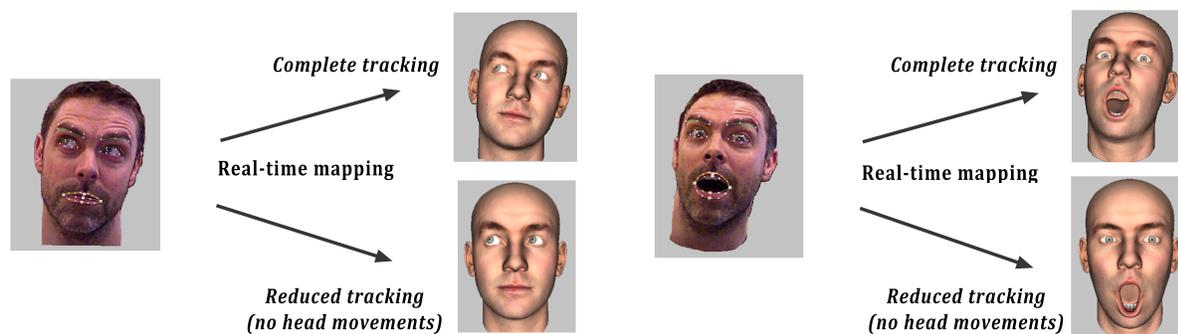


# Animation Realism Affects Perceived Character Appeal of a Self-Virtual Face

Elena Kokkinara\*      Rachel McDonnell †  
Graphics, Vision, and Visualisation Group  
Trinity College Dublin



**Figure 1:** Tracking data of facial movements is mapped onto the virtual face in real-time. We measure participants’ feelings of ownership and agency over the virtual face, as well as perceived appeal.

## Abstract

Appearance and animation realism of virtual characters in games, movies or other VR applications has been shown to affect audiences levels of acceptance and engagement with these characters. However, when a virtual character is representing us in VR setup, the level of engagement might also depend on the levels of perceived ownership and sense of control (agency) we feel towards this virtual character. In this study, we used advanced face-tracking technology in order to map real-time tracking data of participants’ head and eye movements, as well as facial expressions on virtual faces with different appearance realism characteristics (realistic or cartoon-like) and different levels of animation realism (complete or reduced facial movements). Our results suggest that virtual faces are perceived as more appealing when higher levels of animation realism are provided through real-time tracking. Moreover, high-levels of face-ownership and agency can be induced through synchronous mapping of the face tracking on the virtual face. In this study, we provide valuable insights for future games that use face tracking as an input.

**CR Categories:** I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Rendering;

**Keywords:** face ownership, face tracking, animation realism, character appeal, engagement

\*e-mail:ekokkina@scss.tcd.ie

†email:ramcdonn@scss.tcd.ie

## 1 Introduction

Recent advances in facial tracking technologies have allowed us to create realistic animations of virtual faces that would function even in real-time. A number of systems have been developed for gaming, movies and VR platforms, mainly aiming to track actors’ expressions and use them for off-line editing. However, these new technologies could be now used for designing new types of video-games, where real-time face-tracking can be used as an input for controlling virtual faces. As with body tracking, facial tracking as the main input for controlling self-avatars might effectively change the game experience, since participants might feel higher levels of immersion and engagement with their virtual representation. Hence, knowing the factors that contribute to engaging with self-virtual faces could be of importance.

People show various levels of acceptance towards virtual characters in movies, games and other VR applications, depending on how these characters look. Rendering realism has been shown to affect the audience reactions [Seyama and Nagayama 2007]. Highly realistic, human-like characters seem to be perceived as uncanny and produce negative reactions [Geller 2008; Levi 2004], hence designers are advised to create less realistic, cartoon-like humanoid characters. Moreover, animation realism of virtual characters’ faces has been considered highly important for conveying emotions and intent [Hyde et al. 2013].

However, to our knowledge, no perceptual experiments have assessed the way that participants engage with an animated virtual face that represents them in a virtual setup and what are the influencing factors, when they see a real-time mirrored representation of their facial expressions mapped on the virtual face. It is also possible that self-identification, ownership and agency feelings can be induced towards the virtual face, through a realistic mapping of the tracked facial movements of the participant on the virtual face [Slater et al. 2009].

We designed a study where one of two different appearance styles (realistic or cartoon) and two different levels of animation realism (complete or reduced) of self-virtual faces were presented to the participants. A computer screen-based VR scenario was designed to provoke head movements, speaking and facial expressions from

the participant. Eye and head movements and expressions were tracked in real-time and were mapped on the virtual face model, providing synchronous visual-proprioceptive correlations. In control conditions, the mapping of the head movements was removed, while only eye movements and facial expressions were correctly mapped, thus reducing the animation realism (Figure 1). One goal of this study was to examine whether ownership towards a mirrored virtual face can be induced, through real-time animation of eye and head movements and facial expressions. Moreover, we assumed that agency towards the facial movements will be induced due to synchronous animation of the real and virtual face. We hypothesised that unresponsive head movements would affect the levels of ownership and agency. Another goal of this study was to investigate whether the appearance style can affect the potential feelings of ownership and agency towards the virtual face that represents them in VR, as well as engagement with the self-avatar and the VR scenario.

We evaluated the VR experience through a specially designed questionnaire and through physiological reactions (skin conductance response) to an unexpected event in the virtual scenario.

## 2 Related Work

A theory called the ‘Uncanny Valley’ has been used to describe the fact that negative audience reactions occur when the movement of a virtual character does not match its realistic appearance [Mori 1970]. Hodgins et al. [2010] conducted perceptual experiments to determine how degradation of human motion affects the ability of a character to convincingly act out a scene. They found that removing facial animation changed the emotional content of the scene. McDonnell et al. [2012] examined this effect further using a range of rendering styles, and found that animation artifacts were more acceptable on cartoon than on realistic virtual humans.

Studies in immersive virtual reality have shown that it is possible to feel an illusory ownership over a virtual body, when the body is seen from a first person perspective and when participants receive synchronous tapping on the virtual body and their hidden real body [Slater et al. 2009]. Similarly, when participants see their collocated virtual body animating in synchrony with their tracked real body, they can feel a sense of ownership and control over their virtual representation, while asynchronous visuo-motor correlations can break the illusion of ownership [Kokkinara and Slater 2014]. Although morphological similarity towards the seen body has been suggested to affect the illusion of body ownership, it seems that if other multisensory correlations are present appearance does not affect the illusion [Maselli and Slater 2013].

In the case of face-identification/ownership, it is known that visual recognition of memorised visual features contributes to self-face recognition [Brady et al. 2005; Brady et al. 2004]. However, multisensory input from touch, similar to the rubber hand illusion [Botvinick et al. 1998], has also been suggested to contribute to self-recognition, even in the absence of morphological similarity with the seen face (e.g stimulation on another person’s face) [Tsakiris 2008].

Here, we consider the possibility to perceive ownership and control over a mirrored virtual face that appears with synchronous animated expressions to those of the tracked real face, while we are also examining the effect of the appearance realism.

## 3 Methods

We recruited 60 male participants, aged 18-49 (mean  $\pm$  SD = 25.95  $\pm$  6.6). We chose to recruit only male participants in order

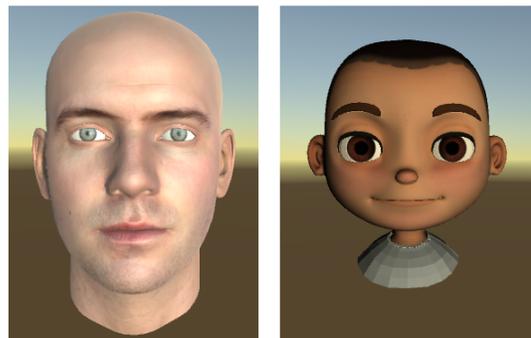
to match the gender of the available face 3D models that would represent them in the virtual scenario, as previous research has actually shown that female motion applied to a male character can appear ambiguous [Zibrek et al. 2013]. Participants were students from different disciplinary backgrounds as well as employed individuals from various fields, recruited mainly via university mailing list and advertisement in the university campus. They were naïve to the purpose of the experiment, and were given book vouchers as a reward for participation.

The experiment had a 2x2 factorial design, with factors character appearance (realistic, cartoon) (Figure 2) and animation realism (complete, reduced face animation) (Figure 1). It was a between-groups design and participants were equally split in four groups (n=15) and experienced only one of the four conditions.

### 3.1 Equipment

The virtual scene was presented on a 24” monitor, adjusted for the height of the participant. An ASUS Xtion PRO LIVE<sup>1</sup> camera, with infrared sensors, adaptive depth detection technology and color image sensing, was adjusted on top of the screen, for motion-sensing of the participant’s face. Real-time motion capture was performed using Faceshift studio<sup>2</sup>. The virtual scenario was realised using the Unity3D game engine<sup>3</sup>. The real-time motion data of the facial movement was streamed from Faceshift to Unity in order to animate the virtual self-avatar face. We used two of the available models (a realistic and a cartoon virtual face) provided with Faceshift (Figure 2). A wireless keyboard was used by the experimenter in order to trigger the AI voice.

Skin conductance signals were recorded at a sampling rate of 100 Hz, using Plux’s portable biosignal acquisition device, Bitalino<sup>4</sup> while the recording and storage of the data was handled directly in Unity. All statistical analysis was carried out using Matlab<sup>5</sup> and R [Team 2000].



**Figure 2:** Self-avatar faces. Left: Realistic face. Right: Cartoon face.

### 3.2 Procedure

Participants read the instructions of the experiment and they signed a consent form. After they filled a form with demographics information, they were seated in front of the computer screen and were asked to train Faceshift, in order to create a personalised profile of

<sup>1</sup><http://bit.ly/1OzfxRb>

<sup>2</sup><http://www.faceshift.com/>

<sup>3</sup><https://unity3d.com/>

<sup>4</sup><http://www.bitalino.com/>

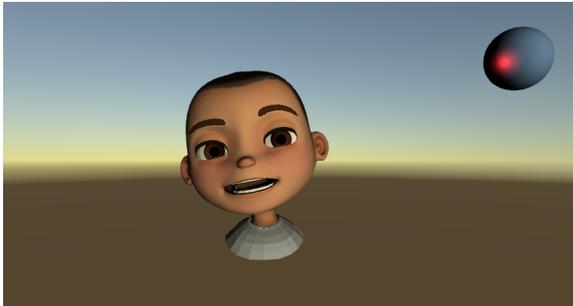
<sup>5</sup><http://mathworks.com/>

their facial expressions. Faceshift would later use this profile in order to map the head, eyes and facial expressions on the virtual face.

After training the system, the experimenter attached the sensors for recording skin conductance and verbally repeated the instructions of the experiment that the participant had previously read.

When the virtual world was first presented on the computer screen, we hid the virtual face and we instructed participants to relax for one minute in order to record a baseline of their skin conductance levels during a relaxation period. After this period the virtual face appeared on the middle of the screen in front of the participant, with face tracking enabled.

An abstract-looking AI agent in the shape of a sphere, speaking in a robot voice, was leading a conversation with the participant (Figure 3, Supplementary video). The questions of the AI agent were triggered by a trained experimenter in the same order for all participants, while some of the responses of the agent were decided by the experimenter on the spot, in order to match participants responds, using the Wizard-of-Oz method [Dahlbäck et al. 1993]. The questions were designed in a way that would provoke head movements, speaking and facial expression from the participant. Examples are: asking about their favourite food, favourite film, or they were even asked to dance (using head movements only) and sing to the rhythm of some famous pop songs. We also asked them to repeat some tongue-twisters, which provoked laughing or smiling. Finally, participants were asked to follow a sphere moving around the screen with their head and eye gaze. This provoked some additional head movements (see Supplementary video). During the reduced animation realism conditions, eye and facial expression tracking was enabled, but there was no head tracking. Participants were asked to focus on their virtual face at all times, avoiding focusing their gaze on the AI robot or outside the screen.



**Figure 3:** Virtual scenario. The virtual face appeared in the middle of the screen and followed participant’s facial expressions, head and eye movements. The AI agent (blue sphere) appeared on the top right corner. A red light was flickering on the blue sphere, in order to simulate speaking.

### 3.3 Response variables

After the last task that the agent assigned to the participant (following the virtual sphere moving on the screen with their head gaze), a simulation of rain was triggered to fall above the virtual face. We expected this to be an arousing event causing skin conductance response. Based on previous studies, we also expected these responses to correlate with the levels of ownership towards the virtual face, as if their own face would get wet from the rain [Armell and Ramachandran 2003; Petkova and Ehrsson 2008; Slater et al. 2010]. We calculated the maximum amplitude of skin conductance levels during the last 10s of the relaxation period and 10s after the

rain started falling.

A 9-item questionnaire, partially based on standardized questionnaires, was used to assess the levels of perceived appeal, ownership and control over the virtual face, as well as the levels of engagement with the AI agent during the experience and the experience of the rain at the end of the experiment (Table 1). Questions were rated on a Likert scale from 1 (totally disagree) to 7 (totally agree).

**Table 1:** The post-experience questionnaire. All questions were rated on a Likert scale from 1 to 7 (Unless stated differently: 1-totally disagree, 7-totally agree).

Question	Statement
Ownership	“During the experiment, I felt that the virtual face was my own face.”
MyMovements	“During the experiment, I felt that the movements of the virtual face were my movements.”
Agency	“During the experiment, I felt that the movements of the virtual face were caused by my movements.”
OtherFace	“During the experiment I felt that the virtual face belonged to someone else.”
RainThreat	“I felt that my own face would get wet when it started raining.”
TrueRain	“For a moment I thought it was raining in the experiment room.”
Appeal	“How appealing did you find this virtual face?”(High appeal rating means that the virtual face is one that you would like to watch more of, and you would be captivated by a movie with that face on a character as the lead actor. Use all cues: appearance, motion) (1-Not appealing at all, 7-Extremely appealing)”
EnjoyAI	“I enjoyed talking to AI (the talking sphere).” (1-Not at All, 7-Very much)
FutureAI	“I would like to talk with AI (the talking sphere) in the future.” (1-Not at All, 7-Very much)

## 4 Results

**Questionnaires:** Table 2 shows the medians and interquartile ranges (IQRs) of the questionnaire responses in the experimental period. Participants reported high levels of ownership towards the virtual face (Ownership), feeling of causing the animation movements (Agency), ownership of the movements (MyMovements) in all conditions and enjoyed interacting with the AI agent (EnjoyAI, FutureAI). Character appeal (Appeal) was high for the complete animation realism, but was lower for the reduced animation realism (see also Figure 4). The control question for ownership (OtherFace) was low in all cases, while participant did not report to feel threatened or perceived as real the simulated rain (RainThreat, TrueRain).

Since the questionnaire scores are ordinal, we used ordered logistic regression on the scores in order to formally test the above results for the 2x2 experimental design. The resulting main effects of the two factors were low for all questions, except for Appeal, where animation realism significantly affected the perceived character appeal [coefficient=-1.3132, SE=0.5633, Z=-2.331, p= 0.019, the interaction term was removed since it was not significant]. Specifically, appeal was rated significantly lower when animation realism was reduced (Figure 4).

**Physiological reactions:** Skin conductance data from two participants in the cartoon/reduced animation condition was removed

**Table 2: Medians and interquartile ranges (IQR) for each condition.**

Question	Complete Animation	
	Realistic Face	Cartoon Face
	Median(IQR) n=15	Median(IQR) n=15
Ownership	5(0)	5(1)
MyMovements	6(1)	6(0)
Agency	7(1)	7(0.5)
OtherFace	3(2)	2(1)
RainThreat	1(1)	1(1)
TrueRain	1(0.5)	1(1)
Appeal	5(0)	5(1)
EnjoyAI	5(1.5)	5(0.5)
FutureAI	5(2.5)	5(3)

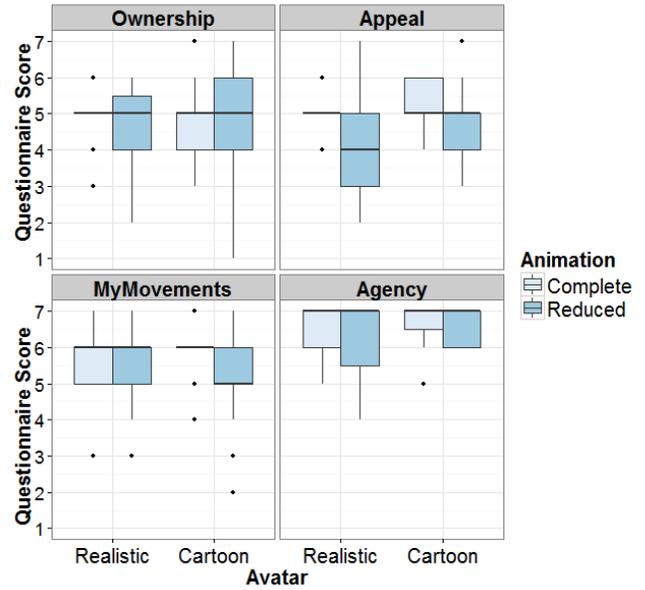
  

Question	Reduced Animation	
	Realistic Face	Cartoon Face
	Median(IQR) n=15	Median(IQR) n=15
Ownership	5(1.5)	5(2)
MyMovements	6(1)	5(1)
Agency	7(1.5)	7(1)
OtherFace	3(3)	2(2)
RainThreat	2(1.5)	1(1.5)
TrueRain	1(1)	1(0.5)
Appeal	4(2)	5(1)
EnjoyAI	5(1)	5(2)
FutureAI	4(2)	5(1.5)

due to a recording error. A small skin conductance response was detected for all cases directly after the simulated rain started falling above the virtual face (Figure 5, red line). In order to test the possible differences between the conditions, we used as a response variable the percentage of change between the maximum skin conductance amplitude in the final 10 seconds of the baseline (relaxation) and the maximum skin conductance amplitude in the first 10 seconds of the stimulation (falling rain) periods. No effects of the two factors were found. It is noticeable that the rain event was not particularly arousing and it was not perceived as threatening. This was also confirmed by the rates on the question RainThreat. Nevertheless, different types of events throughout the experimental phase seem to be particularly arousing, for example when the AI robot asks the participant to sing a song (Figure 5, blue line). However, we could not explain these reactions in relation to high feelings of ownership, but we interpret those as ordinary emotional reactions (e.g. participants were embarrassed to sing in front of the experimenter).

## 5 Discussion

In this paper, we investigated the effect of appearance realism and animation realism of self-virtual faces on face-ownership, agency and perceived character appeal. We conducted a perceptual experiment using a VR scenario where participants were represented by a virtual face that was replicating their facial movements and we used subjective and objective measures to evaluate participants' reactions. Our results suggest that high levels of face ownership and agency towards the virtual face can be induced through real-time mapping of the facial movements. Furthermore, participants perceived both appearance styles highly appealing when complete animation realism was provided, while the perceived appeal of



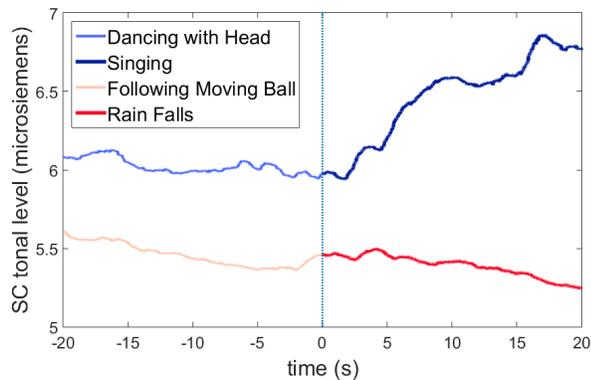
**Figure 4: Box plot for questions Ownership, Appeal, MyMovements and Agency.**

the characters was lower when the animation realism was reduced through unresponsive head movements.

We found no differences on the reported ownership levels between the two appearance styles. This is inline with previous studies that suggest that the level of realism of a virtual body does not affect the illusion of body-ownership when congruent visual-tactile or visual-motor cues are provided [Maselli and Slater 2013]. However, we did not control for physical similarity of participants real and virtual face (e.g. hair/eyes/skin colour, nose, hair). In a post-experimental interview, some participants reported that they did not look like the virtual face, which might have affected the perceived connection with the character. Future studies should further investigate the effect of face-features similarity between participants' real and virtual face.

Opposite to our expectations, ownership levels were not affected by the animation realism. In this case, it seems that a minimum level of visuomotor correlation, coming from synchronous mapping of the facial animation is enough for inducing the illusion. However, it is possible that if participants in the reduced animation realism condition had previously experienced complete (head) animations as a comparison baseline, the subjective reports on ownership might have been different. It is also possible that introducing different types of animation anomalies, such as turning off eye movements and/or facial expressions, might have produced a different result. Similarly, given previous literature, we believe that introducing noise, or a large delay in the tracking would have had an effect of the perceived ownership and agency [Kokkinara and Slater 2014]. Future work should meticulously explore the necessary animation parameters for the induction of face ownership illusion.

The fact that the sense of agency and ownership of the movements were not affected by the animation realism can be justified by the nature of the questions (Agency, MyMovement in Table 1). Participants were asked if they felt that the movements of the virtual face were their own movements or were caused by their movements, which were true for both animation realism conditions, i.e. they



**Figure 5:** Average skin conductance levels of all participants around two different event times. Blue: participants pass from a phase where they move their head in time to a song (light blue) to a phase where they are asked to sing a song (dark blue). Red: participants pass from a phase where they follow a moving ball on the screen using head gaze (light red) to a phase where simulated rain falls above the head of the virtual face (dark red). Time=0 signifies the starting point of each of the two events (singing and rain).

were causing both the complete or the reduced animations. An interesting alternative question would be whether they felt that the virtual face was effectively replicating their movements.

Contrary to the Uncanny Valley theory, the realistic face was rated equally appealing to the cartoon face in both animation realism conditions. Previous studies have shown that lower animation realism was more acceptable on cartoon than on realistic virtual humans [McDonnell et al. 2012]. However, here it is possible that the real-time tracking of the face and possibly the perceived connection (agency and ownership) with the virtual face have equilibrated the perceived appeal for the two appearances. Although the perceived agency and ownership are not affected by animation realism, character appeal seems to be reduced due to the reduced animation realism. This is an additional indication that efficient real-time mapping of the facial movements on the self-virtual face is used as a cue for perceived appeal of the face. Future studies could further assess this view, by comparing perceived appeal on pre-recorded animated faces to self-produced real-time ones.

From the exploratory questions regarding the experience of talking with the AI robot, it seems that appearance and animation realism of the self-virtual face did not affect the reported enjoyment. Furthermore, we found no apparent differences of the physiological reaction to the simulated rain event for the different conditions. We used physiological reaction in this study as an additional objective measurement of ownership. Although, the results are inline with the subjective reports on ownership (no changes across conditions), it seems that the chosen event was not perceived as sufficiently threatening to the self-virtual face, while the stimulus was not presented on an immersive virtual environment like in previous studies [Slater et al. 2010].

This study provides valuable insights regarding engagement with self-virtual faces that are animated through facial tracking. Observing believable representations of our own facial expressions on self-avatars can potentially have a big effect on engagement and enjoyment of new types of games that use face tracking as a main input. Game designers might need to focus on the level of control that the users feel over their virtual representation in order to ensure high levels of appeal.

## Acknowledgements

This research was funded by the Horizon 2020 EU POPULATE (644655) project, funded under the research and innovation programme.

## References

- ARMEL, K. C., AND RAMACHANDRAN, V. S. 2003. Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 1523, 1499–1506.
- BOTVINICK, M., COHEN, J., ET AL. 1998. Rubber hands’ feel’touch that eyes see. *Nature* 391, 6669, 756–756.
- BRADY, N., CAMPBELL, M., AND FLAHERTY, M. 2004. My left brain and me: a dissociation in the perception of self and others. *Neuropsychologia* 42, 9, 1156–1161.
- BRADY, N., CAMPBELL, M., AND FLAHERTY, M. 2005. Perceptual asymmetries are preserved in memory for highly familiar faces of self and friend. *Brain and cognition* 58, 3, 334–342.
- DAHLBÄCK, N., JÖNSSON, A., AND AHRENBORG, L. 1993. Wizard of oz studies: why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces*, ACM, 193–200.
- GELLER, T. 2008. Overcoming the uncanny valley. *IEEE computer graphics and applications*, 4, 11–17.
- HODGINS, J. K., JÖRG, S., O’SULLIVAN, C., PARK, S. I., AND MAHLER, M. 2010. The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception* 7, 4, 1–14.
- HYDE, J., CARTER, E. J., KIESLER, S., AND HODGINS, J. K. 2013. Perceptual effects of damped and exaggerated facial motion in animated characters. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, IEEE, 1–6.
- KOKKINARA, E., AND SLATER, M. 2014. Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception* 43, 1, 43–58.
- LEVI, S. 2004. Why tom hanks is less than human. while sensors can not capture how humans act, humans can give life to digital characters. *Newsweek* 650, 305–6.
- MASELLI, A., AND SLATER, M. 2013. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience* 7.
- MCDONNELL, R., BREIDT, M., AND BUELTHOFF, H. 2012. Render me real Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics* 31, 4, 91:1–91:11.
- MORI, M. 1970. The uncanny valley. *Energy* 7, 4, 33–35.
- PETKOVA, V. I., AND EHRSSON, H. H. 2008. If i were you: Perceptual illusion of body swapping. *PLoS ONE* 3, 12 (12), e3832.
- SEYAMA, J., AND 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.

- SLATER, M., PEREZ-MARCOS, D., EHRSSON, H. H., AND SANCHEZ-VIVES, M. V. 2009. Inducing illusory ownership of a virtual body. *Frontiers in neuroscience* 3, 2, 214.
- SLATER, M., SPANLANG, B., SANCHEZ-VIVES, M. V., BLANKE, O., ET AL. 2010. First person experience of body transfer in virtual reality. *PloS one* 5, 5, e10564.
- TEAM, R. C., 2000. R language definition.
- TSAKIRIS, M. 2008. Looking for myself: current multisensory input alters self-face recognition. *PloS one* 3, 12, e4040.
- ZIBREK, K., HOYET, L., RUHLAND, K., AND MCDONNELL, R. 2013. Evaluating the effect of emotion on gender recognition in virtual humans. In *Proceedings of the ACM Symposium on Applied Perception*, ACM, 45–49.