fronsen, het naar boven kijken naar de hemel en de armen van mijn buurman die over elkaar zijn geslagen zijn betekent dat hij zich ergert, is omdat het kan zijn dat ik daar een theorie over heb geleerd. Het lijkt nog lastig te bepalen dat als je niet hebt geleerd dat mensen zich ergeren om in de regen te moeten lopen omdat ze dan nat worden, je dit via directe waarneming kan begrijpen. Hierbij lijkt het toepassen van simulatie ook lastig als er van te voren geen theorieën gegenereerd wordt om het gedrag van de buurman te begrijpen. Dit voorbeeld lijkt overigens ook niet een interactie waarbij er iets misgaat, maar juist op een voorbeeld uit het alledaagse leven. IT lijkt hier in eerste instantie goed mee om te kunnen gaan door het toepassen van de NPH, maar beweert ook dat het niet de basis kan vormen voor alledaagse interactie.

Een goede conclusie kan dan zijn dat we bij interactie ten eerste bepaalde emoties van iemands gezicht aflezen en we al snel begrijpen wat de emotie betekent, maar dat we ook vaak moeten theoretiseren om te begrijpen waarom iemand zich op een bepaalde manier gedraagt, en dat niet alleen in afwijkende situaties gebeurt, maar juist in het alledaagse leven. De TT lijkt daarom op een goede verklaring van hoe we de ander kunnen begrijpen, omdat we vaak iets van een theorie nodig hebben. De NPH zou ook een goede verklaring kunnen zijn voor het begrijpen van de ander, als het bij de basis van interactie met de ander zou horen.

7. Conclusie

Hoe kunnen we begrip hebben van het gedrag van anderen in alledaagse interactie? Dit is een interessante, en nog steeds onopgeloste, vraag in het debat over de mind. Langere tijd werden de TT en ST als posities neergezet die mindreading vaardigheden zouden kunnen verklaren. Stel dat mijn buurman bij de voordeur staat en weer fronsend met zijn armen over elkaar naar de hemel kijkt. In dat geval probeer ik een verklaring te zoeken voor zijn gedrag. Deze unieke menselijke vaardigheid werd in eerste instantie verklaard doordat we in het bezit zijn van een TT, dat betekent dat we over theorieën bezitten, aangeleerd of aangeboren, waarmee we kunnen verklaren waarom mijn buurman frons. Al snel ontstonden er al een aantal problemen met de TT zoals hoe weet ik dat ik bepaalde theorieën op iedereen kan toepassen? Ik kan er vanuit gaan dat mijn buurman een hekel heeft aan het weer, en daarom frons en met zijn armen over elkaar slaat, want dat deed hij de vorige keer ook. Logisch gezien blijkt het dat ik deze theorie weer aan hem kan toeschrijven. Verschillende bewijsgronden, zoals de false-belief test zou dan kunnen aantonen op welke leeftijd we over een ToM kunnen beschikken en overtuigingen aan anderen kunnen toeschrijven. We zouden dan volgens TT-critici wel over oneindig veel theorieën moeten beschikken en het zou niet erg efficiënt zijn om steeds weer de juiste theorie te moeten kiezen om iemand te kunnen begrijpen. De ST daarentegen beweert dat om de ander te kunnen begrijpen, we onszelf in de mentale schoenen van de ander moeten verplaatsen. Het is hierbij niet nodig om te theoretiseren en zou op cognitief gebied economischer en sneller zijn. De ontdekking van spiegelneuronen leek het bewijs voor de ST in de jaren negentig te zijn. Later bleek toch moeilijk verklaarbaar, omdat het gecompliceerd is een vergelijking te maken met de spiegelneuronen die bij het brein van aapjes zijn ontdekt toe te passen op het brein van mensen.

Een argument tegen de ST is dat er verschillende soorten gedragingen aan mentale toestanden kan worden gekoppeld, het hoeft dus niet zo te zijn dat een bepaald gedrag aan een bepaalde mentale toestand kan worden gelinkt. Er bestaat een oneindig aantal motieven, verlangens en overtuigingen dat het nogal ingewikkeld lijkt om het beperkt aantal aan simulatie ervaringen waarover je bezit toe kan passen op de ander. Daarnaast is er moeilijk een grip op te leggen dat we alleen een simulatieproces nodig hebben om iemand te kunnen begrijpen. Het lijkt vanzelfsprekender om bepaalde

theorieën te kennen om te kunnen weten waarom iemand, of zelfs een object zoals in het voorbeeld van Dennett, zich op een bepaalde manier gedraagt. Aan de ene kant zou een ST aannemelijker zijn in de vorm van een hybride theorie waarbij de TT en ST gecombineerd kunnen worden. Aan de andere kant lijkt de hybride theorie onnodig als verklaring voor alledaagse interactie omdat TT vaak al in eerste instantie kan dienen als verklaring voor het gedrag van de ander.

Als reactie op de problematiek bij TT en ST die ik in de eerste hoofdstukken illustreerde, is er recent een andere theorie, de IT ontstaan. Deze theorie stelt dat we door directe waarneming al kunnen begrijpen wat de ander bedoelt. Om terug te komen op het hierboven beschreven voorbeeld, zou dit betekenen dat op het moment ik naar mijn buurman kijk en hem zie fronsen, ik direct waarneem dat hij er geïrriteerd uitziet, en ik volgens de IT niet eerst een theorie of simulatie hoeft af te gaan om te begrijpen dat hij zich ergert aan het weer. Ik zie namelijk geen rimpels en opgehaalde wenkbrauwen op zijn voorhoofd waarvan ik de betekenis moet zoeken, maar ik kan gelijk zijn emotie waarnemen. Volgens de IT is het wel af en toe nodig om een volkspychologie te gebruiken, maar dan alleen bij interacties waar het mis gaat en daarbij hoeven we niet noodzakelijkerwijs een beroep te doen op TT of ST. Volkspychologie wordt dan geleerd door verhalen waarin duidelijk wordt wat de regels van de wereld zijn, waarbij je dus leert hoe jij en de ander zich horen te gedragen in bepaalde situaties. Het probleem bij IT is dat we dus alleen een volkspychologie gebruiken in situaties waarin we de ander niet helemaal begrijpen of waarin het misgaat. Zoals ik eerder liet zien, maken we juist vaker gebruik van theorieën om precies te voorspellen of te verklaren wat de ander bedoelt, zelfs als er wordt gesimuleerd moet er eerst iets van een theorie bekend zijn.

8. Bibliografie

- Bishop, Michael A. en Stephen M. Downes. "The theory theory thrice over: the child as scientist, superscientist or social institution?" *Stud. Hist. Phil. Sci.* 33 (2002): 121-136.
- Bloom, Paul en Tim P. German. "Two reasons to abandon the false belief task as a test of theory of mind." *Cognition* 77 (2000): 25-31.
- Bohl, Vivian en Nivedita Gangopadhyay. "Theory of mind and the unobservability of other minds." *Philosophical Explorations* 17:2 (2014): 203-222. doi: 10.1080/13869795.2013.821515.
- Brewer, Willam F., Clark A. Chinn en Ala Samarapungavan. "Explanation in scientists and children." *Minds and Machines* 8 (1998): 119-136.
- Churchland, Paul M. "Eliminative materialism and the propositional attitudes." *The Journal of Philosophy* 78:2 (1981): 67-90.
- Cosmides, Leda en John Tooby. *Mapping the mind domain specificity in cognition and culture*. Redactie Lawrence A. Hirschfeld en Susan A. Gelman. Cambridge: Cambridge University Press, 1994.
- Danforth, Bryan N. "Social insects: are ants just wingless bees?" *Current Biology* 23:22 (2013): 1011-1012. doi:10.016/j.cub.2013.10.018.
- Dennett, Daniel. *The intentional stance*. Massachusetts: MIT Press, 1998.
- Frith Chris D. en Uta Frith. "Social cognition in humans," *Current Biology* 17 (2007): 724-732. doi:10.1016/j.cub.2007.05.068.
- Gallagher, Shaun. "Simulation trouble." *Social Neuroscience* 2:3-4 (2007): 353-365. doi:10.1080/17470910601183549.
- Gallagher, Shaun en Daniel Hutto. "Understanding other through primary interaction and narrative practice." In *The Shared Mind: Perspectives on Intersubjectivity*, edited by J. Zlatev, T. Racine, C. Sinha en E. Itkonen, 17-38. Amsterdam: John Benjamins, 2008.
- Gallese Vittorio en Alvin Goldman. "Mirror neurons and the simulation theory of mind-reading." *Trends in Cognitive Sciences* 2:12 (1998): 493-501.
- Goldman, Alvin I. *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. Oxford: Oxford University Press, 2006.
- Gopnik, Alison. "Developing the idea of intentionality: children's theories of mind."

- Canadian Journal of Philosophy.*” 20:1 (1990): 89-113.
- Gopnik, Alison. “Scientific thinking in young children: theoretical advances, empirical research, and policy implications.” 337 (2012): 1623-1627. doi: 10.1126/science.1223416.
- Gopnik Alison en Henry M. Wellman. “Why the child’s theory of mind really is a theory.” *Mind & Language* 7:1&2 (1992): 145-171.
- Herschbach, Mitchell. “Folk psychological and phenomenological accounts of social perception.” *Philosophical Explorations* 11:3 (2008): 223-235. doi: 10.1080/13869790802239268.
- Hickok, Gregory. “Eight problems for the mirror neuron theory of action understanding in monkeys and humans.” *Journal of Cognitive Neuroscience* 21:7 (2008): 1229-1243.
- Lavelle, Jane Suilin. “Theory-Theory and the direct perception of mental states.” *Rev. Phil. Psych.* 3 (2012): 213-230. doi: 10.1007/s13164-012-0094-3.
- Mahy, Caitlin E.V., Louis J. Moses en Jennifer H. Pfeifer. “How and where: Theory-of-mind in the brain.” *Developmental Cognitive Neuroscience* 9 (2014): 68-81. doi: 10.1016/j.dcn.2014.01.002.
- Maiese, Michelle. *Embodiment, emotion, and cognition*. London: Palgrave Macmillan, 2011.
- Molenberghs Pascal, Christopher Brander, Jason B. Mattingley en Ross Cunnington. “The role of the superior temporal sulcus and the mirror neuron system in imitation.” *Human Brain Mapping* 31 (2010): 1316-1326. doi: 10.1002/hbm.20938.
- Perner, Joseph et al. “Thinking of mental and other representations: The roles of left and right temporo-parietal junction.” *Social Neuroscience* 1:3-4 (2006): 245-258. doi:10.1080/17470910600989896.
- Ravenscroft, Ian en Stephen Stich. “What is folk psychology?” *Cognition* 50 (1994): 447-468.
- Scholl, Brian J. en Alan M. Leslie. “Modularity, development and ‘Theory of Mind’.” *Mind & Language* 14:1 (1999): 131-153. doi: 10.1111/1468-0017.00106.
- Schurz, Matthias et al. “Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging

- studies.” *Frontiers in Human Neuroscience* 7:712 (2013): 1-14. doi: 10.3389/fnhum.2013.00712.
- Scott, Rose M. en Renée Baillargeon. “Early false-belief understanding.” *Trends in Cognitive Sciences* 21:4 (2017): 237-249. doi:10.1016/j.tics.2017.01.012.
- Spaulding, Shannon. “Mirror neurons are not evidence for the Simulation Theory.” *Synthese* 189 (2012): 515-534. doi: 10.1007/s11229-012-0086-y.
- Spaulding, Shannon. “Mirror neurons and social cognition.” *Mind & Language* 28:2 (2013): 233-257.
- Wilkerson, William S. “Simulation, theory, and the frame problem: The interpretive moment.” *Philosophical Psychology* 14:2 (2001): 141-153. doi:10.1080/09515080120051535.
- Zahavi, Dan. *Subjectivity and selfhood investigating the first-person perspective*. Cambridge, Massachusetts: MIT Press, 2005.