



Bachelorscriptie: Verwerking van gezichtsexpressies op robotgezichten

A.E.M. Leenders (4083911)

7.5 ECTS

Bachelor Liberal Arts and Sciences UU: Kunstmatige Intelligentie.

1^o beoordelaar: Dr. Marnix Naber

2^o beoordelaar: Dr. Janneke van Lith

Hoofdrichting: *Cognitieve en Neurobiologische Psychologie*

Aantal woorden: 7808

Datum: 14 April 2017

Abstract

In dit onderzoek is getracht een bijdrage te leveren aan de discussie rondom sociale robots. De onderzoeksvraag 'In hoeverre bepaalt de menselijkheid van een robot het verschil in de gemiddelde verandering van hartslag, gezichtsspieractiviteit en GSR tussen robot- en mensengezichten?' is beantwoord door middel van een experiment waarbij de activiteit van de gezichtsspier Zygomaticus Major, de gezichtsspier Corrugator Supercilii, de galvanische huidreactie(GSR) en de hartslag gemeten is. Participanten moesten in dit experiment de emotionele expressies angstig, verdrietig, neutraal, blij en boos op gezichten van mensen en verschillen soorten robots herkennen.

Uit het onderzoek blijkt dat robotkarakters in verschillende mate van menselijkheid niet adequaat van elkaar kunnen worden onderscheiden door het meten van hartslag, GSR en gezichtsspieractiviteit. Ook blijkt dat mensen fysiologisch gezien niet anders reageren op het zien van mensengezichten als op het zien van robotgezichten en dat ze qua reactietijd niet verschillen voor de verschillende karakters. De griezelig- en menselijkheid van een robot zou volgens dit onderzoek geen effect hebben op de fysiologische reactie op het zien van deze emotie. De resultaten uit dit onderzoek zouden kunnen impliceren dat het zinvol is om robots in ons sociale leven te impliceren, maar er is nog verder onderzoek nodig naar de precieze verwerking van robots.

Voorwoord

Voor u ligt het bacheloreindwerkstuk van Anne Leenders. Met dit werkstuk zal ik mijn hoofdrichting ‘Kunstmatige Intelligentie: cognitieve en neurobiologische psychologie’ bij de opleiding *Liberal Arts and Sciences* afsluiten. Dit werkstuk is een individueel werk, maar is in nauwe samenwerking geschreven met Ditte van Haalen, die dezelfde opleiding volgt. We hebben samen de onderwerpen besproken en veel overlegd over ons werkstuk, waarbij we elkaar inspireerden en aanvulden. De samenwerking is dan ook alleen maar goed en soepel verlopen. De experimenten en de voorbereidingen hiervoor zijn samen uitgevoerd, waarbij de taken eerlijk verdeeld waren. We hebben bij de uitvoering van het bacheloreindwerkstuk wel beiden onze eigen invalshoek gekozen, waarbij ik gekozen heb om mij meer te richten de *uncanny valley* op en Ditte heeft gekozen om zich meer op emoties bij robots te richten.

Het gekozen onderwerp, robots, heeft ons geen moment verveeld omdat het zo’n breed onderwerp is waar nog veel over te ontdekken valt. Het is een onderwerp dat op dit moment erg relevant is en veel verschillende invalshoeken heeft. Dit maakt dat we met veel plezier aan ons bacheloreindwerkstuk hebben gewerkt.

Tijdens het schrijven van dit bacheloreindwerkstuk hebben we ontzettend veel geleerd over het schrijven van een onderzoeksverslag, het uitvoeren van een wetenschappelijk experiment en het analyseren van data. We hebben geleerd om wetenschappelijke literatuur te zoeken, te bestuderen en te gebruiken in een wetenschappelijke context. Aan de hand hiervan hebben we geleerd hoe je een experiment kan opzetten en welke moeilijkheden hierbij komen kijken. Het was erg interessant om te leren werken met elektroden en om deze data op een computer gevisualiseerd te zien. Bij het analyseren van de data hebben we ontzettend veel geleerd over dataverwerking. We hebben nieuwe statistiekmogelijkheden ontdekt en geleerd om de resultaten te interpreteren. Ondanks af en toe frustraties maar vooral veel leermomenten, hebben we ons experiment en de analyse met tevredenheid uitgevoerd en zijn wij dan ook zeer tevreden over het resultaat.

Bij het schrijven van deze scriptie werden wij begeleid door dr. Marnix Naber, die wij via deze weg willen bedanken. Vanaf het begin af aan is de begeleiding van Marnix enthousiast, nuttig en leerzaam geweest en zijn inzet heeft ons zeker geholpen om dit werkstuk met een goed gevoel af te sluiten.

Inhoudsopgave

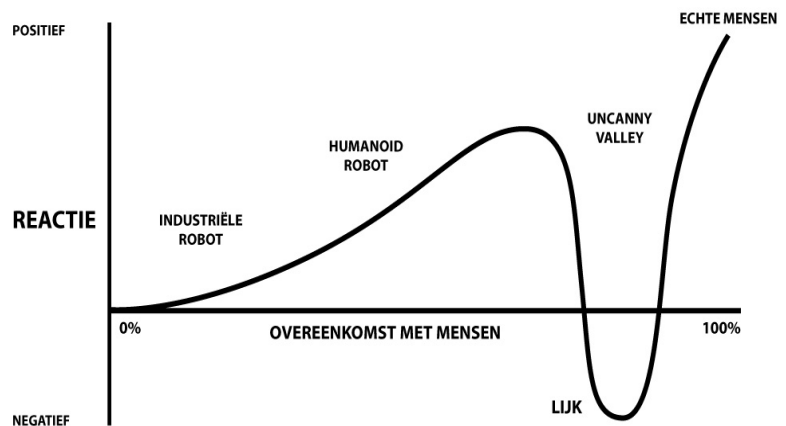
ABSTRACT	1
VOORWOORD	2
INLEIDING	4
METHODEN	8
PARTICIPANTEN	8
MEETINSTRUMENTEN, APPARATUS, STIMULI	8
<i>Meetinstrumenten</i>	8
<i>Apparatus</i>	9
<i>Stimuli</i>	9
PROCEDURE EXPERIMENT	10
STATISTISCHE ANALYSE	11
RESULTATEN	13
VERSCHILLEN IN REACTIETIJD PER ROBOTKARAKTER	13
VERSCHILLEN IN PERCENTAGE GOEDE RESPONSEN PER ROBOTKARAKTER	14
FYSIOLOGISCHE REACTIES VAN GALVANISCHE HUDRESPONS, ELEKTROMYOGRAFIE <i>ZYGOMATICUS MAJOR</i> , ELEKTROMYOGRAFIE <i>CORRUGATOR SUPERCILII</i> EN HARTSLAG PER ROBOTKARAKTER.....	15
<i>Galvanische huidrespons</i>	15
<i>Elektromyografie Zychomaticus Major</i>	16
<i>Elektromyografie Corrugator Supercilii</i>	16
<i>Hartslag</i>	17
DE VOLGORDE VAN GRIEZELIG- EN MENSELIJKHEID.....	18
VERGELIJKING VAN FYSIOLOGISCHE REACTIES OP DE VERSCHILLENDE ROBOTKARAKTERS MET MENSELIJKE STIMULI	20
CONCLUSIE EN DISCUSSIE	21
BRONNENLIJST	25
LITERATUUR	25
APPENDIX	28

Inleiding

Binnen het vakgebied kunstmatige intelligentie is de laatste jaren grote technologische vooruitgang te zien in de ontwikkeling van sociale robots. In Japan, koploper binnen de robotica, kijkt men niet eens meer op van sociale robots. Hiroshi Ishiguro ontwikkelt hier zelfs de zogenaamde *geminoid*, een type robot dat wordt gemaakt naar het evenbeeld van een individu (Nishio, Ishiguro, & Hagita, 2007). In Nederland zijn sociale robots echter nog nauwelijks terug te zien in het dagelijks leven. Wel worden er al experimenten uitgevoerd die de toegevoegde waarde van het gebruik van sociale robots onderzoeken. De ouderenzorg en de behandeling van autisme bij kinderen zijn hier voorbeelden van (Gelderblom, 2014; Kim et al., 2013). Onderliggende vraagstukken die gepaard gaan met deze maatschappelijke ontwikkeling zijn onder andere: Is de mens in staat om robot-mensinteractie op dezelfde manier te verwerken als mens-mensinteractie? Kunnen mensen op dezelfde manier empathie opwekken voor robots als zij voor mensen doen? Factoren die hierbij van belang zijn, zijn de *uncanny valley*, emotieverwerking en empathie. Wanneer sociale robots echt een rol zouden gaan spelen in onze maatschappij, is het van belang om deze vraagstukken te beantwoorden. Sociale robots kunnen dan zo natuurlijk mogelijk opgenomen worden in de maatschappij.

De eerdergenoemde *geminoid* is een voorbeeld van de trend waarin sociale robots steeds meer op mensen gaan lijken. Mori merkte echter al meer dan 45 jaar geleden op dat robots die (te) veel op mensen gaan lijken, een *eerie* (in dit onderzoek vertaald als griezelig) gevoel opwekken bij mensen. Deze aanname is van invloed op

eerdergenoemde vraagstukken omdat griezeligheid een nadelige invloed zou kunnen hebben op robot-mensinteractie. De *uncanny valley* wordt zichtbaar wanneer de *human likeness* (in dit onderzoek vertaald als menselijkheid) uitgezet wordt tegen de affiniteit die men toont tegenover een artificieel object, de negatieve piek bevindt zich bij ongeveer 70 % menselijkheid. (zie figuur 1). De menselijkheid van een robot wordt door Hanson (2006) gedefinieerd als “uitziend binnen de mogelijke natuurlijk-voorkomende verschijningen van mensen” (Mori, MacDorman, & Kageki, 2012).



Figuur 1 - Toont de *uncanny valley*, de voorgestelde relatie tussen de menselijke gedaante van een personage en de affiniteit van de waarnemer hiervoor (Bulckens, 2015).

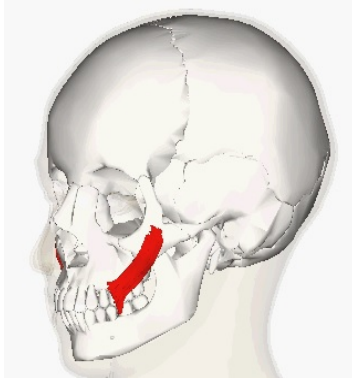
Een andere belangrijke factor in robot-mensinteractie is emotionele verwerking. Het tonen van emoties is een typische sociale eigenschap van de mens. Door onderzoek naar emotieverwerking bij mens-robotinteractie kan deze interactie natuurlijker gemaakt worden. In de literatuur zijn verschillende theorieën over emoties en emotieverwerking te vinden. Eén daarvan is de theorie van Ekman (1992), die de zes basisemoties angst, boos, verdrietig, walging, verbazing en blijdschap erkent. Deze zes emoties zijn door hun evolutionaire oorsprong universeel in alle culturen en voor elke soort (Ekman, Friesen, & Ellsworth, 2013). Omdat elke emotie zijn eigen evolutionair doel dient, hebben zij ook hun eigen fysiologische kenmerken, geïnitieerd door het autonome zenuwstelsel (Levenson, Ekman, & Friesen, 1990; Rainville, Bechara, Naqvi, & Damasio, 2006).

Een derde cruciale factor bij het ontwikkelen van sociale robots is empathie. Onderzoek van onder andere Oberman, McCleery, Ramachandran & Pineda (2007) wijst erop dat de spiegelneuronen in mensen mogelijk actief worden bij zowel het zien van robot acties als het zien van menselijke acties. Dit is essentieel voor sociale interactie en empathie (Gazzola, Rizzolatti, Wicker, & Keysers, 2007). Een ander belangrijk aspect van empathie is het onbewust mimeren van de expressie van een ander die emotie toont, ook bij statische stimuli (De Wied, Van Boxtel, Zaalberg, Goudena, & Matthys, 2006; Dimberg, Thunberg, & Elmehed, 2000; Lang, P. Greenwald, M. K., Bradley, M. M., Hamm, 1993; Sonny-Borgström, 2002). Dit mimeren gebeurt snel, automatisch en grotendeels onbewust (Dimberg et al., 2000).

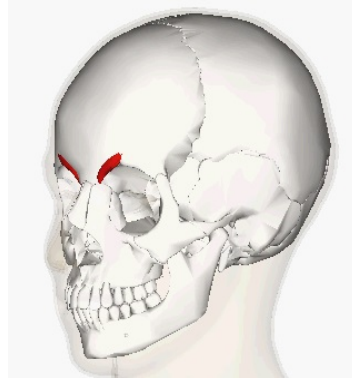
Het is dus mogelijk om objectief reacties op emoties te meten, omdat elke emotie zijn eigen patroon van reacties door het autonome zenuwstelsel kent en we onbewust expressies van anderen imiteren door empathie. In dit onderzoek is ervoor gekozen om de hartslag en *Galvanic Skin Response* (GSR) te meten en een elektromyografie (EMG) van het gezicht uit te voeren. Op deze manier wordt zowel de fysiologische als de empathische reactie op emoties gemeten. De emoties angst, boos, verdrietig en blijdschap worden onderzocht omdat deze emoties het meest gebruikelijk zijn in dit onderzoeksveld (Bos, Jap-Tjong, Spencer, & Hofman, 2016; Levenson et al., 1990; Schwartz, Weinberger, & Singer, 1981).

Dimberg & Thunberg (1998) hebben een onderzoek gedaan naar het gebruik van EMG om gezichtsuitdrukkingen te meten en concludeerden dat verschillende emotionele uitdrukkingen ook verschillende EMGresultaten lieten zien. Deze fysiologische reacties waren al meetbaar na slechts 300-400 ms na de blootstelling aan de stimuli.

De *Musculus Zygomaticus Major* (zie figuur 2), of grote jukspier, is een aangezichtsspier die grotere activiteit vertoont bij gelukkige gezichtsexpressies dan bij de boze stimuli. De *Musculus Corrugator Supercilii* (zie figuur 3), of wenkbrauw spier, is een kleine, smalle piramidevormige spier in



Figuur 2 - *Musculus Zygomaticus Major* (BodyParts3D, 2017).



Figuur 3 - *Musculus Corrugator Supercilii* (BodyParts3D, 2017).

de buurt van het oog en vertoont grotere spieractiviteit in boze gezichtsuitdrukkingen. Door middel van het meten met EMG van veranderingen in deze spierbewegingen in het gezicht kan een onderscheid gemaakt worden tussen positieve en negatieve emoties. De sterkte en richting van de reactie van deze spieren geeft aan om welke emotie het gaat, zo reageert de *Zygomaticus Major* sterker op

blij stimuli dan op boze stimuli en reageert de *Corrugator Supercilii* over het algemeen sterker op boze dan op blij gezichten (Dimberg & Thunberg, 1998).

De hartslag is een maat die veranderingen in het autonome zenuwstelsel registreert. Uit onder andere het onderzoek van Li & Chen (2006) en Critchley, Rotshtein, Nagai, O'doherty, Mathias, & Dolan (2005) blijkt dat emoties accuraat onderscheiden kunnen worden door de hartslag te meten. De samentrekking van het hart door de tijd heen geeft een bepaald patroon. Dit patroon verschilt per emotie. De veranderingen, die volgens het James-Lange model samenhangen met de mate van *arousal*, in dit patroon zijn groter voor de emoties boos, angst en verdrietig dan voor de emotie blij (Collet, Vernet-Maury, Delhomme, & Dittmar, 1997).

Een andere fysiologische maat om veranderingen in het autonome zenuwstelsel te registreren is de GSR. Dit is een veelgebruikte maat die geschikt is voor het meten van verschillen in *arousal* en de bijbehorende emoties, opgewekt door emotionele stimuli (Haag, Goronzy, Schaich, & Williams, 2004; Rosenthal-von der Pütten, Krämer, Hoffmann, Sobieraj, & Eimler, 2013). Uit onderzoek blijkt dat het patroon van de GSR wisselend is voor verschillende emoties, blij en boze emoties wekken bijvoorbeeld een sterke GSR reactie op, terwijl walging samenhangt met een veel minder sterke reactie (Collet et al., 1997). Een stijging in de GSR wordt veroorzaakt door een verhoogde mate van *arousal*, wat in overeenstemming is met het onderzoek van onder andere Collet et al., (1997). Dit onderzoek toont aan dat er een grotere stijging in de GSR is voor negatieve emoties dan voor positieve emoties en dat het hierbinnen vooral mogelijk is om emotie paren van elkaar te onderscheiden (Collet et al., 1997).

Onderzoek met een combinatie van EMG van het gezicht, hartslagmeting, en GSR zorgt ervoor dat emotieverwerking zo breed mogelijk gemeten kan worden. De fysiologische maten vullen elkaar aan waardoor de emoties nog beter van elkaar te onderscheiden zijn. Dit is een veelgebruikte combinatie, zoals te zien is in een overzichtsstudie naar psychofysiologische signalen. Deze studie laat zien dat deze drie meetmethodes vaak en in verschillende combinaties met elkaar voorkomen (Jerritta, Murugappan, Nagarajan, & Wan, 2011).¹

Zoals eerder beschreven, is het *uncanny valley* effect een van de factoren die een rol speelt bij robot-mensinteractie. De menselijkheid kan worden beschouwd vanuit verschillende perspectieven, waarvan statische en dynamische uitstraling er twee zijn. De *Fusiform Face Area* (FFA) is één van de hersengebieden die betrokken is bij het herkennen van gezichten en de menselijke voorkeur om gezichten te herkennen in niet-menselijke objecten (Kanwisher, McDermott, & Chun, 1997). Gezichten worden dus wel herkend, maar worden de gezichtsexpressies bij robots ook hetzelfde verwerkt als menselijke gezichtsexpressies? En bepaalt de menselijkheid van een robot of de gezichtsexpressie als meer griezelig wordt ervaren wanneer een robot zich in de *uncanny valley* zou bevinden? Er zijn meerdere onderzoeken gedaan naar de verschillende perspectieven van de menselijkheid en naar de psychologische validiteit van Mori's hypothese (Mori et al., 2012). Het *uncanny effect* wordt inderdaad waargenomen, maar op welke wijze de esthetiek van de robot invloed heeft op dit effect blijft nog onbeantwoord (Hanson et al., 2005; Seyama & Nagayama, 2007).

Om een bijdrage te kunnen leveren aan deze discussie wordt in dit onderzoek specifiek gekeken naar de menselijkheid van de robots, hun relatie met de *uncanny valley* en empathie. In dit onderzoek wordt getracht de volgende vraag te beantwoorden: *In hoeverre bepaalt de menselijkheid van een robot het verschil in de gemiddelde verandering van hartslag, gezichtsspieractiviteit en GSR tussen robot- en mensengezichten?* De volgende hypothesen zijn gesteld:

1. Robots kunnen adequaat onderscheiden worden van mensen door middel van het meten van hartslag, GSR en gezichtsspieractiviteit.
2. Hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen.

¹ Bovenstaande is geschreven in samenspraak met Ditte van Haalen.

Methoden²

Participanten

Er deden 14 vrouwelijke participanten en 3 mannelijke participanten mee aan het experiment. De participanten zijn geworven via de kenniskring van de onderzoekers en werden gecompenseerd met proefpersoon uren als zij hiervoor in aanmerking kwamen. De andere participanten werkten aan het onderzoek mee op vrijwillige basis. Er waren 3 participanten die niet konden worden meegenomen in het onderzoek vanwege data die niet was opgeslagen en één participant kon niet worden meegenomen in het onderzoek omdat zij tijdens het onderzoek in slaap was gevallen. Er waren voor de data-analyse dus nog 13 participanten over in de leeftijd van 19 tot en met 24 ($M = 20.941$, $SD = 1.003$).

Meetinstrumenten, Apparatus, Stimuli

Meetinstrumenten

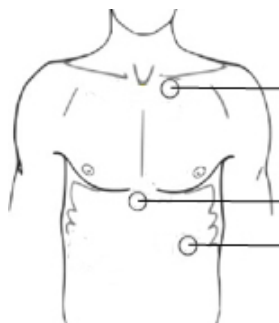
Om de gezichtsspieractiviteit, de hartslag en de galvanische huidreactie (GSR) te meten zijn de volgende invoermogelijkheden van Bio-Semi gebruikt.

De *Galvanic Skin Response* bestaat uit 2 passieve Nihon Kohden elektrodes die werden geplaatst op de linker wijs- en middelvinger zoals afgebeeld in figuur 4.

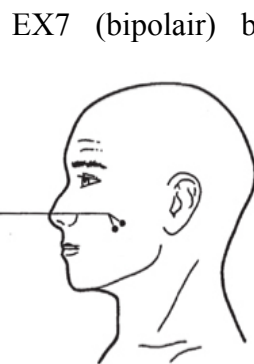
De *Bio-Semi Flat type active-elektrodes* bestaan uit 7+2 elektrodes genaamd EX1-EX7 plus CMS/DRL. EX1-EX3 maten de hartslag van de participant op respectievelijk de ribbenkast, onder het sternum en onder het rechter sleutelbeen zoals afgebeeld in figuur 5. De gezichtsspieractiviteit is gemeten door EX4 en EX5 (bipolair) te plaatsen bij de linkerwang (*Zygomaticus Major*) en EX6-



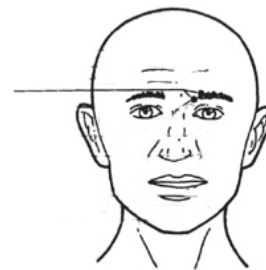
Figuur 4 - Plaatsing van de elektrodes voor de GSR metingen op de linker wijs- en middelvinger (Harmon-Jones & Beer, 2012).



Figuur 5 - Plaatsing van elektroden op de romp voor de hartslag meting. Rechtsonder werd EX1 geplaatst. Bij de sternum, het middenrif, werd EX2 geplaatst en bij het sleutelbeen werd EX3 geplaatst.



Figuur 6 - Plaatsing van de elektroden voor de Zygomaticus Major. De onderste stip is EX4 en de bovenste is EX5 (Harmon-Jones & Beer, 2012).



Figuur 7 - Plaatsing van de elektroden voor de Corrugator Supercilii. De onderste linker stip is EX6 en de bovenste stip is EX7 (Harmon-Jones & Beer, 2012).

EX7 (bipolair) boven de linker wenkbrauw (*Corrugator Supercilii*). DRL- en CMS-elektroden werden geplaatst op het voorhoofd (zie figuur 6 en 7).

² Alle analyses zijn uitgevoerd in samenwerking met Ditte van Haalen

Apparatus

Het *Active Two Biosemi* apparaat (256 channel AD-box met batterij en *Flat-Type Active-electrodes*) is gebruikt om de gezichtsspieractiviteit, de hartslag en de galvanische huidreactie (GSR) gelijktijdig te kunnen meten.

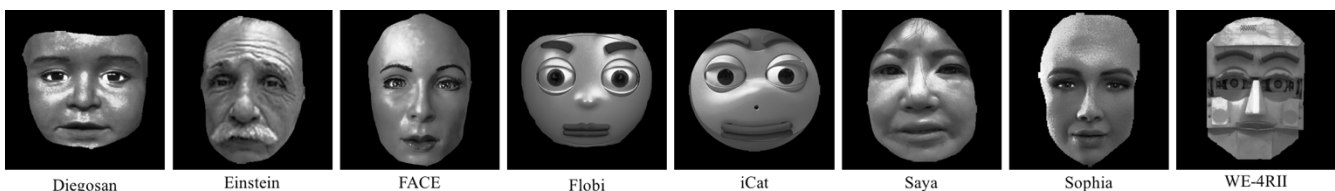
Het beeldscherm (ASUS, 27 inch) dat werd gebruikt om de stimuli op te tonen heeft een schermresolutie van 2560x1440 en een vernieuwingssnelheid van 59Hz. Het experiment is door Dr. Marnix Naber geprogrammeerd in Matlab (2012). Het numpad van het toetsenbord werd gebruikt om een reactie te geven. De participanten moesten hun hoofd op een hoofdsteun, een zogenaamde *chinrest*, leggen, zodat ze niet zouden bewegen (zie figuur 8).



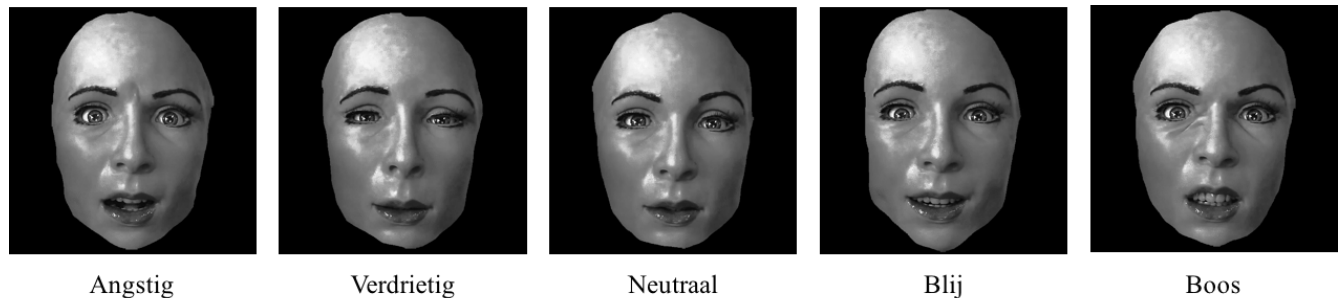
Figuur 8 - Proefopstelling met hoofd in de chinrest.

Stimuli

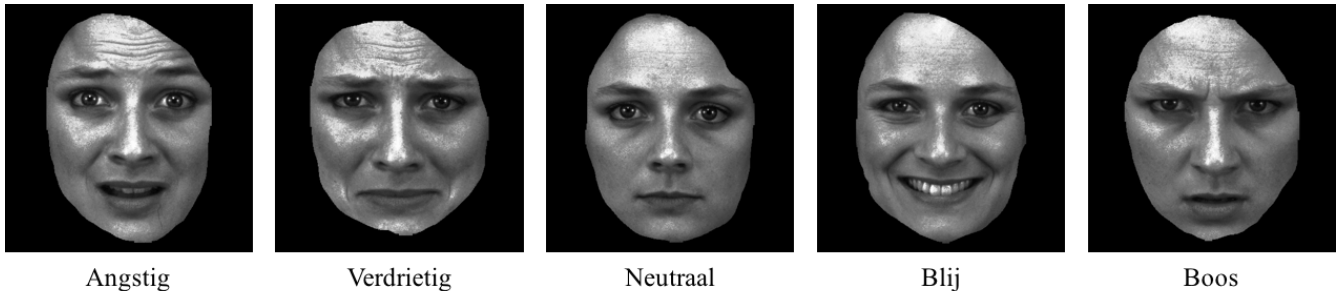
Er zijn 8 statische afbeeldingen van robotkarakters gebruikt in dit experiment waarvan de oren en het haar zijn weggewerkt. De gebruikte robots in dit onderzoek zijn: Diegosan, Einstein, FACE, Flobi, iCat, Saya, Sophia en WE-4RII (figuur 9). Hetzelfde aantal menselijke gezichten (vier mannelijke en vier vrouwelijke) is gebruikt waarvan ook de oren en het haar zijn weggewerkt. De gezichten toonden een angstige, verdrietige, neutrale, blijde of boze uitdrukking, zoals weergegeven in figuur 10 voor robot FACE en figuur 11 voor een vrouwelijk gezicht. Niet elke robot beschikte over alle gezichtsuitdrukkingen. In tabel 1 in de appendix staat een overzicht van de gebruikte emotionele expressies per robot. De gebruikte stimuli zijn afkomstig uit de dataset van overkoepelend onderzoek van Dr. Marnix Naber. De afbeeldingen zijn gecorrigeerd voor verschillen in luminantie en resolutie, met behulp van histogram egalisatie is dit uitgevoerd in Matlab (2012). De afmetingen van alle afbeeldingen zijn aangepast tot 800 x 800 pixels en de achtergrondkleur van het beeldscherm was [116 116 116].



Figuur 9 - Alle verschillende robotkarakters met een neutrale gezichtsuitdrukking. De namen staan afgebeeld onder iedere individuele karakter.



Figuur 10 - Voorbeeldstimuli met robotkarakter FACE in alle verschillende emoties.



Figuur 11 - Vrouwelijke voorbeeldstimuli in alle verschillende emoties.

Procedure experiment

Alle participanten gingen voorafgaand aan het experiment akkoord met de *informed consent* en wisten voorafgaand aan het experiment niets over het specifieke doel van het experiment. Hierbij werd ook verteld dat ze tijdens het experiment niet mochten bewegen. Ze moesten met hun hoofd steunen in de hoofdsteun en hun linkerhand moest rusten op een daarvoor bestemd kussentje. Wanneer een participant vroeg wat er gemeten werd, werd verteld dat er hersenactiviteit gemeten werd om te voorkomen dat de gezichtsspieractiviteit van de participant werd beïnvloed. De participanten werden individueel getest in een laboratorium op de Universiteit Utrecht. In dit laboratorium werden zoals boven beschreven de elektroden op het gereinigde gezicht, de vingers en de romp aangebracht.

De participanten kregen eerst tien test-stimuli te zien om aan het experiment te wennen, waar zij

ook feedback kregen op hun respons. Deze testtrials werden ook meegenomen in de dataverwerking. Eerst werd een fixatiepunt getoond gedurende een gerandomiseerde tijd tussen 0.5 en 2 seconden, daarna werd de stimulus 6 seconden getoond. Vervolgens had de participant 2 seconden de tijd om aan te geven welke emotie hij gezien had door op een toets op het numpad te drukken. De emoties waren als volgt ingedeeld op het numpad: 2 - *angstig*; 4 - *verdrietig*; 5 - *neutraal*; 6 - *blij*; 8 - *boos* (zie figuur 12). Hierna werd weer een fixatiepunt getoond waarmee de cyclus weer opnieuw begon (zie figuur 13). Elke stimulus werd 3 keer herhaald, zodat er in totaal



Figuur 12 - Numpad met bijbehorende emoties weergegeven op de toetsen.

228 trials inclusief testtrials waren. De stimuli werden in gerandomiseerde volgorde getoond zodat de volgorde geen invloed kon hebben op de fysiologische reactie van de participant.

Na het tonen van de 228 stimuli werden subjectieve vragen gesteld in het 2e deel van het experiment. Eerst werden vragen gesteld over alle personages, daarna alleen over de robots. De participanten konden antwoord geven op deze vragen door middel van een 100-

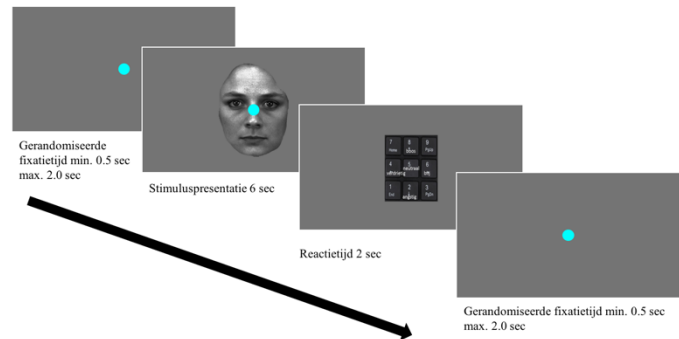
puntsschaal. De vragen die de participanten moesten beantwoorden over alle personages zijn: *Ik vind het gezicht onaantrekkelijk - aantrekkelijk; ik vind het gezicht kunstmatig - natuurlijk; ik vind het gezicht griezelig - geruststellend; ik vind het gezicht onmenselijk - menselijk.*

De vragen die de participanten moesten beantwoorden over alleen de robots waren: *Ik zou het comfortabel vinden als deze robot mij helpt in het huishouden; ik zou het comfortabel vinden als deze robot mij op de hoogte brengt van informatie zoals nieuwsberichten; ik zou het comfortabel vinden als deze robot mijn gesprekspartner is; ik zou het comfortabel vinden als deze robot mij helpt bij het studeren; ik zou het comfortabel vinden als deze robot mijn financiën en administratie bijhoudt.*

Statistische analyse

Voor de analyse van de resultaten werden Matlab (2012), IBM SPSS Statistics versie 23.0 (IBM Corp. Released, 2015). en Microsoft Excel gebruikt. De gestandaardiseerde scores van de gemiddelde verandering van GSR (tijdspad van 6-8 seconden), EMG *Zygomaticus Major* (tijdspad van 1-6 seconden), EMG *Corrugator Supercilii* (tijdspad van 1-6 seconden) en hartslag (tijdspad van 1-6 seconden) werden gebruikt voor de analyses, omdat er een grote spreiding was. De metingen werden gefilterd voor ruis met een *low-pass* en *high-pass* filter. Er werd gekeken naar de respons emotie van de participant in plaats van de getoonde emotie omdat verwacht werd dat de fysiologische reactie van de participant gebaseerd is op de perceptie van de getoonde emotie. Alle trials waarbij te laat respons is gegeven, zijn niet meegenomen in de analyse.

Om een algemeen beeld te krijgen van de resultaten werden de reactietijd en het percentage goede responsen bekeken. Hoewel dit experiment geen test was om de reactietijd en accuraatheid van de participanten te meten, zou dit beiden iets kunnen zeggen over de verwerking van de stimuli. Het is



Figuur 13 - Procedure van het experiment met een cyaan fixatiepunt, gevolgd oor de stimulus en later de tijd om te reageren terwijl de afbeelding van het numpad is weergegeven op het scherm. Daarna wordt weer een fixatiepunt zichtbaar totdat de volgende stimulus verschijnt.

daarom van belang om deze informatie ook mee te nemen in het onderzoek. Er is daarom een niet-parametrische herhaalde metingen rangtekentoets van Wilcoxon uitgevoerd tussen de verschillende reactietijden van de robots en de menselijke stimuli, alsook het gemiddelde van alle robots samen en alle menselijke stimuli samen. Ook werd het percentage goede responsen bij de verschillende robots vergeleken met het percentage goede responsen bij menselijke stimuli, net als het gemiddelde van de robots in het algemeen en de menselijke stimuli. Beide vergelijkingen werden uitgevoerd door een gepaarde t-toets.

Om de eerste hypothese '*Robots kunnen adequaat onderscheiden worden van mensen door middel van het meten van hartslag, GSR en gezichtsspieractiviteit*' te toetsen werd voor iedere fysiologische maat een herhaalde metingen éénweg ANOVA uitgevoerd, met de 8 robotkarakters als factor.

Om de tweede hypothese '*Hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen*' te toetsen werd naar de antwoorden op de subjectieve vragen gekeken. Als eerst is de data gecorrigeerd, er bestonden waarden van onder de 0 en boven de 100, dit had te maken met een technisch aspect van het geprogrammeerde experiment. De scores onder de 0 zijn gecorrigeerd naar 0 en de scores boven de 100 werden vervangen door 100. Een factoranalyse van de antwoorden op vraag 1 t/m 9 op de neutrale robotstimuli werd uitgevoerd om te kijken of er meerdere factoren te onderscheiden zijn. Eerst werden de robots aan de hand van gemiddelde antwoorden op de subjectieve vragen gerangschikt naar gemiddelde score. Daarna werden de robots gerangschikt op griezeligheid naar aanleiding van de gemiddelde score van factor 2. Vervolgens werden gepaarde metingen t-toetsen uitgevoerd per fysiologische maat om ieder robotkarakter apart te vergelijken met de gemiddelde reactie op de menselijke stimuli. Dit werd alleen uitgevoerd met de neutrale stimuli, omdat de te beantwoorden onderzoeksvraag onafhankelijk is van de getoonde emotie.

Resultaten³

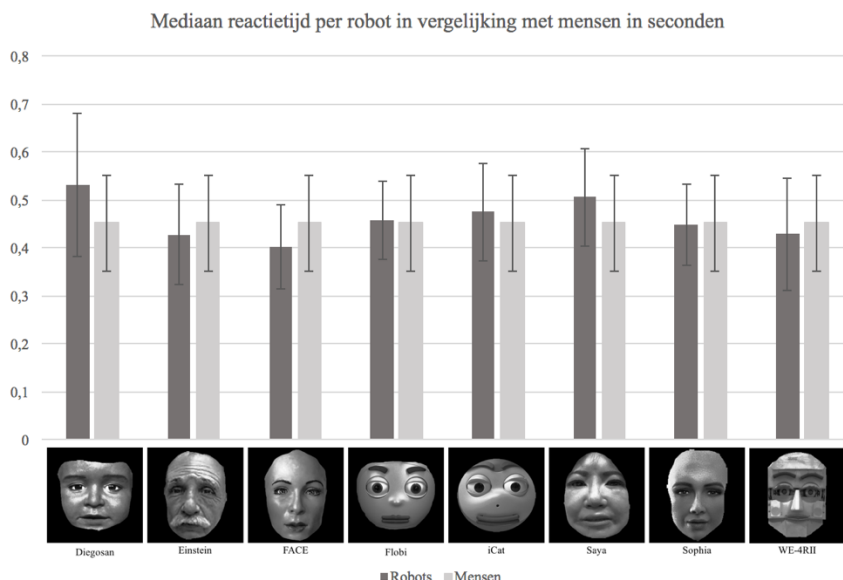
Verschillen in reactietijd per robotkarakter

Om te onderzoeken of mensen hetzelfde reageren op robots als op mensen, zijn de reactietijden met elkaar vergeleken. De mediaan van de reactietijd voor menselijke stimuli was 0.453 ($IQR = 0.220$) seconden en de mediaan voor de reactietijd voor robotstimuli was 0.443 ($IQR = 0.200$). Een niet-parametrische herhaalde metingen rangtekentoets van Wilcoxon toonde aan dat deze tijden niet significant van elkaar verschilden, waarbij $Z = -0.489$ en $p = 0.625$.

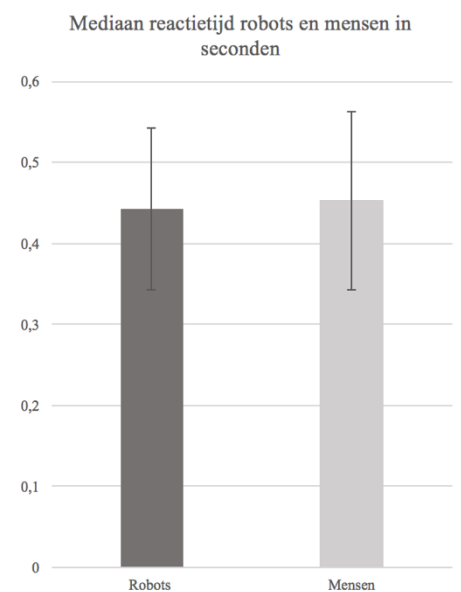
Daarnaast zijn er ook niet-parametrische herhaalde metingen rangtekentoetsen van Wilcoxon uitgevoerd tussen de reactietijden voor de verschillende robotstimuli om te onderzoeken of hier een verschil tussen zat, dit zou ook een verschil in verwerking kunnen aantonen.

Hieruit bleek dat Diegosan ($Mdn = 0.531$, $IQR = 0.298$), Einstein ($Mdn = 0.428$, $IQR = 0.209$), FACE ($Mdn = 0.403$, $IQR = 0.175$), Flobi ($M = 0.459$, $SD = 0.162$), iCat ($Mdn = 0.475$, $IQR = 0.204$), Saya ($Mdn = 0.506$, $IQR = 0.201$), Sophia ($Mdn = 0.448$, $IQR = 0.169$) en WE-4RII ($Mdn = 0.429$, $IQR = 0.233$) allemaal onderling niet verschilden in mediaan van reactietijd in seconden. Deze resultaten zijn zichtbaar gemaakt in figuur 14 en figuur 15. In tabel 2 en tabel 3 in de appendix zijn deze resultaten samengevat.

Uit deze testen bleek dus dat mensen over het algemeen niet sneller of langzamer reageerden op de verschillende robotstimuli.



Figuur 14 - Mediaan reactietijd per robot in vergelijking met mensen in seconden. Geen enkele robot wordt significant sneller herkend dan een andere robot.



Figuur 15 - Mediaan reactietijd voor robotstimuli en menselijke stimuli in seconden. De reactietijden verschillen niet significant van elkaar.

³ De resultaten zijn verkregen verwerkt in samenwerking met Ditte van Haalen.

Verschillen in percentage goede responsen per robotkarakter

Om verder te onderzoeken of mensen hetzelfde reageren op robots als op mensen, is het percentage goede responsen vergeleken, hiermee worden de juist herkende emoties in de gezichten van de stimuli bedoeld. Het gemiddelde percentage goede responsen voor robotstimuli was 70.58 % ($SD = 7.418$) en het gemiddelde percentage roede responsen voor menselijke stimuli was 87.76 % ($SD = 4.192$). Door middel van een gepaarde t-toets met een α van 0.05 is aangetoond dat robotstimuli gemiddeld slechter worden herkend dan menselijke stimuli, waarbij $t(12) = -7.327, p < 0.001$ en $d = -2.936$.

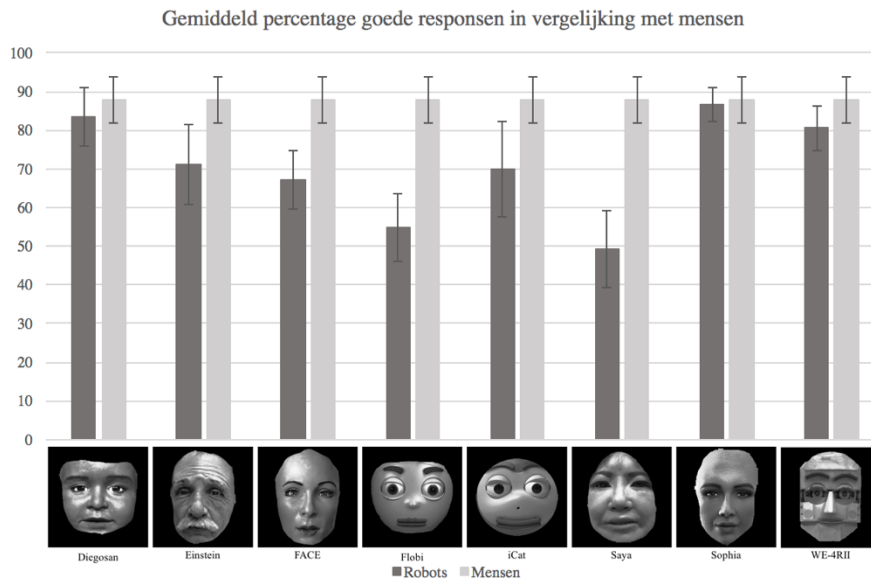
Daarnaast werden er ook gepaarde t-toetsen uitgevoerd om het percentage van de goede responsen voor alle verschillende robotkarakters te vergelijken met elkaar. Het bleek dat een aantal robots beter werden herkend dan anderen.

Diegosan ($M = 83.333, SD = 15.210$) werd beter herkend dan Einstein ($M = 71.154, SD = 20.590$) $t(12) = 2.333, p = 0.038, d = 0.673$, FACE ($M = 67.179, SD = 15.012$) $t(12) = 3.659, p = 0.003, d = 1.069$, Flobi ($M = 54.872, SD = 17.460$) $t(12) = 3.949, p = 0.002, d = 1.738$ en Saya ($M = 49.359, SD = 19.971$) $t(12) = 4.971, p < 0.001, d = 1.914$. Diegosan werd niet significant beter herkend dan iCat ($M = 69.872, SD = 24.892$), niet slechter dan Sophia ($M = 86.667, SD = 9.027$) en niet beter dan WE-4RII ($M = 80.513, SD = 11.373$).

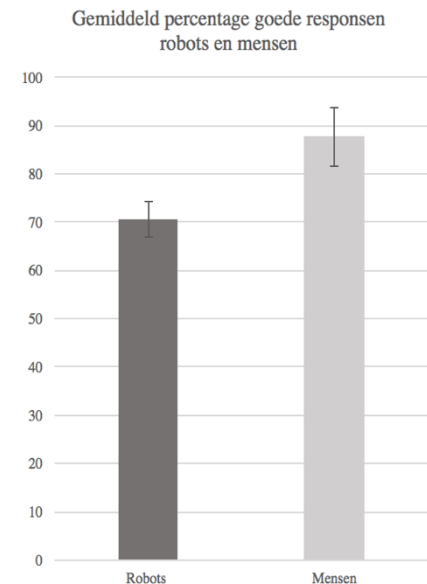
Einstein werd niet significant beter herkend dan FACE, Flobi, iCat en WE-4RII. Einstein werd significant beter herkend dan Saya, $t(12) = 3.461, p = 0.005, d = 1.075$ en slechter herkend dan Sophia, $t(12) = -2.665, p = 0.021, d = -0.976$. FACE werd niet significant beter herkend dan Flobi en iCat. FACE werd wel beter herkend dan Saya, $t(12) = 2.416, p = 0.033, d = 1.009$. Ook werd FACE slechter herkend dan Sophia, $t(12) = -4.531, p = 0.001, d = -1.573$. en WE-4RII, $t(12) = -2.523, p = 0.027, d = -1.001$. Flobi werd niet significant beter herkend dan iCat en Saya. Flobi werd slechter herkend dan Sophia ($M = 86.667, SD = 9.027$), $t(12) = -4.993, p < 0.001, d = -2.288$. en WE-4RII $t(12) = -4.963, p < 0.001, d = -1.740$. iCat werd beter herkend dan Saya $t(12) = 2.587, p = 0.024, d = 0.909$ en slechter herkend dan Sophia, $t(12) = 2.318, p = 0.039, d = -0.897$. iCat werd niet significant beter herkend dan WE-4RII. Saya werd slechter herkend dan Sophia, $t(12) = -5.713, p < 0.001, d = -2.407$ en WE-4RII, $t(12) = -5.072, p < 0.001, d = -1.917$. Sophia werd niet significant beter herkend dan WE-4RII.

Deze resultaten zijn zichtbaar gemaakt in figuur 16 en figuur 17, in tabel 4 en tabel 5 in de appendix zijn deze resultaten samengevat.

Uit deze resultaten bleek dus dat mensen over het algemeen emoties van menselijke stimuli beter kunnen herkennen dan die van niet menselijke stimuli, hoewel WE-4RII daar weer een tegenvoorbeeld van is, WE-4RII werd ook erg goed herkend.



Figuur 15 - Gemiddeld percentage goede responsen per robot in vergelijking met mensen. Geen enkele robot werd significant sneller herkend dan een andere robot.



Figuur 17 - Gemiddeld percentage goede responsen voor robotstimuli en menselijke stimuli. De reactietijden verschilden niet significant van elkaar. Er wordt gemiddeld minder goed herkend welke emotie werd getoond bij robots, dan bij mensen.

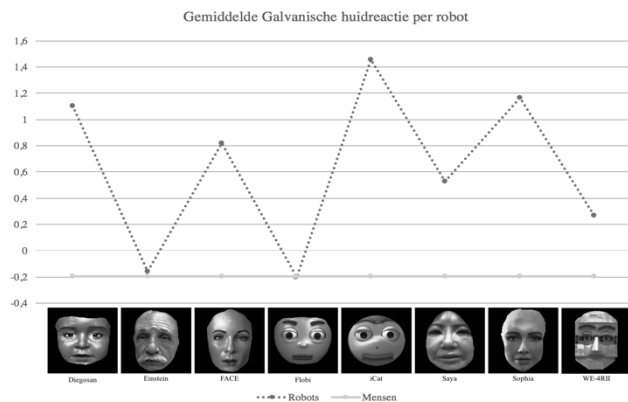
Fysiologische reacties van Galvanische huidrespons, Elektromyografie *Zygomatikus Major*, Elektromyografie *Corrugator Supercilii* en Hartslag per robotkarakter.

Om de eerste hypothese, 'Robots kunnen adequaat onderscheiden worden van mensen door middel van het meten van hartslag, GSR en gezichtsspieractiviteit' te toetsen werd een herhaalde metingen éénwegANOVA uitgevoerd per fysiologische maat. Hiermee werd getoetst of er een verschil zat tussen de fysiologische reacties voor de verschillende karakters van de robots. De gemiddelde reacties van de robots per fysiologische maat zijn samengevat in tabel 6 in de appendix, de resultaten van deze herhaalde metingen éénwegANOVA zijn samengevat in tabel 7 in de appendix.

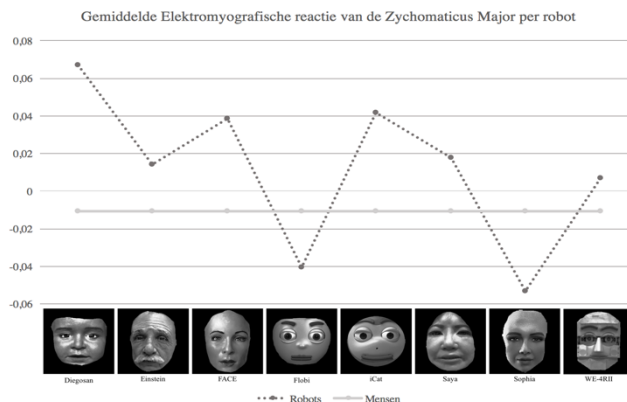
Galvanische huidrespons

Boxplots en Shapiro-Wilk statistiek toonden aan dat de assumptie van normaliteit geschonden was; de homogeniteit van varianties was niet aangetoond, $F_{\max} = 10.584$. Volgens Mauchly's test ($p < 0.001$) kon sphericity niet worden aangenomen, de Huynh-Feldt aanpassing werd vervolgens gebruikt.

De resultaten van de herhaalde metingen éénwegANOVA lieten geen significant verschil zien tussen de reactie op de verschillende robotkarakters $F(4.061, 48.728) = 1.132$ en $p = 0.353$. De gemiddelde galvanische huidreacties op de robotkarakters zijn weergegeven in figuur 18. Het bleek dus dat de robots niet adequaat onderscheiden konden worden van elkaar door middel van het meten van de galvanische huidrespons reactie tussen 6 en 8 seconden na het tonen van de stimuli.



Figuur 18 - Gemiddelde galvanische huidreactie per robotkarakter en het totaal van alle menselijke stimuli. Er bleek geen significant verschil tussen de robots onderling te zijn.



Figuur 19 - Gemiddelde Elektromyografische reactie van de Zychomaticus Major reactie per robotkarakter en het totaal van alle menselijke stimuli. Er bleek geen significant verschil tussen de robots onderling te zijn.

Elektromyografie Zychomaticus Major

Boxplots en Shapiro-Wilk statistiek toonden aan dat de assumptie van normaliteit geschonden was; de homogeniteit van varianties was aangetoond, $F_{\max} = 5.238$. Volgens Mauchly's test ($p = 0.070$) kon sphericity worden aangenomen, er was dus geen aanpassing in de analyse nodig.

De resultaten van de herhaalde metingen éénwegANOVA lieten geen significant verschil zien tussen de reactie op de verschillende robots $F(7, 84) = 0.703$ en $p = 0.669$. De gemiddelde elektromyografische reacties van de *Zychomaticus Major* op de robotkarakters zijn weergegeven in figuur 19. Het bleek dus dat de robots niet adequaat onderscheiden konden worden van elkaar door middel van het meten van de gezichtsspierreacties van de *Zychomaticus Major* tussen 1 en 6 seconden na het tonen van de stimuli.

Elektromyografie Corrugator Supercilii

Boxplots en Shapiro-Wilk statistiek toonden dat de assumptie van normaliteit net geschonden was, Diegosan $p = 0.048$ en Einstein $p = 0.046$; de homogeniteit van varianties was aangetoond, $F_{\max} = 8.484$. Volgens Mauchly's test ($p = 0.327$) kon sphericity worden aangenomen, ook hier was dus geen aanpassing in de analyse nodig.

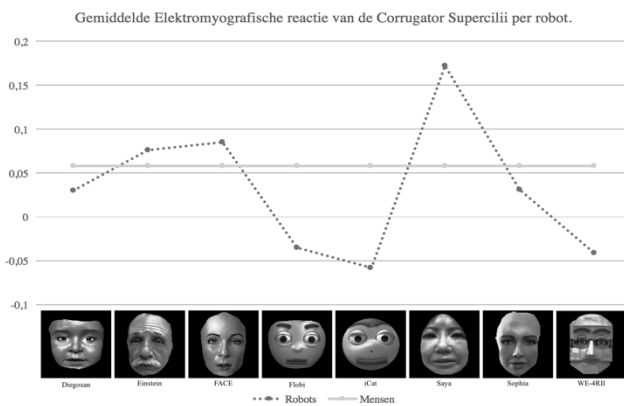
De resultaten van de herhaalde metingen éénwegANOVA geen significant verschil zien tussen de reactie op de verschillende robotkarakters $F(7, 84) = 1.272$ en $p = 0.227$. De gemiddelde

elektromyografische reacties van de *Corrugator Supercilii* op de robotkarakters zijn weergegeven in figuur 20. Het bleek dus dat de robots niet adequaat onderscheiden konden worden van elkaar door middel van het meten van de gezichtsspierreacties van de *Corrugator Supercilii* tussen 1 en 6 seconden na het tonen van de stimuli.

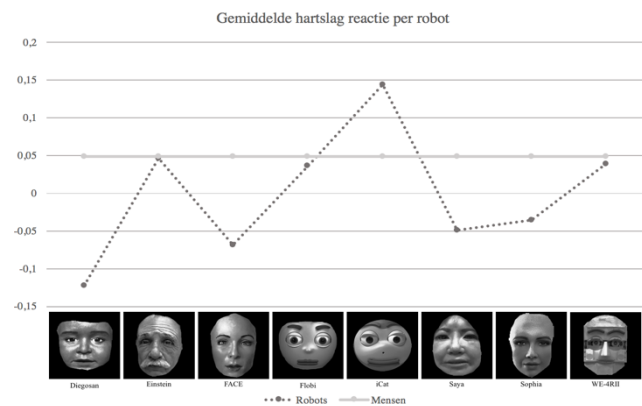
Hartslag

Boxplots en Shapiro-Wilk statistiek toonden dat de assumptie van normaliteit aangenomen kan worden; de homogeeniteit van varianties was aangetoond, $F_{\max} = 6.486$. Volgens Mauchly's test ($p = 0.168$) kan sphericity worden aangenomen, ook hier was dus geen aanpassing in de analyse nodig.

De resultaten van de herhaalde metingen éénwegANOVA lieten geen significant verschil zien tussen de reactie op de verschillende robotkarakters $F(7, 84) = 1.096$ en $p = 0.373$. Het bleek dus dat de robots niet adequaat onderscheiden konden worden van elkaar door middel van het meten van de gemiddelde hartslag verandering tussen 1 en 6 seconden na het tonen van de stimuli.



Figuur 20 - Gemiddelde Elektromyografische reactie van de *Corrugator Supercilii* reactie per robotkarakter en het totaal van alle menselijke stimuli. Er bleek geen significant verschil tussen de robots onderling te zijn.

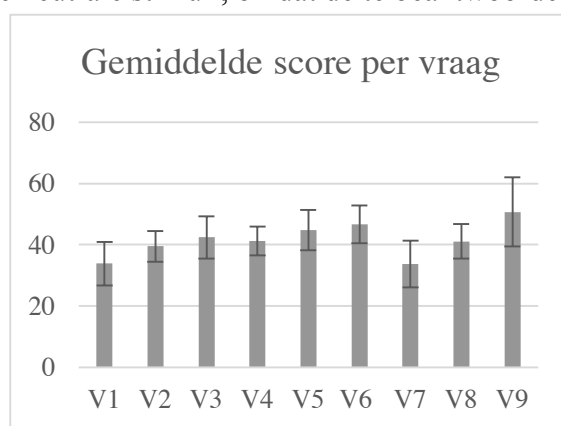


Figuur 21 - Gemiddelde hartslag reactie per robotkarakter en het totaal van alle menselijke stimuli. Er bleek geen significant verschil tussen de robots onderling te zijn.

De volgorde van griezelig- en menselijkheid

Om het eerste deel van de tweede hypothese ‘*Hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen*’ te toetsen werd naar de antwoorden op de subjectieve vragen gekeken. Een factoranalyse van de antwoorden op vraag 1 t/m 9 op de robotstimuli werd uitgevoerd om te kijken of er meerdere factoren te onderscheiden zijn. Eerst werden de robots aan de hand van gemiddelde antwoorden op de subjectieve vragen gerangschikt naar gemiddelde score. Daarna werden de robots gerangschikt op griezeligheid naar aanleiding van de gemiddelde score van factor 2. Vervolgens werden gepaarde metingen t-toetsen uitgevoerd per fysiologische maat om ieder robotkarakter apart te vergelijken met de gemiddelde reactie op de menselijke stimuli. Dit werd alleen uitgevoerd met de neutrale stimuli, omdat de te beantwoorden onderzoeksvraag onafhankelijk is van de getoonde emotie.

Na de correctie van de antwoorden buiten de scorering van 0-100 werd gecontroleerd op normaliteit van de antwoorden en bleek alleen vraag 8 (zie tabel 8 in de appendix) niet normaal verdeeld te zijn. Omdat dit alleen vraag 8 betrof, is gekozen te werken met de gemiddelde antwoorden per vraag en is hiermee een factoranalyse uitgevoerd, de gemiddelden zijn afgebeeld



Figuur 22 - Gemiddelde score per subjectieve vraag op een schaal van 0 tot 100 met standaardafwijking. Het gemiddelde is berekend over alle antwoorden op de robotstimuli met neutrale gezichten. V1 t/mV9 staan respectievelijk voor vraag 1 tot en met vraag 9

Tabel 9

Structuurmatrix van de correlatie na promax rotatie die de vragen binnen elke factor met elkaar hebben.

Vraag	Factor		
	1	2	3
V1	-	-	0.894
V2	-	-	0.572
V3	-	0.957	-
V4	-	0.929	-
V5	0.893	-	-
V6	0.680	-	-
V7	0.923	-	-
V8	0.850	-	-
V9	-	-	0.823

in figuur 22.

Een KMO test

gaf aan dat de data niet goed genoeg was voor een factor

analyse. $p = 0.349$. Barlett's test of sphericity geeft aan dat de

data wel goed was voor aan factor analyse: $\chi^2(36) = 74.531$

en $p = < 0.001$. Omdat er niets te doen was aan het kleine aantal

participanten is besloten wel een factoranalyse uit te voeren.

Uit de factoranalyse kwamen 3 factoren die voldoen aan het

Kaiser criterium met een eigenwaarde van > 1 . Deze drie

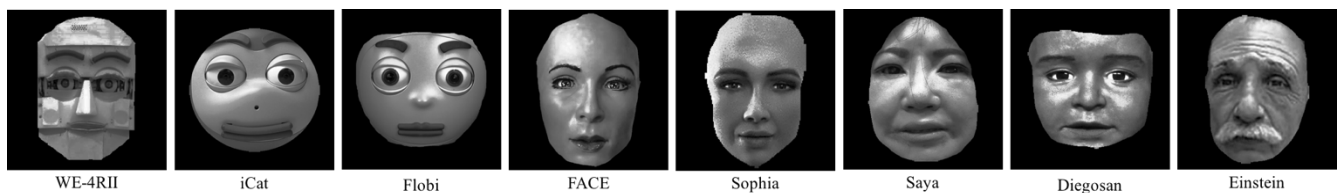
factoren verklaren samen de variantie voor 82.31%. Er is een

structuurmatrix opgesteld van waarin de correlatie van de

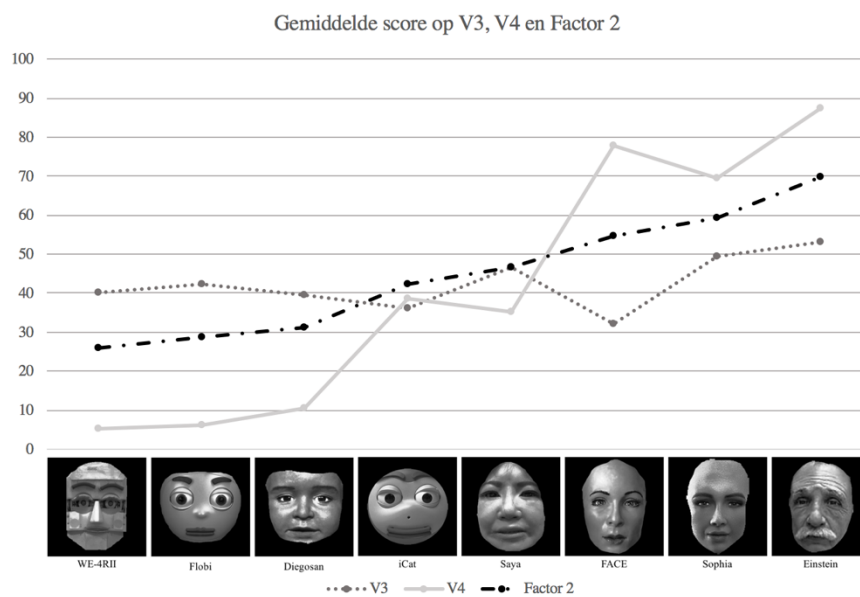
negen vragen binnen de factoren is samengevat in tabel 9 in de

appendix.

Factor 2 is de factor die V3 en V4 heeft gereduceerd. Een lage waarde betekent griezelig en onmenselijk, een hoge waarde betekent geruststellend en menselijk. Aan de hand van deze factor zou een volgorde opgesteld kunnen worden van de verschillende robots die overeen zou komen met de *uncanny valley*. Hoe hoger de gemiddelde score op deze factor, hoe menselijker deze dan zou moeten zijn en hoe geruststellender. Om deze reden zijn twee verschillende reeksen opgesteld; de gemiddelde score op alle vragen in figuur 23 en de gemiddelde score op alleen factor 2 in figuur 24 in tabel 10 in de appendix staan de gemiddelden per robot per vraag genoteerd. In de eerste figuur zijn de robotkarakters



op volgorde geplaatst van laag naar hoog op de gemiddelde score van alle vragen. In de tweede figuur zijn de robotkarakters op volgorde geplaatst van laag naar hoog op de gemiddelde score van alleen factor 2, daarnaast zijn de gemiddelde scores op V3, V4 en Factor 2 weergegeven. Hoe hoger de score op factor 2, hoe groter het verschil tussen V4 en V3. Hierdoor wordt dus zichtbaar dat de meest griezelige en menselijke karakter FACE is.



Figuur 24 - Gemiddelde score op V3, V4 en Factor 3 met de robotkarakters op volgorde van 'laag scorend op Factor 2' naar 'hoog scorend op Factor 2'

Vergelijking van fysiologische reacties op de verschillende robotkarakters met menselijke stimuli

Om het tweede deel van de tweede hypothese '*Hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen*' werd nogmaals gekeken naar de fysiologische reacties van de participanten op de verschillende stimuli. Er werden gepaarde metingen t-toetsen per fysiologische maat uitgevoerd om ieder robotkarakter apart te vergelijken met de gemiddelde reactie op de menselijke stimuli. Dit werd alleen uitgevoerd met de neutrale stimuli, omdat de te beantwoorden onderzoeksvraag onafhankelijk is van de getoonde emotie. De gemiddelde reacties van de robotkarakters en van de menselijke stimuli die zijn verwerkt in de gepaarde metingen t-toets per fysiologische maat zijn samengevat in tabel 11 in de appendix, de resultaten zijn samengevat in tabel 12 in de appendix.

Uit de resultaten van de gepaarde metingen t-toetsen bleek dat het voor geen enkele robot uitmaakt hoe griezelig en menselijk de robot is, de fysiologische reacties van de participanten was bij geen enkele vergelijking significant.

Conclusie en Discussie⁴

In dit onderzoek is getracht een bijdrage te leveren aan de discussie rondom sociale robots en de verwerking van emotionele expressies hiervan. De centrale vraag in dit onderzoek, *‘In hoeverre bepaalt de menselijkheid van een robot het verschil in de gemiddelde verandering van hartslag, gezichtspieractiviteit en GSR tussen robot- en mensengezichten?’* is beantwoord door middel van een experiment waarbij hartslag-, EMG- en GSRmetingen met elektrodes gedaan zijn. Hoewel er al veel onderzoek gedaan is naar emotieverwerking en sociale robots, is er nog weinig onderzoek gedaan dat deze factoren combineert. Dit explorierend onderzoek tracht hier een begin mee te maken. De hypothesen in dit onderzoek zijn:

1. Robots kunnen adequaat onderscheiden worden van mensen door middel van het meten van hartslag, GSR en gezichtsspieractiviteit.
2. Hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen.

In dit onderzoek zijn uiteindelijk slechts 13 metingen meegenomen in de analyse. Vanwege de korte duur van dit bachelorproject was het niet mogelijk om meer tijd te besteden aan het uitvoeren van het experiment. Voor de analyses die zijn uitgevoerd zijn eigenlijk een veel hoger aantal metingen nodig om tot goede, betrouwbare resultaten te leiden. Een poweranalyse wijst uit dat er een minimumaantal van 40 participanten nodig is voor significante resultaten. De kans dat de verkregen resultaten in dit onderzoek op toeval berusten is dan ook erg groot. Dit kan verklaren waarom sommige resultaten elkaar tegenspreken.

De gemiddelde reactietijd voor robotgezichten verschilt niet significant van de gemiddelde reactietijd voor menselijke gezichten, ook verschillen de individuele robot karakters niet van het gemiddelde van de menselijke gezichten. Als gekeken wordt naar het aantal goede responsen, scoren de participanten het best op robots Diegosan, Sophia en WE-4RII. Opvallend hieraan is dat Diegosan en Sophia beiden erg menselijk zijn, maar dat WE-4RII juist een mechanisch uiterlijk heeft doordat het ontwerp geen egale ‘huid’ heeft. Wel zijn de emotionele expressies van de stimuli van WE-4RII erg duidelijk en misschien wel overdreven te noemen. Dat zou het makkelijker kunnen maken voor de

⁴ In samenspraak met Ditte van Haalen is het algemene deel van de discussie geschreven, de structuur is ook samen opgezet. De rest van de discussie is eigen werk.

participanten om de getoonde emoties te herkennen. Dit kan duiden op het nut van het menselijker maken van een robot als het doel van de robot sociale interactie is: emoties worden namelijk beter herkend (Dautenhahn, 2007). Het kan erop wijzen dat de menselijkheid van de robot invloed heeft op de emotionele perceptie in sociale interactie. Zoals eerdergenoemd, is er nog vrij weinig literatuur over dit onderwerp gepubliceerd zodat noch bevestigd noch tegengesproken kan worden dat deze resultaten op toeval berusten.

Het blijkt niet mogelijk om robots adequaat te kunnen onderscheiden van mensen door te kijken naar de gemiddelde verandering in hartslag, GSR en gezichtsspieractiviteit van de *Corrugator Supercilii* en *Zygomaticus Major*. Ook onderling zijn de robots niet te onderscheiden op deze fysiologische maten. Het feit dat de minder menselijke robots ook niet te onderscheiden zijn van de menselijkere robots is consistent met het resultaat van de analyse tussen alle robots en alle mensen. Deze resultaten zouden elkaar echter tegenspreken wanneer de meest menselijke robots wel andere fysiologische reacties bij de participanten zouden opwekken dan de minst menselijke robots, terwijl er geen verschil zit tussen de robotgezichten en menselijke gezichten in het algemeen.

Hoewel er in dit onderzoek wel enkele niet significante verschillen waar te nemen zijn, is de kans groot dat deze resultaten op toeval berusten omdat de herhaalde metingen éénwegANOVA beschrijft dat er in het geheel geen significant verschil in fysiologische reactie is. Als er een grotere steekproef was geweest, waren er wellicht meer significante waarden uitgekomen. Zoals al eerder genoemd is deze kans door het geringe aantal participanten erg aannemelijk. Er is nauwelijks onderzoek uitgevoerd naar deze specifieke vraag, waardoor de uitkomsten niet te vergelijken zijn met de resultaten van andere onderzoeken. Daarnaast is het voor de experimentleiders de eerste keer dat zij een experiment met elektroden uitvoeren, dit kan een negatieve invloed hebben op de betrouwbaarheid van de data. Doordat de analyse geen enkel significant resultaat geeft, moet de eerste hypothese verworpen worden.

Factor 2 neemt de resultaten van V3 en V4 samen. Een lage waarde betekent griezelig en onmenselijk, een hoge waarde betekent geruststellend en menselijk. Dit is echter niet in overeenstemming met Mori's hypothese van de *uncanny valley*, die stelt dat hoe menselijker een robot is, hoe griezeliger deze is. Als kanttekening moet daarbij wel gesteld worden dat de participant de stimulus wel moet herkennen als robot. Dat deze twee vragen als factor uit de factoranalyse zijn gekomen, kan worden verklaard doordat de participanten dat robots niet als robot herkennen. Ook zijn er alleen statische en zwart-witte stimuli getoond, wat ook invloed zal hebben op deze resultaten. Voor vervolgonderzoek valt te overwegen om de stimuli in kleur weer te geven. Dynamische stimuli zullen in

alle waarschijnlijkheid ook meer verschil in reactie opwekken. De verschillscores van V3 en V4 geven een betere volgorde van de verwachte griezelligheid weer. Of dat het geval is, is niet duidelijk en niet te achterhalen.

Deze resultaten zijn ook in overeenstemming met de uitkomst van de gepaarde t-toetsen, die aantonen dat alle fysiologische reacties op de verschillende robots niet significant verschillen met de gemiddelde fysiologische reacties op menselijke gezichten. Als menselijkheid een belangrijk aspect zou zijn in het verwerken van emotionele gezichtsuitdrukkingen, zouden hier wel verschillen verwacht worden. De hypothese, hoe griezelig- en menselijker de robot, hoe minder de gemiddelde reactie van de fysiologische maat vergelijkbaar is met die van de reactie op mensen, moet aan de hand van de resultaten verworpen worden.

Bij de uitvoering van experiment worden ook enkele beperkingen beschreven. De gebruikte robotstimuli zijn af en toe onduidelijk voor de participanten of ze roepen associaties op die niets te maken hebben met de robots. De robot Einstein wordt door participanten betrouwbaarder gevonden omdat Einstein geassocieerd wordt met een slim, betrouwbaar persoon. Niet iedereen herkent Einstein echter als Einstein en ook niet als robot. De robot Diegosan, een kind, roept juist een minder vertrouwd gevoel op omdat mensen het een kind vinden. Een oplettende participant merkte op dat het ongemakkelijk voelt om de administratie te laten doen door een baby, dat is robot Diegosan. Diezelfde participant herkende Einstein als zodanig en zei hier dan ook over dat het wel gewenst is hem de administratie te laten doen. Dit is duidelijke feedback voor vervolgonderzoek, want de vraag bleek hierdoor niet valide voor elke robot. Deze vraag is minstens door deze ene participant beantwoord door invloed van andere kenmerken van de robot dan kenmerken die met de *uncanny valley* te maken hebben.

Het experiment zelf duurt ook erg lang: de participanten moeten een half uur stilzitten. De meeste participanten gaven aan erg slaperig te worden, één van de participanten is zelfs in slaap gevallen en deze gegevens hebben we dus niet mee kunnen nemen. Deze slaapverwekkendheid van het experiment kan de resultaten ook beïnvloeden.

Voor vervolgonderzoek kunnen al deze factoren meegenomen worden. Het belangrijkste is dat vervolgonderzoek meer participanten test en zich of toespitst op minder verschillende robots of een analyse uitvoert waarbij deze verschillen worden meegenomen. Hiermee kan er een duidelijker resultaat behaald worden.

De onderzoeksvraag ‘*In hoeverre bepaalt de menselijkheid van een robot het verschil in de gemiddelde verandering van hartslag, gezichtsspieractiviteit en GSR tussen robot- en mensengezichten?*’

is dan als volgt te beantwoorden: De gemiddelde verandering in hartslag, gezichtsspieractiviteit van de Corrugator Supercilii en Zygomaticus Major en GSR wordt in dit onderzoek niet bepaald door de menselijkheid van de robots. Mensen reageren dus fysiologisch hetzelfde op robots met verschillende mate van menselijkheid, als op mensen. De reactietijd en percentage goede responsen verschillen niet, even als de fysiologische reacties op de verschillende robots ongeacht de menselijkheid van deze robots. Deze resultaten zetten dan ook geen kracht bij de hypothese van Mori over de *uncanny valley*. In dit explorerend onderzoek zijn wel vast de eerste valkuilen ontdekt en hopelijk is hiermee de weg vrij gemaakt voor meer onderzoek.

Bronnenlijst

Literatuur

- Bos, P. A., Jap-Tjong, N., Spencer, H., & Hofman, D. (2016). Social context modulates facial imitation of children's emotional expressions. *Plos One*, *11*(12), e0167991. <https://doi.org/10.1371/journal.pone.0167991>
- Collet, C., Vernet-Maury, E., Delhomme, G., & Dittmar, A. (1997). Autonomic nervous system response specificity to basic emotions. *Journal of the Autonomic Nervous System*, *62*, 45–57.
- Critchley, H. D., Rotshtein, P., Nagai, Y., O'Doherty, J., Mathias, C. J., & Dolan, R. J. (2005). Activity in the human brain predicting differential heart rate responses to emotional facial expressions. *NeuroImage*, *24*(3), 751–762. <https://doi.org/10.1016/j.neuroimage.2004.10.013>
- Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human--robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 679–704.
- De Wied, M., Van Boxtel, A., Zaalberg, R., Goudena, P. P., & Matthys, W. (2006). Facial EMG responses to dynamic emotional facial expressions in boys with disruptive behavior disorders. *Journal of Psychiatric Research*, *40*(2), 112–121. <https://doi.org/10.1016/j.jpsychires.2005.08.003>
- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, *39*(1), 39–45.
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. <https://doi.org/10.1111/1467-9280.00221>
- Ekman, P. (1992). Are there basic emotions?
- Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). *Emotion in the human face: Guidelines for research and an integration of findings*. Elsevier.
- Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage*, *35*(4), 1674–1684. <https://doi.org/10.1016/j.neuroimage.2007.02.003>
- Gelderblom, G. J. (2014). Zorgrobotica. In J. van Hoof & E. J. M. Wouters (Eds.), *Het verpleeghuis van de toekomst is (een) thuis* (pp. 75–77). Houten: Bohn Stafleu van Loghum. https://doi.org/10.1007/978-90-368-0693-0_21

- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. *Affective Dialogue Systems, i*, 36–48. https://doi.org/10.1007/978-3-540-24842-2_4
- Hanson, D. (2006). Exploring the aesthetic range for humanoid robots. *Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*, 39–42.
- Hanson, D., Olney, A., Prilliman, S., Mathews, E., Zielke, M., Hammons, D., ... Stephanou, H. (2005). Upending the uncanny valley. In *Proceedings of the national conference on artificial intelligence* (Vol. 20, p. 1728).
- IBM Corp. Released. (2015). IBM SPSS Statistics for Windows, Version 23.0. 2015.
- Jerritta, S., Murugappan, M., Nagarajan, R., & Wan, K. (2011). Physiological signals based human emotion recognition: a review. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on* (pp. 410–415).
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311.
- Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders, 43*(5), 1038–1049. <https://doi.org/10.1007/s10803-012-1645-2>
- Lang, P. Greenwald, M. K., Bradley, M. M., Hamm, A. O. (1993). Looking at pictures-afffective, facial, visceral and behavioral reactions, 261–273.
- Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27*(4), 363–384.
- Li, L., & Chen, J. (2006). Emotion recognition using physiological signals. *Advances in Artificial Reality and Tele-Existence - Lecture Notes in Computer Science, 4282*, 437–446. <https://doi.org/10.1007/11941354>
- MATLAB. (2012). MATLAB. *MATLAB*. <https://doi.org/10.1201/9781420034950>
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley. *IEEE Robotics and Automation Magazine, 19*(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>

- Nishio, S., Ishiguro, H., & Hagita, N. (2007). Geminoid: Teleoperated android of an existing person. *Humanoid Robots: New Developments*, (June), 343–352. <https://doi.org/10.5772/4876>
- Oberman, L. M., McCleery, J. P., Ramachandran, V. S., & Pineda, J. A. (2007). EEG evidence for mirror neuron activity during the observation of human and robot actions: Toward an analysis of the human qualities of interactive robots. *Neurocomputing*, 70(13–15), 2194–2203. <https://doi.org/10.1016/j.neucom.2006.02.024>
- Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, 61(1), 5–18. <https://doi.org/10.1016/j.ijpsycho.2005.10.024>
- Rosenthal-von der Pütten, A. M., Krämer, N. C., Hoffmann, L., Sobieraj, S., & Eimler, S. C. (2013). An Experimental Study on Emotional Reactions Towards a Robot. *International Journal of Social Robotics*, 5(1), 17–34. <https://doi.org/10.1007/s12369-012-0173-8>
- Schwartz, G. E., Weinberger, D. a., & Singer, J. a. (1981). Cardiovascular differentiation of happiness, sadness, anger, and fear following imagery and exercise. *Psychosomatic Medicine*, 43(4), 343–364. <https://doi.org/10.1097/00006842-198108000-00007>
- Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, 16(4), 337–351.
- Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, 43(5), 433–443. <https://doi.org/10.1111/1467-9450.00312>

Figuren

- Bulkens, B. (2015, 23 December). Uncanny Valley Nederlands De Animatier [Illustratie]. Geraadpleegd op 16 Februari 2017 van <http://www.animatier.nl/kerstfilms-en-de-uncanny-valley/>
- "BodyParts3D, © The Database Center for Life Science licensed under CC Attribution-Share Alike 2.1 Japan." Geraadpleegd op 16 Februari 2017.
- Harmon-Jones, E., & Beer, J. S. (Eds.). (2012). *Methods in social neuroscience*. Guilford Press. Geraadpleegd op 28 Februari 2017.

Appendix

Tabel 1*Overzicht van aanwezige robotstimuli per emotie*

Stimulus	Emotie				
	<i>Angstig</i>	<i>Verdrietig</i>	<i>Neutraal</i>	<i>Blij</i>	<i>Boos</i>
Diegosan	x	x	x	x	
Einstein	x		x	x	x
FACE	x	x	x	x	x
Flobi	x	x	x	x	x
iCat	x		x	x	x
Saya		x	x	x	x
Sophia	x	x	x	x	x
WE-4RII	x	x	x	x	x

*Notitie: x = De emotionele expressie is aanwezig in de stimulus set.***Tabel 3***Resultaten niet-parametrische herhaalde metingen rangtekentoets van Wilcoxon.*

Stimulus		<i>Z</i>	<i>p-waarde</i>
Diegosan	Einstein	-1.573	0.116
	FACE	-1.572	0.116
	Flobi	-1.782	0.075
	iCat	-1.119	0.263
	Saya	-0.943	0.345
	Sophia	-1.782	0.075
	WE-4RII	-1.713	0.087
Einstein	FACE	-0.804	0.422
	Flobi	-0.559	0.576
	iCat	-0.035	0.972
	Saya	-1.153	0.249
	Sophia	-0.524	0.600
	WE-4RII	-0.734	0.463
FACE	Flobi	-0.105	0.916
	iCat	0.000	1.000
	Saya	-0.664	0.507
	Sophia	-0.035	0.972
	WE-4RII	-0.419	0.675
Flobi	iCat	-0.943	0.345
	Saya	-1.398	0.162
	Sophia	-0.454	0.650
	WE-4RII	-0.035	0.972
iCat	Saya	-0.594	0.552
	Sophia	-0.874	0.382
	WE-4RII	-0.524	0.600
Saya	Sophia	-1.363	0.173
	WE-4RII	-1.013	0.311
Sophia	WE-4RII	-0.140	0.889

*Notitie: De significantie is tweezijdig getest.***Tabel 2***menselijke stimuli samen.*

Stimulus	<i>n</i>	<i>Mdn</i>	<i>IQR</i>
Diegosan	13	0.531	0.298
Einstein	13	0.428	0.209
FACE	13	0.403	0.175
Flobi	13	0.459	0.162
iCat	13	0.475	0.204
Saya	13	0.506	0.201
Sophia	13	0.448	0.169
WE-4RII	13	0.429	0.233
Robots	13	0.443	0.200
Mensen	13	0.453	0.220

Tabel 4*Gemiddeld percentage goede responsen per robot en voor alle menselijke stimuli samen.*

Stimulus	<i>n</i>	<i>M</i>	<i>SD</i>
Diegosan	13	83.333	15.210
Einstein	13	71.154	20.590
FACE	13	67.179	15.012
Flobi	13	54.872	17.460
iCat	13	69.872	24.892
Saya	13	49.359	19.971
Sophia	13	86.667	9.027
WE-4RII	13	80.513	11.373
Robots	13	70.584	7.418
Mensen	13	87.763	4.192

Notitie: De gemiddelde scores staan in percentage.

Tabel 5

Resultaten gepaarde metingen *t*-test met vrijheidsgraden, *p*-waarde en voor alle metingen een effectgrootte.

Stimulus		<i>t</i> -waarde	<i>df</i>	<i>p</i> -waarde	Cohen's <i>d</i>
Diegosan	Einstein	2.333	12	0.038*	0.673
	FACE	3.659	12	0.003*	1.069
	Flobi	3.949	12	0.002*	1.738
	iCat	2.048	12	0.0630	0.653
	Saya	4.971	12	0.000*	1.914
	Sophia	-0.749	12	0.468	-0.267
	WE-4RII	0.428	12	0.676	0.210
Einstein	FACE	0.800	12	0.439	0.221
	Flobi	2.166	12	0.051	0.853
	iCat	0.185	12	0.856	0.056
	Saya	3.461	12	0.005*	1.075
	Sophia	-2.665	12	0.021*	-0.976
	WE-4RII	-1.295	12	0.220	-0.563
FACE	Flobi	1.692	12	0.116	0.756
	iCat	-0.379	12	0.711	-0.131
	Saya	2.416	12	0.033*	1.009
	Sophia	-4.531	12	0.001*	-1.573
	WE-4RII	-2.523	12	0.027*	-1.001
Flobi	iCat	-1.906	12	0.081*	-0.698
	Saya	0.746	12	0.470	0.294
	Sophia	-4.993	12	0.000*	-2.288
	WE-4RII	-4.963	12	0.000*	-1.740
iCat	Saya	2.587	12	0.024*	0.909
	Sophia	-2.318	12	0.039*	-0.897
	WE-4RII	-1.357	12	0.200	-0.550
Saya	Sophia	-5.713	12	0.000*	-2.407
	WE-4RII	-5.072	12	0.000*	-1.917
Sophia	WE-4RII	1.350	12	0.202	0.599

Notitie: De significantie is tweezijdig getest. Een * betekent het resultaat is significant. Er is gebruik gemaakt van een α van 0.05

Tabel 6

Gemiddelde fysiologische reacties van GSR, EMG Zygomaticus Major, EMG Corrugator Supercilii en Hartslag per robot.

Fysiologische maat	Stimulus	<i>n</i>	<i>M</i>	<i>SD</i>
GSR	Diegosan	13	1.104	2.467
	Einstein	13	-0.160	1.153
	FACE	13	0.816	2.068
	Flobi	13	-0.200	3.718
	iCat	13	1.455	3.914
	Saya	13	0.531	2.420
	Sophia	13	1.165	3.106
	WE-4RII	13	0.267	1.498
EMG Zygomaticus	Diegosan	13	0.067	0.257
	Einstein	13	0.014	0.144
	FACE	13	0.039	0.194
	Flobi	13	-0.041	0.095
	iCat	13	0.042	0.103
	Saya	13	0.018	0.144
	Sophia	13	-0.053	0.276
	WE-4RII	13	0.007	0.073
EMG Corrugator	Diegosan	13	0.031	0.266
	Einstein	13	0.08	0.199
	FACE	13	0.09	0.324
	Flobi	13	-0.035	0.225
	iCat	13	-0.058	0.223
	Saya	13	0.173	0.498
	Sophia	13	0.032	0.301
	WE-4RII	13	-0.041	0.137
Hartslag	Diegosan	13	-0.12	0.292
	Einstein	13	0.047	0.256
	FACE	13	-0.069	0.226
	Flobi	13	0.036	0.281
	iCat	13	0.144	0.494
	Saya	13	-0.049	0.281
	Sophia	13	-0.036	0.189
	WE-4RII	13	0.039	0.258

Tabel 7
fysiologische reacties van GSR, EMG Zygomaticus Major, EMG Corrugator

Fysiologische maat	Mauchly's $\text{sp}F$	df	p -waarde
GSR	< 0.001	1.132 8.728, 4.061)	0.353
EMG Zychomaticus	0.070	0.703 (84, 7)	0.669
EMG Corrugator Supr	0.327	1.272 (84, 7)	0.227
Hartslag	0.168	1.096 (84, 7)	0.373

Notitie: Het aantal vrijheidsgraden voor Mauchly's test is 27. Wanneer de sphericity niet kan worden angenommen, is een Huynh-Feldt aanpassing gedaan en is hiermee de herhaalde metingen éénweg ANOVA uitgevoerd.

Tabel 8

Mediaan reactietijden per robot en voor alle menselijke stimuli samen.

Vraagnummer met vraag in tekst	n	M	SD
V1 <i>Ik vind het gezicht onaantrekkelijk - aantrekkelijk.</i>	13	33.836	14.351
V2 <i>Ik vind het gezicht kunstmatig - natuurlijk.</i>	13	39.469	10.108
V3 <i>Ik vind het gezicht griezelig - geruststellend.</i>	13	42.465	13.754
V4 <i>Ik vind het gezicht onmenselijk - menselijk.</i>	13	41.287	9.445
V5 <i>Ik zou het comfortabel vinden als deze robot mij helpt in het huishouden.</i> <i>Ik zou het comfortabel vinden als deze robot mij op de hoogte brengt van</i>	13	44.838	13.218
V6 <i>informatie zoals nieuwsberichten.</i>	13	46.723	12.242
V7 <i>Ik zou het comfortabel vinden als deze robot mijn gesprekspartner is.</i>	13	33.689	15.085
V8 <i>Ik zou het comfortabel vinden als deze robot mij helpt bij het studeren.</i> <i>Ik zou het comfortabel vinden als deze robot mijn financiën en</i>	13	41.057	11.29
V9 <i>administratie bijhoudt</i>	13	50.678	22.647

Notitie: Bij de eerste vier vragen staat links het begrip genoteerd dat bij de 0-waarde op de schaal hoort, en rechts het woord dat bij 100 hoort als waarde op de schaal.

Tabel 10

Gemiddelde scores per vraag per robotkarakter

Vraag	n	Robotkarakter															
		Diegosan		Einstein		FACE		Flobi		iCat		Saya		Sophia		WE-4RII	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
V1	13	42.548	24.059	20.030	14.343	51.196	24.105	27.206	23.974	25.307	28.503	23.095	20.720	58.017	26.386	23.293	23.055
V2	13	54.279	39.514	83.786	18.302	36.358	24.223	16.977	27.909	44.111	8.469	76.004	23.522	40.781	27.360	31.550	5.472
V3	13	49.399	29.732	53.047	33.397	36.208	11.752	39.615	25.745	42.362	24.088	32.151	19.006	46.731	22.373	40.210	23.129
V4	13	69.387	23.252	87.452	14.219	38.702	21.886	10.541	14.525	61.058	8.635	77.680	21.741	35.313	24.303	51.142	7.464
V5	13	22.109	20.527	18.389	15.645	56.983	34.495	57.242	21.880	69.994	21.459	33.359	34.692	57.452	34.022	43.173	24.122
V6	13	25.841	18.664	54.826	23.931	51.689	34.260	46.713	24.045	55.559	24.579	30.060	29.996	60.944	27.515	48.155	25.381
V7	13	26.496	22.787	38.660	29.662	49.621	34.830	27.566	23.879	29.718	23.883	28.155	26.395	45.721	35.327	23.576	25.401
V8	13	16.767	20.335	48.960	28.725	48.684	33.197	41.941	24.664	46.388	16.734	31.767	30.505	50.697	34.175	43.251	30.565
V9	13	21.466	18.025	55.343	37.520	58.444	36.629	57.091	32.841	61.593	31.740	32.404	34.660	63.323	35.252	55.763	31.760
Totaal		35.718		51.166		47.543		36.099		37.937		40.519		50.998		31.743	

Tabel 11

De gemiddelde fysiologische reacties per fysiologische maat van de verschillende robotskarakters en mensen die zijn gebruikt voor de gepaarde metingen t-testen.

Fysiologische maat	Karakter	n	Robots		Mensen	
			M	SD	M	SD
GSR	WE-4RII	13	-1.521	3.710	0.748	2.073
	Flobi	11	0.229	2.567	0.339	0.996
	Diegosan	12	1.589	4.178	0.5879	2.080
	iCat	12	2.435	4.0102	0.861	2.123
	Saya	6	-0.8556	3.503	0.253	0.801
	FACE	12	1.816	4.720	0.861	2.123
	Sophia	13	2.327	8.623	0.748	2.073
	Einstein	7	-2.966	5.603	0.083	0.841
EMG Zygomaticus	WE-4RII	13	0.0252	0.150	0.064	0.316
	Flobi	11	-0.067	0.187	0.125	0.297
	Diegosan	12	-0.122	0.269	0.0623	0.330
	iCat	12	0.0086	0.092	-0.013	0.154
	Saya	6	0.194	0.315	0.034	0.098
	FACE	12	0.175	0.468	-0.013	0.154
	Sophia	13	0.110	0.547	0.064	0.316
	Einstein	7	-0.212	0.318	0.022	0.058
EMG Corrugator	WE-4RII	13	0.049	0.226	0.062	0.199
	Flobi	11	-0.101	0.220	0.060	0.216
	Diegosan	12	0.075	0.432	0.034	0.182
	iCat	12	0.181	0.639	0.027	0.162
	Saya	6	0.328	1.707	-0.043	0.112
	FACE	12	0.110	0.374	0.027	0.162
	Sophia	13	-0.152	0.433	0.062	0.199
	Einstein	7	0.114	0.362	-0.033	0.107
Hartslag	WE-4RII	13	0.068	0.527	0.066	0.231
	Flobi	11	-0.130	0.691	0.064	0.249
	Diegosan	12	0.0178	0.351	0.052	0.236
	iCat	12	0.022	0.745	0.073	0.240
	Saya	6	0.214	0.807	0.057	0.296
	FACE	12	-0.037	0.505	0.073	0.240
	Sophia	13	-0.067	0.726	0.066	0.231
	Einstein	7	-0.624	1.128	0.100	0.244

Tabel 12

Resultaten gepaarde metingen t-test waarbij per fysiologische maat iedere robotkarakter wordt vergeleken met de mensen.

Fysiologische maat	Stimulus	t- waarde	df	p-waarde
GSR	WE-4RII	-1.488	12	0.162
	Flobi	-0.147	10	0.886
	Diegosan	1.413	11	0.185
	iCat	1.670	11	0.123
	Saya	0.637	12	0.536
	FACE	0.688	11	0.506
	Sophia	-0.676	5	0.529
EMG Zygomaticus	Einstein	-1.422	6	0.205
	WE-4RII	-0.444	12	0.665
	Flobi	-2.037	10	0.069
	Diegosan	-1.402	11	0.189
	iCat	0.085	11	0.934
	Saya	0.264	12	0.796
	FACE	1.263	11	0.233
EMG Corrugator	Sophia	1.790	5	0.133
	Einstein	-1.569	6	0.168
	WE-4RII	-0.041	12	0.968
	Flobi	-1.567	10	0.148
	Diegosan	0.387	11	0.706
	iCat	0.916	11	0.379
	Saya	-1.751	12	0.105
Hartslag	FACE	0.831	11	0.424
	Sophia	0.548	5	0.607
	Einstein	0.903	6	0.401
	WE-4RII	0.013	12	0.990
	Flobi	-0.899	10	0.390
	Diegosan	-0.272	11	0.791
	iCat	-0.233	11	0.820
	Saya	-0.582	12	0.571
	FACE	-0.730	11	0.481
	Sophia	0.596	5	0.577
	Einstein	-1.548	6	0.173