

Using Facial Animation to Increase the Enfacement Illusion and Avatar Self-Identification

Mar Gonzalez-Franco, *Member, IEEE*, Anthony Steed, *Member, IEEE*, Steve Hoogendyk, and Eyal Ofek, *Member, IEEE*

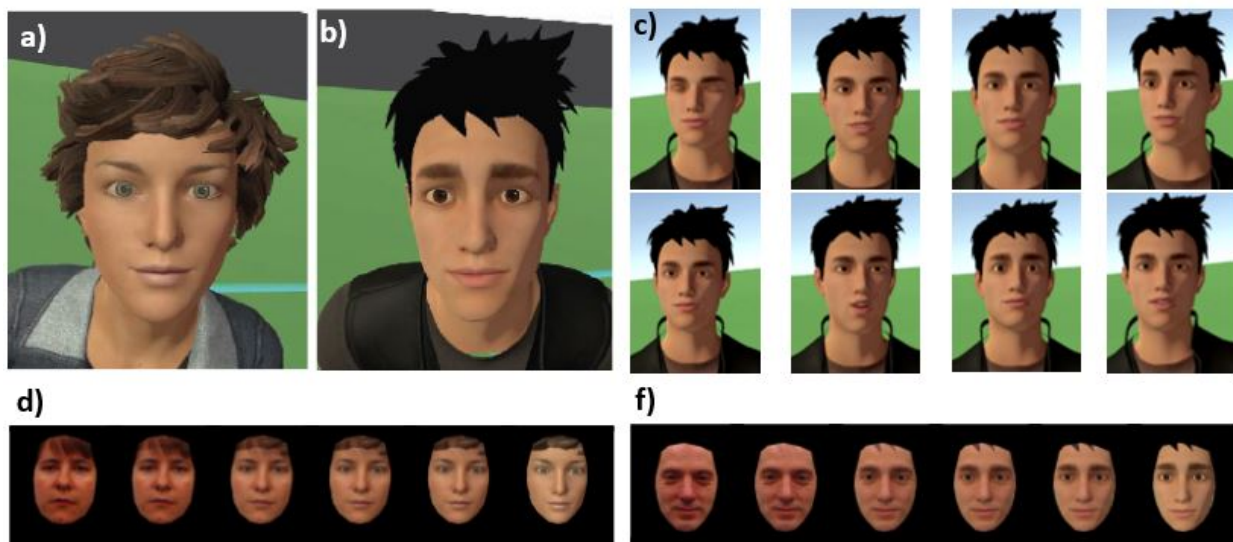


Fig. 1. A face morphing test (d,f) showed higher enfacement levels towards avatars (a,b) that had additional facial animations (c). Examples of the face of the male avatar showing frames mimicking speech and also idle animations such as eye blinks can be observed in (c).

Abstract—Through avatar embodiment in Virtual Reality (VR) we can achieve the illusion that an avatar is substituting our body: the avatar moves as we move and we see it from a first person perspective. However, self-identification, the process of identifying a representation as being oneself, poses new challenges because a key determinant is that we see and have agency in our own face. Providing control over the face is hard with current HMD technologies because face tracking is either cumbersome or error prone. However, limited animation is easily achieved based on speaking. We investigate the level of avatar enfacement, that is believing that a picture of a face is one's own face, with three levels of facial animation: (i) one in which the facial expressions of the avatars are static, (ii) one in which we implement lip-sync motion and (iii) one in which the avatar presents lip-sync plus additional facial animations, with blinks, designed by a professional animator. We measure self-identification using a face morphing tool that morphs from the face of the participant to the face of a gender matched avatar. We find that self-identification on avatars can be increased through pre-baked animations even when these are not photorealistic nor look like the participant.

Index Terms—self-avatars, virtual reality, embodiment, face animation, enfacement

1 INTRODUCTION

An increasing proportion of consumer immersive virtual reality (VR) systems support six degrees of tracking for the head and both hands. This has enabled development of a very broad range of novel content experiences in which users can move and interact with their bodies. To some extent users, can utilise motions and skills that they have learnt outside VR. However, the user of a head-mounted display (HMD) system cannot see their own body. Thus a key decision for the design of immersive experiences is how the user is represented to themselves, and once a representation has been constructed, whether and how users

identify with that representation.

The literature on impact of avatars is extensive. Very realistic avatars of users can be made (see Section 2.1), but these can elicit negative reactions from viewers because although they appear realistic they do not have all the behaviours or responses of the real body. This has become known as the uncanny valley, see [26, 27]. However, recent work has shown a very clear type of impact of an avatar that represents the self. With avatars of a wide varieties of appearances, users can feel embodied within that avatar to the extent that they identify with it and have some degree of ownership (e.g. [14, 17]). As we discuss further in Section 2.2, having a virtual body different from our own can change our interactions and biases (e.g. [24]). Self-animated avatars have been shown to improve communication in virtual environments [8].

While most related work in VR has focused on ownership of the body, our self-identity is very strongly associated with our face. While the self-avatar's face may not be usually visible to the user in an immersive experience, mirrors or mirrored surfaces will allow the users to inspect themselves [16]. Using mirrors, researchers have shown that Enfacement (see Section 2.3), and self-identity is a fluid concept [46].

• Mar Gonzalez-Franco, Steve Hoogendyk and Eyal Ofek are with Microsoft Research. Anthony Steed was with Microsoft Research and is now at University College London. Corresponding E-mail: margon@microsoft.com.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

In fact, look-a-like avatars, that are not perfect pictures, are also self-recognised at a neurological level differently than other avatars [12]. But given the state of current consumer equipment and content production pipelines, most consumer VR experiences do not support realistic self-avatars and do not support tracking of the face, though these technologies are under development, see Section 2.1. Thus we now ask whether users in consumer VR self-identify with an avatar having a face different than their own.

Most immersive consumer experiences use cartoon-style or more abstract representations. Indeed, such representations might be preferred by some designers, even if real-time photorealism were available. However, despite currently not having the systems to track the face, we have access to the user's voice and we can drive the facial animation from that.

A common way of inducing ownership illusions is through tactile induction. While you could do tactile induction on the face, we consider this as less appropriate given that the face is so sensitive. Tactile induction is also not currently deployable within a consumer context. We can however have users inhabit and explore use of an avatar face. Thus, in this paper we explore whether well-designed facial animations can affect self-identification. Our experience uses a measurement of self-identification similar to that of Tsakiris [46]. We find that merely the embodiment of an avatar in front of a mirror already increases the self-recognition of the avatar. Furthermore, users embodied in avatars that exhibit both lip-sync and face idling mechanisms (such as blinks) self-recognised even more as the avatar than those who only had lip-sync. This has important implications for production of VR experiences because it is very easy to implement and thus can be broadly used if an affinity with the character is desirable.

2 RELATED WORK

2.1 Self-Avatar Construction

The form of the self-avatar is a key set of decision points in the design of an immersive experience. The designer must choose whether or not to draw a full body avatar, and subsequently choose between a photorealistic avatar or a cartoony avatar. One common choice is to draw the self-avatar as a pair of hands or tools in the hands along with a simple head or representation of a HMD. This is faithful to the positional tracking typically available [38], but it is limited in expressiveness.

While photo-realistic avatars of the quality achievable in off-line rendering are coming (e.g. see the Digital Emily project [1]), the use of a photorealistic model might not be appropriate for the design of the environment. On one hand it might reduce the anonymity of the participants, and on the other, the characteristics of the game design might require a cartoony look.

In some situations faithful tracking of the face will not be appropriate or desired by users that are maybe performing other tasks simultaneously or want to mask their emotional state. The face is an important resource for non-verbal communication and thus it is very common to provide some animation of faces even if this cannot be tracked [31]. One common way of modelling faces is to implement facial bones or blend shape rigs, [21, 30], which puts the control of the face in a structure that is amenable to real-time animation and at the same time it reduces the number of parameters to be controlled. These facial rigs can be driven by cameras that track the full face (e.g. [40]) or cameras can be mounted on a HMD to track part of the face and infer the full face [20, 29, 43]. However, at the current time, consumer immersive systems do not provide face tracking and it is not clear if precise tracking will be acceptable or useful in all situations.

An alternative that is commonly used is to animate the face with pre-canned animations where those animation might be triggered by user interaction and voice. However, animating something as apparently simple as eye-animation and eye blinks needs significant attention to detail in order to appear naturalistic [39, 45]. For example, Trutoiu et al. investigate how smile and blinks are temporally correlated [44]. These results suggest that animation of the face requires careful design in order to appear plausible.

In this paper, we will be concerned with humanoid characters with a cartoony or stylised appearance. We use lip-sync and professionally-

authored idle animations to explore the importance of these factors on self-recognition and enfacement illusions on non look-a-like avatars.

2.2 Self-Avatar Impact

There has been a very wide variety of work on the impact of avatars. In a social situation, the appearance and behaviour of the avatar is extremely important. However, there is a complex interaction between the appearance and behaviour of the avatar. For example, in an early study in the area, Garau et al. [10] found that inferred eye-gaze animations had a positive on user response to a high quality avatar, but a negative impact on response on a more abstract avatar. This suggests that the uncanny valley effect [26, 27] is a complex interaction between appearance and behaviour. Users seem to be particularly sensitive to representations of the face [35]. In more complex simulations of character and personality of animated characters, realism seems to be a positive attribute [48].

As noted in the introduction, one clear impact of avatar representation is when the avatar represents the user themselves. The sense of *embodiment* in VR now attracts a lot of attention [17]. Building on the seminal work of Botvinick and Cohen [5] and Petkova and Ehrsson [32], who showed that synchronous visual and tactile stimulation could induce an illusion of being embodied within a physical prop resembling an arm or body, Slater et al. showed that a similar illusion could be achieved with a virtual arm [37]. Subsequent work showed that an embodiment illusion could be generated without tactile stimulation by having the immersive representation of the hand with tracked fingers [33] or of the whole arm when playing simple VR games [47]. Embodiment has since been shown to have significant potential impacts on the user such as changing biases [24], or changing distance judgements [41]. Work specifically on whether the self avatar has to be realistic has shown that the avatar can vary quite a lot in appearance and still induce an illusion of ownership [3, 22]. There does however appear to be an interaction between the gender of the user and the appearance of the gender of the self-avatar [34] that highlights the need to have gender-specific representations.

2.3 Self-Identification through the Face

The prior work on body ownership has mostly focused on movement of the hands and arms. However, one thread focuses on ownership of faces and the impact of this. Tsarkis's early work had the user in front of a laptop that showed a virtual avatar head [46]. Participants experienced tactile stimulation on the face as they saw the avatar face touched. This impacted the participants' ratings of self-identity in that they were biased towards rating a representation of their face that was morphed into the avatar as their own face, see Section 3.2 for more detail of the protocol. This illusion is known as the *enfacement illusion*. It has now been studied in quite some detail, exploring questions of individual sensitivity, sub-components of self-identity and relation to theories of self-perception [2, 36, 42].

Bringing this toward VR, Kokkinara and McDonnell investigated ownership of the face of a self-avatar that is animated with face tracking information [19]. They found that participants felt more ownership over faces with higher levels of animation realism. It has been shown also that smile animations can create strong responses on participants [28], even if they do not completely match the exact facial expression of the participants. This then leads to the question of whether an induction of an ownership over the face is possible without explicit tracking and without having a very detailed avatar of the user's own face.

We are particularly interested in the use case of consumer VR and the use of stylised avatars rather than photorealistic captures of the users. A key question is thus whether the enfacement illusion can work with a highly animated, but somewhat abstract avatar. The work of Fuentes et al. suggests that spatial knowledge of one's own face is quite poor [9]. Self-identification with another's body has also shown to alter self-other face distinction, as a result of the enfacement illusion [7]. This suggests that faces with quite different structure might support the illusion.

3 MATERIALS AND METHODS

3.1 Apparatus

Participants were equipped with an HTC Vive VR Headset. A Unity scene was constructed that used an inverse kinematics toolkit (Final IK) to transfer the tracking of the participant's motion to the avatar. Additionally, the lip-sync animations were posed in Maya, brought into Unity as .fbx files and then setup with a state machine to control blending between poses. The state machine used the loop back microphone tool provided by Oxygen also available on the Unity Asset Store. The scene was otherwise very simple with a mirror positioned so that the user could see their self-avatar (see accompanying video).

The idle facial system of the animated condition was comprised of a number of ambient animations intended to help make the avatar feel alive. A professional animator designed the facial system by authoring a number of looping animations on various parts of the face. For example, the nostrils scale slightly to give the impression of breathing, eyelids blink occasionally, eyes have small quick movement to simulate saccades, and lips purse at times to help simulate skin pliability. All of these various facial animations have different loop lengths to hide the repeating nature of the animations. The animations were then brought into Unity and set up on an animation layer above a lip-sync layer. This allowed both animations systems (lip-sync and facial ambient animation) to play at the same time.

For the self-recognition task, the facial morph was created from a frontal picture that we took of the participant and a frontal capture of the gender matching avatar. We used the facemorpher library built on Python and OpenCV to create the 100 frames that correspond to the 1% increments of morph transformation. The videos were also assembled directly with the same library inside Ubuntu with a 5 frames per second refresh. Participants then viewed the morphing videos (20 seconds) on a wide screen (with life-size faces) to perform the self-face recognition task, again 3 times in each direction after each VR exposure (Figure 3).

3.2 Procedure

Participants ($n=20$, 4 females, age from 27 to 51), first completed a self-face recognition task [46]. Then they entered a VR setup in which they embodied that virtual avatar that tracked the movements of the participants (Figure 2). The procedure was approved by the Microsoft Research Ethics Committee and participants gave informed consent following the standards of the Declaration of Helsinki.

Participants could see their virtual body in a mirror 40 cm away. Mirrors have been used in the past to enhance body ownership [16]. After a 1 minute of accommodation to the virtual world, in their experimental condition, participants completed a short performance in which they were asked to give a pep talk. This required them to further move their hands and body in front of the mirror as if they were acting the scene in front of an audience. They were however asked to remain in place since no foot tracking was available.

Phrases of the performance were extracted from J.K. Rowling's speech at Harvard Commencement in 2008 [11]. It included the sections: "You might never fail on the scale I did, but some failure in life is inevitable."; "Every day of my working week in my early 20s I was reminded how incredibly fortunate I was, to live in a country with a democratically elected government, where legal representation and a public trial were the rights of everyone."; "I am not dull enough to suppose that because you are young, gifted and well-educated, you have never known hardship or heartbreak. Talent and intelligence never yet inoculated anyone against the caprice of the Fates."; "You might be driven by a fear of failure quite as much as a desire for success.". Each of the phrases was read a loud by the experimenter and the participant was asked to repeat them. None of the phrases were repeated between the different conditions and they were presented in a random order. The total pep talk lasted approximately 1 additional minute. After the VR experience participants completed again the self-face morphing recognition task in both directions (self to avatar, and avatar to self).



Fig. 2. A participant in front of the mirror performing the pep talk.

3.3 Stimuli

In a counterbalanced within subjects design, all participants experienced three conditions:

1. **Still Face:** during this condition participants completed the experiment while the avatar face remained static.
2. **Lip-Sync:** in this condition the avatar face was driven by a lip-sync engine that captured the speech of participants from the microphone and moved the lips of the avatar in real-time.
3. **Animated + Lip-Sync:** in this condition as well as the lip-sync animations, the avatar included additional eye blinks and subtle facial animations that increased the vividness of the avatar face. The animations were created by a professional animator and a sample of them can be seen in Figure 1.

All participants completed all conditions (which were counterbalanced in order) and after each condition they performed a self-recognition morphing task. All in all participants completed the morphing task right before the experiences (baseline), and after each of the conditions. In all cases they did the morphing task three times in each direction. See the video in the supplementary material to better observe the quality of the facial animations and lip-sync animations, as well as the morphing task.

3.4 Measurements

3.4.1 Self Identification: Morphing Task

Participants completed a self-face recognition task multiple times [46]. In this task participants watched movies that contained a morphing sequence in 1% morphing transitions, from self to a gender matched avatar (100% self to 0% self) or from the avatar to the self (0% self to 100% self).

Participants were instructed to stop the movie when they felt the face was starting to look more like the self than other or vice versa, depending on the morphing direction displayed in the movie. They repeated this test 3 times in each direction (see Figure 3 for example frames cover morphs which are overlaid with summative results as described in Section 4.1). Each movie lasted 20 seconds and it was presented on a desktop screen in life-size. This morphing method can help explore the levels of self-other face distinction that will vary as a result of the enfacement illusion [7].

3.4.2 Gaze Tracking

Using a ray-casting technique we were able to evaluate where participants were looking while performing in front of the mirror. The gaze maps were calculated as a straight line from the direction at which the HMD was oriented. Our system did not support eye tracking, hence the gaze is an estimate only. However, previous work has shown that HMD direction is a good estimate of eye tracking [4, 13, 18, 25]

3.4.3 Embodiment

We used a subset of the standard embodiment questions to evaluate the embodiment of participants [15]. For example, questions regarding touch were not used because the current experiment did not include

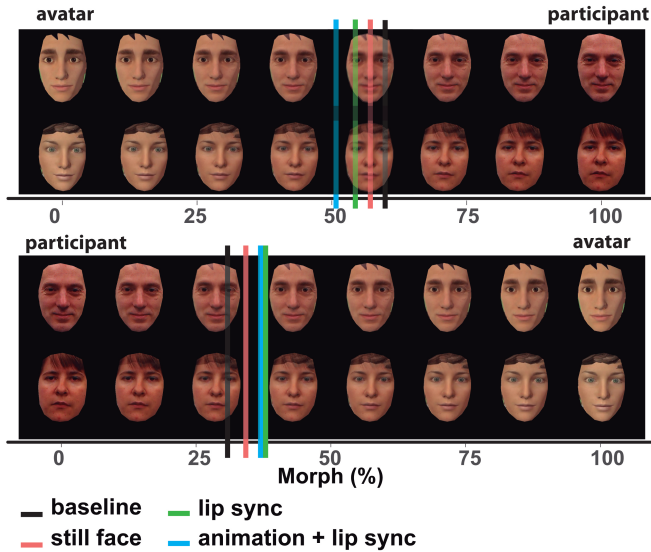


Fig. 3. The participant’s face was morphed to the gender matched avatar face for the self-recognition task. The vertical lines indicate the average moment at which participants stopped or started self-recognising in the test for all conditions in both directions.

tactile stimulation. The questionnaire was completed after each condition and had a 7-Likert scale for each question in which participants stated their level of agreement with the following questions from -3 (completely disagree) to 3 (completely agree):

1. I felt as if the virtual body I saw when I looked down was my body
2. It seemed as if I might have more than one body
3. It felt like I could control the virtual body as if it was my own body
4. I felt as if my body was located where I saw the virtual body
5. I felt out of my body
6. I felt as if my (real) body were drifting towards the virtual body or as if the virtual body were drifting towards my (real) body
7. I felt like I was wearing different clothes from when I came to the laboratory

4 RESULTS

4.1 Self Identification: Morphing Task

One characteristic of self-recognition morphing is that typically the task has a fading effect depending on whether the video morphs from the self to the other or from the other to the self (see Figure 3 which plots the difference in means of recognition point and Figure 4 which plots hysteresis results across all participants). This fading is represented by the different times at which people will stop the task which is described by an hysteresis. The smaller the area of the hysteresis the harder the self-identification task was, as the ability to distinguish the two ends of the morph was lower. In essence, in our experiment a smaller area means a higher enfacement illusion.

We find a self-recognition hysteresis for the baseline (the measure that was recorded before the VR exposure started) as well as for all the conditions (Figure 4).

We ran a Friedman rank sum test on the hysteresis data (response times in Morph %, values in Table 1) to further understand the effects of the VR exposure on the self-recognition on avatar faces. We test the response time on a paired manner, within subjects, in both directions,

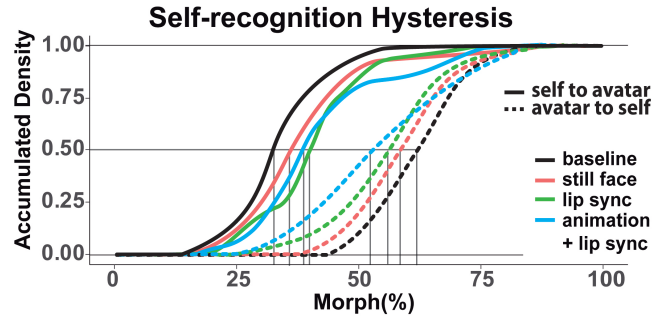


Fig. 4. Hysteresis results grand averages across all participants. A reduction on the hysteresis was found after embodying the avatar and the hysteresis was even smaller even when the avatar had lip-sync and idle facial expressions. Dotted lines represent the morphing from the avatar to the self. Solid lines represent the morphing from the self to the avatar.

Table 1. Self-Recognition Response Times (Morph % and SD)

Condition Direction	Baseline	Still Face	Lip-Sync	Animated + Lip-Sync
IN	62 ± 8.5	59 ± 10	55.5 ± 11	54 ± 13.5
OUT	33 ± 8.5	38 ± 12.5	42.5 ± 10.5	41.5 ± 13.5

from the self to the avatar (*OUT*) and from the avatar to the self (*IN*). We found a significant effect for conditions (still face, lip-sync and animated + lip-sync) in both directions *IN* $\chi^2 = 12.7, df = 3, p = 0.005$, *OUT* $\chi^2 = 19.8, df = 3, p = 0.0002$. We performed a post-hoc pairwise comparisons using Conover tests with Bonferroni adjustments to explore how the different conditions compared. On one hand, after the VR exposure, participants stopped the morphing earlier going in the *IN* direction as compared to the Baseline for all conditions, Table 1, $p = 0.0009$. The lip-sync and animated+lip-sync conditions also showed earlier stops, meaning they self recognized themselves earlier in the avatar to self morph, than the still face condition $p = 0.01$. There were no significant differences between lip-sync and animated+lip-sync conditions for the *IN* morph. However in the other direction *OUT*, participants in the animated+lip-sync condition took significantly longer, than in the lip-sync only, to stop the morphing going from themselves into the avatar ($p = 0.029$). This bias towards self-recognition in the avatar can be understood as a greater enfacement illusion.

4.2 Gaze Tracking

The gaze maps in Figure 5 use warmer colours on avatar parts that were often gazed at, and colder colours for less attended parts.

Using paired comparisons Conover tests we found significant differences in the percentage of gaze of the Chest and Face. Participants looked significantly more to the face during the animated + lip-sync or the lip-sync conditions than during the still face condition ($p < 0.0001$). Instead participants in the still face condition gazed more often to their chest area ($p = 0.004$). In essence, with the animated + lip-sync condition participants reduced the number of times they gazed at their chest, even more than the lip-sync condition ($p = 0.014$).

We hypothesise that this prolonged observation of the avatar face in the lip-sync and animated + lip-sync condition impacted the self-recognition.

4.3 Embodiment

The scores to the embodiment questionnaire in all conditions are shown in Figure 7.

Doing a paired comparison (Friedman and Conover test with Bonferroni adjustment) we found participants were most embodied (q1) in the animated + lip-sync and lip-sync conditions than in the still face

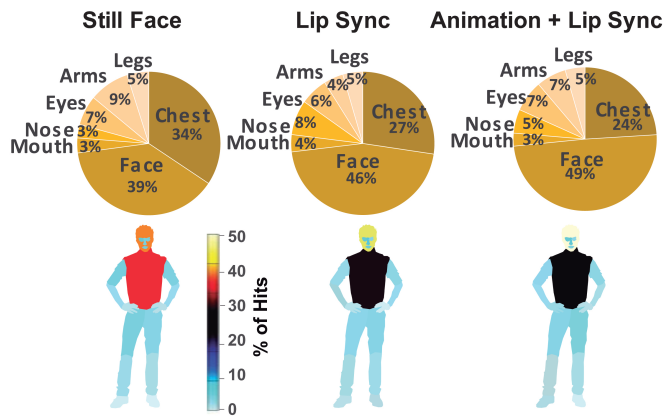


Fig. 5. Gaze distribution results. Using a ray-casting technique out of the HMD we evaluated which parts of the avatar were more frequently looked at during the experiment in each condition.

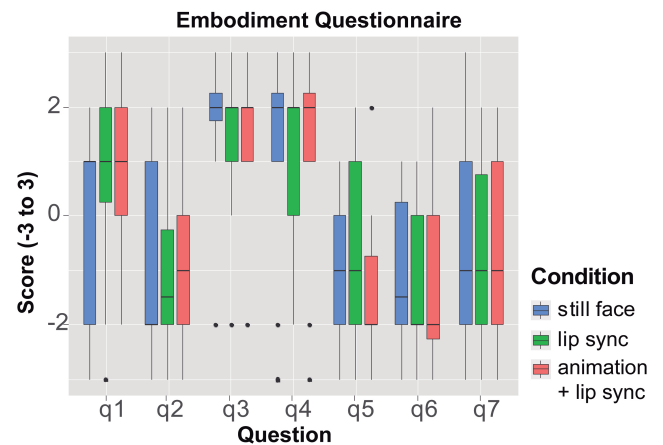


Fig. 7. Scores of the embodiment questions.

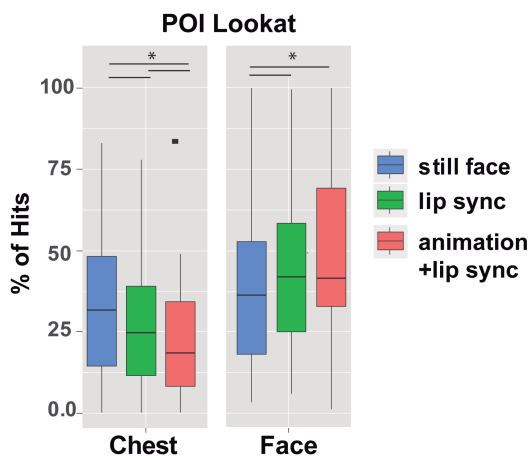


Fig. 6. Comparison of gazing percentage of the Chest and Face under different conditions.

condition ($p = 0.006$). But there were no differences between these two conditions.

Further Wilcoxon paired analysis between control questions (q1 and its control q5) showed significant differences between the two questions for all conditions ($p = 0.004$). Other questions did not show significant differences, nor did the aggregated score.

5 DISCUSSION

The main findings show the significant relationship between the way in which the avatar’s face is animated and the amount of avatar enfacement exhibited by the participants. Self-recognition on the avatar was maximised when face animations were used, even though those animations were idling animation not driven by the user’s actual facial movement.

We find that merely the embodiment of an avatar in front of a mirror already increased the enfacement of the avatar when compared to the baseline self-recognition test. We hypothesise that enfacement and embodiment are closely linked, and the former might not be possible to achieve if the avatar is not correctly embodied. Along those lines, the animated + lip-sync and lip-sync conditions not only produced higher enfacement levels but also were the conditions in which participants reported higher levels of embodiment, as shown by the differences in the questionnaire. We believe that the lack of further differences between conditions in embodiment score was due to the fact that high embodiment levels were achieved on all conditions. But embodiment was even higher on the lip-sync and animated + lip-sync.

Furthermore, when the users embodied avatars that exhibited both lip-sync and face animated with idling mechanisms, they experienced greater fading effect on the morphing task, that is they reduced the natural hysteresis of self-identification during the morphing task. This means that they internally underwent an enfacement of the virtual avatar’s face. This has additional interest given the context of the experiment, in which the virtual face was not a look-a-like nor a photorealistic representation.

One possible reasoning is that the reduction in the discrepancies between the virtual representation and the motions of the user was the cause of the higher embodiment and ulterior increased self-recognition on the avatar. Similar to the work of Ma et al. who showed that enfacing smiles made participants happier [23], in our setup, participants could have had a stronger emphatic connection with the animated avatar. In fact, in the conditions where the lips of the avatar mimicked the user’s speech and there was animation of the face, the avatar looked more alive.

On the other hand, it may be that adding animations to the face attracts more attention of the user to the face, and through larger exposure we improved the user’s familiarity with the avatar’s appearance. In our experiment we showed that the lip-sync and the animations attracted more gaze time towards the face while the static face, with no natural motion in it, did not.

Either way, the current setup was able to introduce a larger self-identification on the avatar, despite the fact that the avatar facial animations were not always triggered by the participant.

6 CONCLUSIONS

The current results have significant implications for the VR and perception communities. On one hand, as we demonstrate, facial animation can increase self-identification on virtual characters in VR. Our results complement prior work on embodiment and facial animation. Interestingly, increasing self-identification on virtual characters might be key to increase accountability and good behaviours.

On the other hand, our results are applicable now as they do not require of any additional technologies to be deployed. Current VR devices already integrate a microphone and the animations can be baked to the system, which could simplify a faster adoption of these self-identification tools.

Future work should look at using actual sensing of the user to animate the eyes or other parts of the face, and measure enfacement levels. New commercial HMDs, such as the Vive Pro Eye, can enable correct rendering of the avatar motions, which may again improve the reported results. Additionally, lip-sync systems for avatar animation keep evolving and they are currently reaching human perception levels [6]. Hence we hypothesise that lip-sync will become an even more common form of facial animation.

In addition, other tests, along the lines of predefined facial animations, could look at the use of more abstract or affective motions on the face, to either attract attention or provoke emotion contagion. These could also increase the self-identification and enfacement illusion.

In particular we envision our results will increase in importance when communication between users in VR becomes ubiquitous. Social VR platforms such as AltSpace, or Facebook Spaces strive to facilitate connection between users, represented by 3D avatars. Usage of animated lip-syncing and eye tracking has a potential to improve self-identification, communication, empathy and even spread moods and behaviours in immersive VR.

REFERENCES

- [1] O. Alexander, M. Rogers, W. Lambeth, J.-Y. Chiang, W.-C. Ma, C.-C. Wang, and P. Debevec. The Digital Emily Project: Achieving a Photorealistic Digital Actor. *IEEE Computer Graphics and Applications*, 30(4):20–31, July 2010. doi: 10.1109/MCG.2010.65
- [2] M. A. Apps and M. Tsakiris. The free-energy self: A predictive coding account of self-recognition. *Neuroscience and biobehavioral reviews*, 0:85–97, Apr. 2014. doi: 10.1016/j.neubiorev.2013.01.029
- [3] F. Argelaguet, L. Hoyet, M. Trico, and A. Lecuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Virtual Reality (VR)*, pp. 3–10, Mar. 2016. doi: 10.1109/VR.2016.7504682
- [4] D. Borland, T. Peck, and M. Slater. An evaluation of self-avatar eye movement for virtual embodiment. *IEEE transactions on visualization and computer graphics*, 19(4):591–596, 2013.
- [5] M. Botvinick, J. Cohen, and others. Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669):756–756, 1998.
- [6] D. Cudeiro, T. Bolkart, C. Laidlaw, A. Ranjan, and M. J. Black. Capture, learning, and synthesis of 3d speaking styles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 10101–10111, 2019.
- [7] M. Dobricki and B. J. Mohler. Self-identification with another’s body alters self-other face distinction. *Perception*, 44(7):814–820, 2015.
- [8] T. J. Dodds, B. J. Mohler, and H. H. Bühlhoff. Talk to the virtual hands: Self-animated avatars improve communication in head-mounted display virtual environments. *PLoS one*, 6(10):e25759, 2011.
- [9] C. T. Fuentes, C. Runa, X. A. Blanco, V. Orvalho, and P. Haggard. Does My Face FIT?: A Face Image Task Reveals Structure and Distortions of Facial Feature Representation. *PLOS ONE*, 8(10):e76805, Oct. 2013. doi: 10.1371/journal.pone.0076805
- [10] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed, and M. A. Sasse. The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’03, pp. 529–536. ACM, New York, NY, USA, 2003. event-place: Ft. Lauderdale, Florida, USA. doi: 10.1145/642611.642703
- [11] T. H. Gazette. Text of J.K. Rowling’s speech. <https://news.harvard.edu/gazette/story/2008/06/text-of-j-k-rowling-speech/>, 2008. [Online; accessed 22-May-2019].
- [12] M. Gonzalez-Franco, A. I. Bellido, K. J. Blom, M. Slater, and A. Rodriguez-Fornells. The neurological traces of look-alike avatars. *Frontiers in human neuroscience*, 10:392, 2016.
- [13] M. Gonzalez-Franco and P. A. Chou. Non-linear modeling of eye gaze perception as a function of gaze and head direction. In *2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 275–280. IEEE, 2014.
- [14] M. Gonzalez-Franco and J. Lanier. Model of illusions and virtual reality. *Frontiers in psychology*, 8:1125, 2017.
- [15] M. Gonzalez-Franco and T. C. Peck. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5:74, 2018.
- [16] M. Gonzalez-Franco, D. Perez-Marcos, B. Spanlang, and M. Slater. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE virtual reality conference (VR)*, pp. 111–114. IEEE, 2010.
- [17] K. Kiltner, R. Groten, and M. Slater. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, Nov. 2012. doi: 10.1162/PRES.a.00124
- [18] S. Kishore, M. González-Franco, C. Hintemüller, C. Kapeller, C. Guger, M. Slater, and K. J. Blom. Comparison of ssvp bci and eye tracking for controlling a humanoid robot in a social environment. *Presence: Teleoperators and virtual environments*, 23(3):242–252, 2014.
- [19] E. Kokkinara and R. McDonnell. Animation Realism Affects Perceived Character Appeal of a Self-virtual Face. In *Proceedings of the 28th ACM SIGGRAPH Conference on Motion in Games*, MIG ’15, pp. 221–226. ACM, New York, NY, USA, 2015. doi: 10.1145/2822013.2822035
- [20] H. Li, L. Trutoiu, K. Olszewski, L. Wei, T. Trutna, P.-L. Hsieh, A. Nicholls, and C. Ma. Facial Performance Sensing Head-mounted Display. *ACM Trans. Graph.*, 34(4):47:1–47:9, July 2015. doi: 10.1145/2766939
- [21] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero. Learning a model of facial shape and expression from 4d scans. *ACM Transactions on Graphics (TOG)*, 36(6):194, 2017.
- [22] L. Lin and S. Jrg. Need a Hand?: How Appearance Affects the Virtual Hand Illusion. In *Proceedings of the ACM Symposium on Applied Perception*, SAP ’16, pp. 69–76. ACM, New York, NY, USA, 2016. event-place: Anaheim, California. doi: 10.1145/2931002.2931006
- [23] K. Ma, R. Sellaro, D. P. Lippelt, and B. Hommel. Mood migration: How enfacing a smile makes you happier. *Cognition*, 151:52–62, June 2016. doi: 10.1016/j.cognition.2016.02.018
- [24] L. Maister, M. Slater, M. V. Sanchez-Vives, and M. Tsakiris. Changing bodies changes minds: owning another body affects social cognition. *Trends in Cognitive Sciences*, 19(1):6–12, Jan. 2015. doi: 10.1016/j.tics.2014.11.001
- [25] S. Marwecki, A. D. Wilson, E. Ofek, M. Gonzalez Franco, and C. Holz. Mise-unseen: Using eye tracking to hide virtual reality scene changes in plain sight. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST 19, p. 777789. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3332165.3347919
- [26] M. Mori. Bukimi no tani [the uncanny valley]. *Energy*, 7:33–35, 1970.
- [27] M. Mori, K. F. MacDorman, and N. Kageki. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine*, 19(2):98–100, June 2012. doi: 10.1109/MRA.2012.2192811
- [28] S. Y. Oh, J. Bailenson, N. Krämer, and B. Li. Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PLoS one*, 11(9):e0161794, 2016.
- [29] K. Olszewski, J. J. Lim, S. Saito, and H. Li. High-fidelity Facial and Speech Animation for VR HMDs. *ACM Trans. Graph.*, 35(6):221:1–221:14, Nov. 2016. doi: 10.1145/2980179.2980252
- [30] V. Orvalho, P. Bastos, F. I. Parke, B. Oliveira, and X. Alvarez. A Facial Rigging Survey. In *Eurographics*, 2012. doi: 10.2312/conf/EG2012/stars/183-204
- [31] F. I. Parke and K. Waters. *Computer Facial Animation*. A K Peters/CRC Press, Wellesley, Mass, 2 edition ed., Sept. 2008.
- [32] V. I. Petkova and H. H. Ehrsson. If I Were You: Perceptual Illusion of Body Swapping. *PLoS ONE*, 3(12):e3832, 2008. doi: 10.1371/journal.pone.0003832
- [33] M. V. Sanchez-Vives, B. Spanlang, A. Frisoli, M. Bergamasco, and M. Slater. Virtual Hand Illusion Induced by Visuomotor Correlations. *PLoS ONE*, 5(4):e10381, 2010. doi: 10.1371/journal.pone.0010381
- [34] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. “These Are Not My Hands!”: Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI ’17, pp. 1577–1582. ACM, New York, NY, USA, 2017. event-place: Denver, Colorado, USA. doi: 10.1145/3025453.3025602
- [35] J. Seyama and R. S. Nagayama. The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence: Teleoperators and Virtual Environments*, 16(4):337–351, July 2007. doi: 10.1162/pres.16.4.337
- [36] A. Sforza, I. Bufalari, P. Haggard, and S. M. Aglioti. My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Social Neuroscience*, 5(2):148–162, Apr. 2010. doi: 10.1080/17470910903205503
- [37] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. V. Sanchez-Vives. Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, 2, 2008. doi: 10.3389/neuro.09.006.2008
- [38] B. Spanlang, J.-M. Normand, D. Borland, K. Kiltner, E. Giannopoulos, A. Pomes, M. Gonzalez-Franco, D. Perez-Marcos, J. Arroyo-Palacios, X. N. Muncunill, et al. How to build an embodiment lab: achieving body representation illusions in virtual reality. *Frontiers in Robotics and AI*, 1:9, 2014.
- [39] W. Steptoe and A. Steed. High-Fidelity Avatar Eye-Representation. In *2008 IEEE Virtual Reality Conference*, pp. 111–114, Mar. 2008. doi: 10.

- [40] C. Strassberger and R. Sikkema. Democratising Mocap: Real-time Full-performance Motion Capture with an iPhone X, Xsens, and Maya. In *ACM SIGGRAPH 2018 Real-Time Live!*, SIGGRAPH '18, pp. 2:1–2:1. ACM, New York, NY, USA, 2018. event-place: Vancouver, British Columbia, Canada. doi: 10.1145/3229227.3229236
- [41] A. Tajadura-Jiménez, D. Banakou, N. Bianchi-Berthouze, and M. Slater. Embodiment in a Child-Like Talking Virtual Body Influences Object Size Perception, Self-Identification, and Subsequent Real Speaking. *Scientific Reports*, 7(1):9637, Aug. 2017. doi: 10.1038/s41598-017-09497-3
- [42] A. Tajadura-Jiménez, M. R. Longo, R. Coleman, and M. Tsakiris. The person in the mirror: using the enfacement illusion to investigate the experiential structure of self-identification. *Consciousness and cognition*, 21(4):1725–1738, Dec. 2012. doi: 10.1016/j.concog.2012.10.004
- [43] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner. FaceVR: Real-Time Gaze-Aware Facial Reenactment in Virtual Reality. *ACM Trans. Graph.*, 37(2):25:1–25:15, June 2018. doi: 10.1145/3182644
- [44] L. C. Trutoiu, E. J. Carter, I. Matthews, and J. K. Hodgins. Modeling and Animating Eye Blinks. *ACM Trans. Appl. Percept.*, 8(3):17:1–17:17, Aug. 2011. doi: 10.1145/2010325.2010327
- [45] L. C. Trutoiu, J. K. Hodgins, and J. F. Cohn. The temporal connection between smiles and blinks. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pp. 1–6, Apr. 2013. doi: 10.1109/FG.2013.6553736
- [46] M. Tsakiris. Looking for Myself: Current Multisensory Input Alters Self-Face Recognition. *PLoS ONE*, 3(12), Dec. 2008. doi: 10.1371/journal.pone.0004040
- [47] Y. Yuan and A. Steed. Is the rubber hand illusion induced by immersive virtual reality? In *Virtual Reality Conference (VR), 2010 IEEE*, pp. 95–102. IEEE, 2010.
- [48] K. Zibrek, E. Kokkinara, and R. McDonnell. The Effect of Realistic Appearance of Virtual Characters in Immersive Environments - Does the Character's Personality Play a Role? *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1681–1690, Apr. 2018. doi: 10.1109/TVCG.2018.2794638