

## Real feelings for virtual humans

Sebas Nouwen

“*There are millions of species on our planet, and we hope to add one more*”. Dit zijn de woorden van Pranav Mistry, hoofd van Samsung Technology and Advanced Research (STAR) Labs (BBC, 2020). Het is niet niks om een technologische uitvinding op gelijke voet te plaatsen met een nieuwe levensvorm. Desondanks werden deze buitengewone claims gedaan in de dagen voorafgaand aan de onthulling van NEON tijdens de Consumer Electronics Show (CES) op 7 januari 2020. Hoewel de details voor de aankondiging nog vaag waren, werd NEON op de website omschreven als “*Our first artificial human is here. NEON is a computationally created virtual being that looks and behaves like a real human, with the ability to show emotions and intelligence.*” (NEON, 2020). Niet alleen werd NEON gehypt als een revolutionaire sprong in realisme en capaciteiten, maar er werd ook anders naar ‘virtual humans’ (technologie die mensen imiteert) gekeken dan we gewend zijn. “*Not an AI assistant. Not an interface to the Internet. Not a music player. Simply, a friend.*” Simpelweg een vriend.

Met zulke hoge verwachtingen is het niet vreemd dat STAR Labs tijdens de officiële onthulling de hype niet waar kon maken (BBC, 2020). Het illustreert echter wel twee interessante ontwikkelingen van dit moment. Ten eerste zijn de verwachtingen van technologie in 2020 zo hoog, dat een computer als vriend niet bij voorbaat afgeschreven wordt. Deze revolutionaire sprong is blijkbaar realistisch genoeg om waar te kunnen zijn. Anders had NEON nooit zoveel hype kunnen creëren bij zoveel tech journalisten. Daarnaast valt op dat NEON niet gepositioneerd wordt als een *virtual assistant* zoals we die nu kennen. Het is geen Siri van Apple, Google assistent of Billie van Bol.com. Al deze ‘virtual humans’ worden heel functioneel ingezet; bijvoorbeeld om een bestelling te doen, de lampen te dimmen of het weerbericht op te vragen. Neon is wat dat betreft substantieel anders. Ondanks

dat Siri best een gevoel voor humor heeft (vraag haar maar eens om een grap te vertellen), is het idee van een ‘echte’ relatie met een virtueel persoon erg vreemd. De positionering van NEON is een duidelijk signaal dat onze relatie met *virtual humans* op het punt staat om ingrijpend te veranderen. Het zal niet lang meer duren voordat we digitale ‘vrienden’ krijgen. Mocht de voorgaande zin als science fiction klinken, dan is dat waarschijnlijk een goede illustratie van de relevantie van dit onderzoek.

We staan aan de vooravond van een revolutie die voor de meeste mensen als een grote verrassing zal komen. Films als *Her* (2013) en *Blade Runner 2049* (2018) geven misschien wel een realistischer beeld van de toekomst dan de gemiddelde kijker in 2020 denkt. In beide films is de technologie zo ver ontwikkeld dat relaties met kunstmatige intelligentie, in de vorm van virtuele mensen, een nu nog ongekende diepgang heeft. In *Her* gaat het om een stem zonder lichaam, een soort Siri, terwijl het virtuele menselijke karakter in *Blade Runner* wel een fysieke (zij het holografische) vorm heeft. Er is anno 2020 nog zo weinig bekend over ‘virtuele’ vriendschappen en relaties, terwijl de impact ontzettend groot kan zijn. Zowel op de individuele gebruiker als de maatschappij als geheel.

## Sociale virtual humans

Technologie heeft in het verleden spierkracht en cognitieve capaciteit kunnen vervangen, maar het is niet ondenkbaar dat binnenkort een aantal sociale ‘taken’ vervangen kunnen worden. De term sociale technologie refereert over het algemeen naar social media of communicatie technologie, oftewel technologie die sociaal contact tussen mensen faciliteert. In de toekomst zullen we het steeds vaker hebben over technologie die zelf sociale interactie heeft met mensen; een heel andere vorm van sociale technologie. Virtual humans worden al gebruikt voor het coachen van interculturele communicatie (Lane et al., 2008), onderhandelingsvaardigheden (Core et al., 2006) of als lifestyle coaches (Kulyk, Op den

Akker, Klaassen, & Van Gemert-Pijnen, 2014). Bill Gates noemde het “the most exciting thing going on” tijdens de Recode tech conference (Clark, 2016). "It's the big dream that anybody who's ever been in computer science has been thinking about." Grappig, want Microsoft's chatbot Tay deed het niet zo goed. Tay was een chatbot, ontworpen om menselijke interactie te leren door communicatie met echte mensen op Twitter. Binnen no-time sloeg het systeem op hol met afschuwelijk racistische posts. Als reactie op Tay's falen schreef Peter Lee, vice president van Microsoft Research, “In that sense, the challenges are just as much social as they are technical” (Lee, 2016). Tay werd na 16 uur offline gehaald door Microsoft. Er zitten veel haken en ogen aan het creëren van een goede virtual human en daarbij is kennis van mensen net zo belangrijk als kennis van technologie.

Uiteraard zijn er ook meer succesvolle projecten en is er veel concurrentie om de beste digitale assistent af te leveren. Google heeft tijdens haar I/O event in 2018 een indrukwekkende demo gegeven van Google assistant, die complexe natuurlijke gesprekken kon voeren, zonder dat gebruikers door hadden dat ze met een machine aan het praten waren. Het is een van de best werkende voorbeelden, maar heeft niet echt een identiteit. In tegenstelling tot Siri, de concurrerende kunstmatige intelligentie van Apple, die duidelijk een gevoel voor humor en een ‘mind of her own’ heeft. Overigens is humor ook een slimme strategie om te verhullen dat de robot niet altijd weet wat de juiste reactie op de gebruiker is.

Het afbakenen van mogelijkheden is een optie om het system bruikbaar te maken. De University of California gebruikt virtual humans in een Intelligent Tutoring System (ITS) (Lane et al., 2008). Een serious game, waarbij de gebruiker het vertrouwen van een virtual human, met een andere culturele achtergrond, moet winnen om een overeenkomst te sluiten. Het sluiten van die overeenkomst is mogelijk door de juiste keuzes te maken in de communicatie met deze virtual humans. Het meerkeuze aspect zorgt voor een eindige hoeveelheid opties en dus een vooraf realiseerbare set van reacties voor de virtual humans.

Tegelijkertijd is het ontbreken van interactie door middel van natuurlijke taal een belangrijke beperking als het gaat om de ecologische validiteit van de te trainen vaardigheden. Het bleek dat gebruikers vooral vooruitgingen in het herkennen van fases van onderhandelen.

Nederlanders zijn bijvoorbeeld meer to the point, terwijl het in andere culturen belangrijk is om langer bezig te zijn met het opbouwen van een relatie. Door de context smal te houden en de interactie beperkt, zal dit natuurlijk nooit uitmonden in een algemene intelligentie waar je voor ‘alles’ terecht kan.

Een belangrijk punt is dat een virtual human kan aansluiten op de context van de gebruiker. Cougar's virtual human agent is gemaakt om via een camera mensen te herkennen, over het weer te praten en kalenderinformatie te gebruiken in gesprekken (Connectome, z.d.). Omdat het systeem lijkt te reageren op omgevingsfactoren en unieke informatie, maakt dit het systeem al veel realistischer en persoonlijker. Soul Machines maakt eveneens virtual humans met hun Soul Engine (Soul machines, z.d.). Ze laten zich hierbij inspireren door onderzoek van neurowetenschappers. De virtual humans van Soul Machines, de naam zegt het al, worden gemaakt met een focus op menselijkheid en niet alleen maar functionaliteit. Omdat deze virtual humans zo menselijk zijn, worden ze ingezet in situaties met veel menselijk contact, zoals customer servicepersoneel, doktersassistent of als docent. In een van hun voorbeelden geven ze aan dat wanneer je een nieuwe Mercedes koopt, je een eigen persoonlijke vertegenwoordiger krijgt, die je wegwijs maakt met je nieuwe auto. Soul Machines geeft aan dat hun virtual humans uniek zijn door een specifieke persoonlijkheid, humor en natuurlijke gespreksvaardigheden, het tonen van en reageren op emoties en interactie afgestemd op de eindgebruiker.

Ook binnen de geestelijke gezondheidszorg wordt er onderzoek gedaan naar de bijdrage die virtual humans kunnen leveren. Een interessant voorbeeld is Vincent, een virtual human in de vorm van een chatbot (Lee et al., 2019). In het onderzoek werden twee condities

vergeleken. De meer voor de hand liggende waarin Vincent fungeerde als alternatief voor een zorgprofessional die de gebruiker helpt en daarnaast een conditie waarin de gebruiker juist degene is die zorg verleent aan Vincent. Het bleek dat gebruikers die compassie toonden voor Vincent uiteindelijk ook meer compassie voor zichzelf voelden. Uit dit onderzoek blijkt in ieder geval dat een sterkere band met een virtual human een positief effect op de gebruiker kan hebben.

## TamaCoachee

Vanuit de gedachte dat een sterkere band met een virtual human invloed heeft op diens effectiviteit is het TamaCoachee onderzoek gestart. Binnen de Fontys hogeschool wordt sinds september 2019 onderzoek gedaan naar een virtuele coachee, waarmee studenten tijdens hun studie hun coachvaardigheden kunnen oefenen. Daarbij is de relatie tussen coach en coachee in het bijzonder een punt van aandacht. Onder de naam TamaCoachee, een combinatie van Tamagotchi (het virtuele huisdier uit de jaren '90) en coachee, wordt gewerkt aan een tool die gebruikt kan worden in een onderwijscontext. Studenten kunnen op hun smartphone een gesprek voeren met een digitaal virtueel persoon, vergelijkbaar met een gesprek via Skype of FaceTime. Waar de Tamagotchi een virtueel diertje was dat gevoed moest worden om in leven te blijven, is de TamaCoachee een cliënt die jouw coachvaardigheden nodig heeft om geholpen te worden. De gebruiker heeft dus goede coach- en gespreksvaardigheden nodig om met de TamaCoachee om te gaan. Door een analyse van deze gespreksdata worden studenten individueel of in een klassikale setting voorzien van feedback en inzicht in hun coachingsstijl. Ook bestaat de mogelijkheid om met een docent in gesprek te gaan om gemaakte keuzes en de gevolgen daarvan te bespreken. Doordat de TamaCoachee als virtual human feedback geeft op datgene dat mensen zo uniek maakt,

complexe sociale interactie, is het een ideale context om de grenzen van kunstmatige intelligentie en verbintenis met virtual humans te onderzoeken.

## Learning Society

In het kader van TEC for Society is er binnen het thema Learning Society een eerste onderzoek gedaan naar de TamaCoachee. Reacties van de eindgebruikers, studenten psychologie, zijn onderzocht door het geven van een demonstratie.

Er is een 'visie' over de TamaCoachee als een instrument dat gebruikt kan worden om coachvaardigheden te trainen. Op dit moment is deze visie op de TamaCoachee nog geen realiteit, maar de ontwikkeling gaat snel. Om rekening te houden met de eindgebruikers en onverwachte effecten van de TamaCoachee is er een demonstratie prototype ontwikkeld. De demonstratie is zo opgezet dat deze een realistisch beeld geeft van het gebruik van de TamaCoachee.

In een aantal focus groepen is deze TamaCoachee demo besproken met de toekomstige gebruikers. In dit vooronderzoek kwamen een aantal opvallende resultaten aan het licht, welke zijn ingedeeld in 5 categorieën: relevantie, verbondenheid, ontwerp, effectiviteit en neveneffecten. De eerste categorie illustreert de relevantie van het TamaCoachee project en het nut om dit onderzoek uit te voeren. De andere vier categorieën sluiten aan bij de richting die gegeven wordt aan de onderzoeksvragen van vervolgonderzoek.

### Relevantie

1. Studenten geven aan dat hun eerste ervaring in de rol van coach tijdens de opleiding Toegepaste Psychologie erg spannend was. Er is een grote angst om "iemand's leven te verpesten", wanneer er geoefend wordt op een echt persoon.

2. Sommige studenten zien vooral het nut van de TamaCoachee bij het trainen van heel specifieke (persoonlijke) leerdoelen. Het stellen van individuele leerdoelen waaraan gewerkt kan worden lijkt daarom een belangrijke ontwerpeis.
3. De TamaCoachee wordt door veel studenten gezien als een *objectievere* manier van beoordelen. Met name omdat de casus voor iedere student gelijk is, terwijl er met de echte coachees grote onderlinge verschillen zijn.

#### Verbondenheid

4. Sommige studenten geven aan dat het oefenen op een virtueel persoon minder persoonlijk en minder prettig lijkt, omdat deze ervaring “niet echt” is. Het gaat dan niet om missende kenmerken, maar puur om ‘echtheid’ of ‘menselijkheid’.

#### Ontwerp

5. In elke groep waren wel één of twee deelnemers die niet doorhadden dat het ging om een Wizard of Oz opstelling. Ze dachten dat de demonstratie echt was en waren erg onder de indruk van het resultaat. Er is niet aangekondigd dat de demonstratie ‘nep’ was.
6. Sommige studenten missen de non-verbale communicatie en lichaamshouding in de demo. De demo van de TamaCoachee bestond uit een hoofd en schouders, terwijl de rest van het lichaam niet zichtbaar was.
7. Geen fysieke nabijheid zorgt ervoor dat bepaalde oefeningen onmogelijk gemaakt worden. Voorbeelden waren bijvoorbeeld oefeningen waarbij coach en coachee anders gingen zitten ten opzichte van elkaar in een fysieke locatie of het gebruik van rekwisieten zoals kaarten.

## Effectiviteit

8. Studenten willen een veilige plek om te oefenen, maar erkennen ook dat het gebrek aan (permanente) gevolgen hun motivatie kan verlagen.
  - a. Het overnieuw mogen doen zorgt ervoor dat er geen consequenties aan acties verbonden worden en de opbouw van een relatie met de TamaCoachee sterk verzwakt wordt.
  - b. Een stuk gesprek overnieuw doen kan bijdragen aan het leren als studenten in dezelfde situatie een andere tactiek kunnen proberen.
9. De meeste studenten zien het niet zitten om de TamaCoachee in te zetten als een middel om summatief te toetsen. Genoemde bezwaren zijn dat het niet goed werkt in vergelijking met een docent als beoordelaar en dat het anders (onecht) is ten opzichte van een coachgesprek met een persoon (en dat er dus iets anders getoetst wordt). Er wordt dus getwijfeld aan de validiteit, terwijl de objectiviteit (zie punt 3) juist positief benoemd wordt.

## Neveneffecten

10. Het delen van data is een punt van discussie. Sommige studenten willen graag hun data delen met de docent of zelfs de hele klas, omdat juist hier feedback opgehaald kan worden. Anderen vinden het te 'persoonlijk' om deze data te delen.

## Vervolgonderzoek

Om de TamaCoachee daadwerkelijk te kunnen ontwikkelen, op een manier dat deze zowel effectief is als ons iets leert over deze nieuwe mens-machine interactie, wordt een vervolgonderzoek opgezet. Het onderzoek zal zich richten op de onderstaande vier thema's.



## 1. Verbondenheid

- In welke mate voelen mensen zich verbonden met virtual humans?
- Welke factoren zijn van invloed op de mate van verbondenheid met virtual humans?
- Is de mate van verbondenheid veranderd in de afgelopen 55 jaar?

## 2. Ontwerp

- Welke uiterlijke kenmerken beïnvloeden de verbondenheid met virtual humans?
- Welke persoonlijkheidsfactoren beïnvloeden de verbondenheid met virtual humans?
- Welke vormen van uniciteit beïnvloeden de verbondenheid met virtual humans?

## 3. Effectiviteit

- Leidt een hogere mate van verbondenheid (1) tot effectievere training van coaches (in opleiding)?

## 4. Neveneffecten

- Wat zijn de neveneffecten van een sterke verbondenheid met een virtual human?

De ontwikkelingen gaan snel en de verwachting is dat een systeem zoals de TamaCoachee in de komende jaren werkelijkheid zal worden, maar op dit moment is het onmogelijk om onderzoek te doen naar iets dat nog niet bestaat. Tegelijkertijd is het onverstandig om te wachten tot de technologie beschikbaar is, voordat we gaan nadenken wat het effect is op het individu en de maatschappij. Door gebruik te maken van huidige beschikbare technologie in combinatie met onderzoeksmethoden zoals de ‘Wizard of Oz’ techniek, is het mogelijk om verder vooruit te kijken. We kunnen effecten onderzoeken van technologie die pas over een paar jaar beschikbaar zal worden. Als het gaat om onderzoek naar virtual humans, dan ligt de focus vaak op technische aspecten om de virtual humans zo realistisch of natuurlijk mogelijk te maken of op het rendement van virtual humans. Hoeveel geld kan het besparen? Wat zijn

de leeropbrengsten? Een kant die vaak vergeten wordt, is dat mensen een band opbouwen met andere mensen. Het is echter heel menselijk om op een vergelijkbare manier te kijken naar dieren of zelfs objecten (Reeves & Nass, 1996). Het zit in ons om een relatie met een virtueel mens op een vergelijkbare manier aan te gaan als met echte mensen. De door STAR beloofde virtuele vrienden bestaan nu misschien nog niet, maar de komende technologische revolutie zal nieuwe mogelijkheden ontsluiten waarvan de impact nu nog onbekend is.

*Gezien de enorme positieve en negatieve impact van Facebook (2004) of de iPhone (2007) was het verstandig geweest om 10 à 15 jaar geleden vooruit te kijken. Misschien hadden we een aantal maatschappelijke problemen daarmee kunnen voorkomen. Als we wachten tot realistische virtuele mensen een realiteit zijn, dan is het te laat om deze ontwikkeling te sturen. Het is niet vreemd dat een vriendschap met een virtueel persoon een grote individuele impact op iemand leven kan hebben. Maar zodra virtuele vrienden op elke goedkope smartphone hun weg vinden naar miljarden potentiële gebruikers, zal de maatschappelijk impact ook enorm zal zijn.*

Mensen hebben de neiging om te antropomorfyseren (Reeves & Nass, 1996).

Wanneer proefpersonen blootgesteld worden aan video's met geweld tegen mensen, robots en (andere) levenloze voorwerpen, blijkt uit de fMRI scan dat bij de video's met geweld tegen robots en mensen overeenkomsten te vertonen qua emotionele reacties (Pütten et al., 2013). Daarnaast was in de nasleep van een onderzoek van SELEMCA te zien dat echte en virtuele mensen na een periode van samenzijn weer van elkaar gescheiden werden (Burger, 2015). Het leidde tot verdriet en een gevoel van verlies. Wat nou als we 'te veel' waarde gaan hechten aan een relatie met virtual humans?

De kennis op zich is dus van grote maatschappelijke waarde, maar door gebruik te maken van Design Thinking zullen de onderzoeksresultaten leiden tot een nieuw prototype,

aan de hand waarvan weer nieuw onderzoek gedaan kan worden. Binnen de eerste twee jaar zal een prototype ontwikkeld worden dat niet alleen gebruikt kan worden in een experimentele context, maar ook in een ‘real life’ onderwijssetting. Daarmee levert het TamaCoachee project al in een vroeg stadium een bijdrage aan praktijkgericht onderwijs. Uiteraard zal dit prototype door middel van iteratief ontwerp daarna nog verder verbeterd worden. Tijdens de Coronacrisis is gebleken hoe lastig het is om een vaardigheidstraining (deels) online aan te kunnen bieden. In deze context kan de TamaCoachee van grote toegevoegde waarde zijn.

## Referenties

- BCC. (2020, 8 januari). *CES 2020: Neon's artificial humans 'don't live up to hype'*.  
Geraadpleegd op 8 januari 2020, van <https://www.bbc.com/news/technology-51028990>
- Burger, S. (Regisseur). (2015, 26 maart). *Ik ben Alice*. Nederland: VPRO.
- Clark, J. (2016, 23 juni). *Artificial intelligence has a 'sea of dudes' problem*. Geraadpleegd op 8 januari 2020, van <https://www.bloomberg.com/news/articles/2016-06-23/artificial-intelligence-has-a-sea-of-dudes-problem>
- Connectome: "Virtual Human Agent" is a next-generation interface, combined AI, AR, and Blockchain. (z.d.). Geraadpleegd op 8 januari 2020, van <https://connectome.to/>
- Core, M.G., Traum, D., Lane, H.C., Swartour, W., Marsella, S., Gratch, J., & Van Lent, M. (2006). Negotiation Skills through Practice and Reflection with Virtual Humans. *Simulation transaction of the society for modeling and simulation international*, 82(11), 685-701.
- Ireland, C. (2012). *Alan Turing at 100*. Retrieved from:  
<https://news.harvard.edu/gazette/story/2012/09/alan-turing-at-100/>
- Jonze, S. (Regisseur). (2013, 13 oktober). *Her*. Verenigde Staten.
- Kulyk, O. A., op den Akker, H. J. A., Klaassen, R., & van Gemert-Pijnen, J. E. W. C. (2014). Personalized Virtual Coaching for Lifestyle Support: Principles for Design and Evaluation. *International journal on advances in life sciences*, 6(3-4), 300-309.
- Lane, H., Hays, M., Core, M., Gomboc, D., Forbell, E., & Rosenberg, M. (2008). *Coaching Intercultural Communication in a Serious Game*. Geraadpleegd op 8 januari 2020, van [https://www.researchgate.net/publication/228907667\\_Coaching\\_Intercultural\\_Communication\\_in\\_a\\_Serious\\_Game](https://www.researchgate.net/publication/228907667_Coaching_Intercultural_Communication_in_a_Serious_Game)
- Lee, M., Ackermans, S. C. A., van As, N. L., Chang, H., Lucas, E. M., & IJsselsteijn, W. (2019). Caring for Vincent: a chatbot for self-compassion. In *CHI 2019 - Proceedings of*

*the 2019 CHI Conference on Human Factors in Computing Systems* [702]

ACM/IEEE. <https://doi.org/10.1145/3290605.3300932>

Lee, P. (2016, 25 maart). *Learning from Tay's introduction* [Official Microsoft blog].

Geraadpleegd op 8 januari 2020, van <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/>

NEON. (z.d.). Geraadpleegd op 8 januari 2020, van <https://www.neon.life>

Pandorabots (z.d.). Meet Kuki. Geraadpleegd op 8 januari 2020, van

<https://www.pandorabots.com/mitsuku/>

Pütten, A.M., Schulte, F.P., Eimler, S.C., Hoffmann, L., Sobieraj, S., Maderwald, S., Krämer,

N.C., & Brand, M. (2013). Neural correlates of empathy towards robots. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 215-216.

Reeves, B., & Nass, C. (2006). *The media equation: how people treat computers, television, and new media like real people and places*. Stanford, CA: CSLI Publications.

Rosenthal-von der Pütten, A. M., Krämer, N. C., Maderwald, S., Brand, M., & Grabenhorst, F.

(2019). Neural Mechanisms for Accepting and Rejecting Artificial Social Partners in the Uncanny Valley. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 39(33), 6555–6570. <https://doi.org/10.1523/JNEUROSCI.2956-18.2019>

Soul Machines: Changing the face of AI. (z.d.). Geraadpleegd op 8 januari 2020, van

<https://www.soulmachines.com>

Villeneuve, D. (Regisseur). (2017, 3 oktober). *Blade Runner 2049*. Verenigde Staten.