Dancing with the Avatars: Feedforward Learning from Self-Avatars

Isabel Fitton
isf21@bath.ac.uk
University of Bath
Bath, United Kingdom

Jeremy Dalton
jeremy.dalton@pwc.com
PwC
London, United Kingdom

Michael J. Proulx
mjp51@bath.ac.uk
University of Bath
Bath, United Kingdom

Christof Lutteroth
cl2073@bath.ac.uk
University of Bath
Bath, United Kingdom

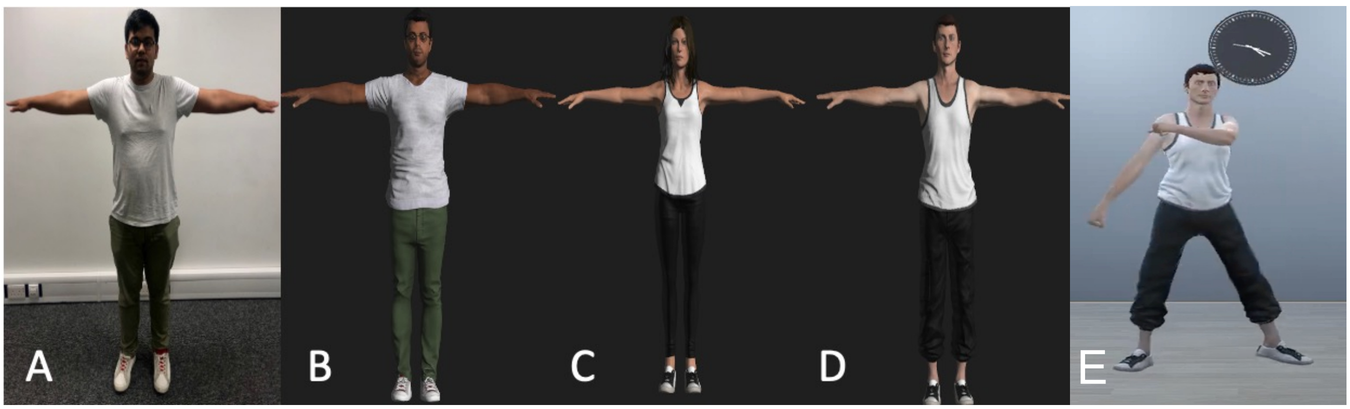

Figure 1: Avatars were generated by Adobe Fuse CC: Self avatars (B) were produced in the likeness of a photograph of the individual (A), while gender-matched generic avatars (C, D) were made using readily available standard components. Avatars were then animated to display Hip-Hop dance moves (E). Users watched a video of the self or generic avatar to practice and learn the dance moves.

ABSTRACT

Feedforward is a training technique where people observe themselves perform a new skill to promote rapid learning, commonly implemented via video self-modelling. Avatars provide a unique opportunity to self-model skills an individual's physical self cannot yet perform. We investigated the use of avatars in video-based learning and explore the potential of feedforward learning from self-avatars. Using modern dancing as a skill to learn, we compared the user experience when learning from a human training video and an avatar training video, considering both self-avatars (n=8) and gender-matched generic avatars (n=8). Our results indicate that learning from avatars can improve the user experience over learning from a human in a video, providing attentional and motivational benefits. Furthermore, self-avatars make the training more relatable and immersive than generic avatars. We discuss the implications from this preliminary work, highlighting methodological considerations for feedforward learning from avatars and promising future work.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Laboratory experiments**; • **Applied computing** → **Education**.

KEYWORDS

Feedforward, Self-Modelling, Avatar Customisation, Skills Training

ACM Reference Format:
Isabel Fitton, Jeremy Dalton, Michael J. Proulx, and Christof Lutteroth. 2022. Dancing with the Avatars: Feedforward Learning from Self-Avatars. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3491101.3519732

CHI '22 Extended Abstracts, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9156-6/22/04…$15.00
https://doi.org/10.1145/3491101.3519732

1 INTRODUCTION

Psychomotor learning is a broad term incorporating any short to long-term change in motor performance induced by training, in essence attaining more effective movement and skill [35, 57]. Psychomotor skills are those which involve physical actions e.g. manipulating objects, control of the body, and perceptual acuity. Learning to dance is a rich example of psychomotor learning [33] where dancers must transfer visual information into motor action, maintaining control over their body, in order to replicate the dance moves [12]. Psychomotor skills are taught through observational learning whereby a learner, with the necessary motor ability and motivation, attends to a 'model' demonstrating the skill and learns a new behavioural pattern [5, 6]. In the context of dance, the learner observes a dance instructor's model of movement and mirrors it [26]. Models which are more closely aligned to the observer, on dimensions such as appearance, personality traits, or demographics, are said to have a more powerful impact on learning [55]. Ordinarily an instructor may or may not happen to be matched with the identity of the learner, e.g. of the same gender or ethnicity, making it difficult to enhance the learning outcomes. However, digital learning provides the opportunity to manipulate the learning experience to maximise success.

Feedforward learning aims to do this by manipulating what is observed, allowing learners to observe a model that they self-identify with performing the desired skill successfully, to promote increased motivation and rapid learning [21]. Feedforward is a continuous effect observed gradually as people identify more with the model. Video is the predominant medium in which the principles of feedforward learning have been explored [22]. Video Self-Modelling (VSM) is achieved by recording an individual's performance and editing the footage to remove mistakes to produce a short video depicting a 'perfect' performance [16, 22]. Individuals watch their own successful performance several times until they are able to improve towards the success they observed. VSM has been used for the acquisition and improvement of many psychomotor skills. For example, right-footed footballers shown videos of themselves with equally good left-foot abilities make significant improvements in left-foot kicks [59], supporting VSM as an efficacious method of feedforward learning.

Contrary to this success, VSM remains relatively unused despite dance being taught using video tutorials [27, 45] There are shortcomings of this method. First, it requires recordings of an individual which may be difficult to obtain particularly in remote learning scenarios, and even when used alongside physical practice editing videos for VSM remains cumbersome. Furthermore it is only possible to portray an individual's existing skills with minor stumbles or mistakes edited out, limiting how much can be learned from a single video. Avatars offer a unique opportunity to overcome these limitations, allowing the simulation of an optimal future state which the individual's physical self cannot yet achieve [29]. It is possible that feedforward learning with self-avatars may have even greater potential as a training method than VSM.

Whilst we speculate the use of avatars in feedforward learning is plausible and beneficial, feedforward learning with self-avatars is largely unresearched. In this paper, we explore the potential role of avatars for video-based learning, and investigate the role of self-avatars in feedforward learning by comparing training outcomes when learning from a self-avatar and a gender-matched generic avatar (see Figure 1). In a user study (n = 16) participants learned two different Hip-Hop dance sequences, one from a human training video and one from a self/generic avatar training video. Regarding these training methods, we address the following research questions:

- RQ1 What is the user experience of learning from an avatar training video compared to a human training video?
- RQ2 How does dance training with self-avatars compare with dance training with generic avatars?

2 RELATED WORK

Learning to dance can increasingly be accomplished in digital domains [17, 25], and is commonly taught using video tutorials [27, 45], especially over the course of the pandemic [9, 36]. Although VSM has not been reportedly used for teaching dance, it has been used successfully for other skills in the psychomotor domain. Its success demonstrates the value of self-modelling in feedforward learning of psychomotor skills, with video providing the opportunity to observe a self-model. However, it is unlikely that raw videos of the self that can be edited for VSM are always available. Virtual avatars can be used to provide a representation of the self when video content is not available, and in different contexts have demonstrated positive effects of self-avatars compared to other avatars on the user. Avatar-based feedforward would allow greater flexibility in its use, and here we identify related work in these areas suggesting that being able to self-identify with an avatar is crucial for obtaining successful feedforward effects.

Video Self-Modelling (VSM). Successful for the acquisition of many skills; including social, behavioural, academic, and psychomotor skills [13, 22, 30, 40, 44, 52], VSM is an effective implementation of feedforward learning. Simply reviewing a 2-minute video, 6 times, over a period of two weeks is sufficient to lead to significant improvements in a self-modelled behaviour in 78% of participants [23]. VSM has been successful for learning and improving complex psychomotor skills such as gymnasts' pommel horse routines [8] and swimming [58], suggesting that feedforward learning is a valid technique for the acquisition of other psychomotor skills such as dance. Self-modelling videos are more effective than other-modelling videos, e.g. where a learner observers a video of someone else [58]. This is explained by the mirror neurons involved in encoding motor components of an action and reconfigured for imitation [15] being most active when observing a self-model [51]. Hence feedforward effects are observed more as an individual identifies more with the model, making a self-model the most effective.

Avatar Self-Modelling. It is widely recognised that avatar appearance impacts users, with certain physical characteristics giving rise to assumptions of their capabilities which influences user behaviour [47], for example by playing better as an athletic avatar [42, 43], perceiving physical tasks as less arduous with muscular avatars [32], or displaying greater confidence with a taller avatar [64]. Self-avatars are those whose appearance has been customised to resemble the user [3]. Customisation of avatars is related to increased levels of enjoyment, immersion, effort, and intrinsic motivation [60], with individuals spending longer in an infinite runner game when using a customised avatar that they identified with compared to a randomly assigned avatar [10]. Self-avatars have been shown to positively influence performance, for example users perform significantly better on a Wii Sports Tennis game when playing as a self-avatar compared to a dissimilar avatar [46]. Observing a self-avatar has also been shown to encourage behaviour change, whereas an avatar only matched on sex and age did not lead to change [28]. Viewing self-avatars from a third-person perspective is associated with more pronounced effects [37] and has been used as a mechanism to increase intrinsic motivation and performance in exergames through self-competition [7, 34]. Avatar feedforward has not yet been implemented for learning psychomotor skills, yet research comparing self and other avatars suggests that a self-avatar, which the user identifies with, could be an effective model to learn from[1].

Avatar Identification. Identification is an imaginative process in which the individual connects cognitively and emotionally with the avatar [18]. It involves taking the perspective of the avatar, adopting the same goals, and temporarily replacing one's self-awareness with that of the avatar [18]. Self-identity encompasses all aspects of an individual, including physical and mental characteristics such as appearance, voice, and personality [47]. Feelings of identification are greatest when interacting with self-avatars as opposed to another avatar. For example, physiological arousal was greatest when participants were exposed to a running avatar which looked like them as opposed to someone else [29, 41]. Thus, it is evident that similarity in terms of appearance and demographics are key factors driving the process of identification with an avatar [19, 34]. We will manipulate the avatar appearance and define identification as the extent to which the avatar is viewed as a self-model, using the physical similarity, liking, and wishful identification dimensions of the polythetic model of player-avatar identification (PAI) [19].

3 METHODOLOGY

We investigated learning of Hip-Hop dance moves from a human and an avatar. The human training video used was a demonstration of three beginner dance moves by professional dancer Mihran Kirakosian [31]. The avatar training videos followed the same format, with the avatar demonstrating three different beginner moves. In a between-participant design, participants were randomly allocated to the self-avatar or generic avatar condition.

Avatars (see Figure 1) were produced using blend shapes in Adobe Fuse CC [39]. The male and female generic avatars were constructed of non-customised components, dressed in loose-fit clothing typical for Hip-Hop dancing. The self-avatars were created in collaboration with each participant to produce an avatar in their own likeness. To shorten the customisation time required by participants, the experimenter was provided with a photo of the participant and used this to model a resembling character (focusing on gender, hair style/colour, skin tone, and facial features). Following this, each participant in the self-avatar condition had an avatar customisation session conducted over Zoom for 30 minutes in which they were given control of the experimenter's screen and guidance of how to use Adobe Fuse CC so that they could make further adjustments to their self-avatar. Most participants were pleased with the facial likeness and focused their time on personalising the body-type and character's clothing. The avatars were imported into the virtual dance studio scene created using freely available assets in Unity. The avatars were animated to demonstrate three Hip-Hop moves in time with the soundtrack added to the virtual dance studio environment. A screen recording software was used to produce a video of each avatar.

3.1 Outcome Variables

A main outcome measure in this research was dance skill. Perceived skill level was measured at the start by asking participants about their prior dance experience ('Please list any/all of your dance experience to date') and rating their own dancing abilities (e.g. 'I am a good dancer') on a 5-point Likert scale. During the study, dance skill was measured by scoring the participants' performances. A bespoke grade of execution was produced for each dance, which resembled the well-used RCampus iRubric for dance [49], but taking a move-by-move approach as with figure skating [56]. The grade of execution contained four categories; the overall quality of the movement ('performance'), memory of the routine ('composition'), the ability to reproduce a hip-hop style ('technical skills'), and the ability to maintain an 8-count beat ('rhythm/tempo'). Each category was graded on a 5-point scale, with 5 indicating no faults, 4 indicating a minor fault, 3 indicating a major fault, 2 indicating two major faults, and 1 indicating more than two major faults. Each dance move was scored out of 20, and each three-move routine scored out of 60.

Competence, that is, how confident people felt in their ability to dance (e.g. 'I think I am pretty good at this activity'), was measured after each dance using the Perceived Competence subscale of the well-validated Intrinsic Motivation Inventory (IMI) [38], which scores perceived competence from 1 (lowest) to 7 (highest). Furthermore, we used the Interest/Enjoyment (e.g. 'I enjoyed doing this'), Effort/Importance (e.g. 'I put a lot of effort into this'), and Pressure/Tension (e.g. 'I felt very tense whilst doing this activity') subscales of the IMI to measure intrinsic motivation, i.e. the enjoyment participants gained from dancing as opposed to motivation from external pressures and rewards [53]. Avatar identification was measured using the Physical Similarity (e.g. 'I physically resemble this avatar'), Liking (e.g. 'I like this avatar'), and Wishful Identification subscales (e.g. 'Sometimes I wish I could be more like this avatar') of the PAI [19], to verify whether the self-avatar manipulation was successful.

To check whether individual differences in immersive tendency affected learning from the avatars, we used the 18-item Immersive Tendency Questionnaire (ITQ)[63] as a pre-study measure to indicate how easily an individual becomes immersed (e.g. 'Do you easily become deeply involved in movies or tv dramas?'). After the avatar video training, the Film Immersive Experience Questionnaire (Film IEQ) [50] was used to assess how immersive participants found the training video (e.g. 'To what extent did the video hold your attention'). The original Film IEQ is a 24 item questionnaire, items 17, 18 and 19 were not relevant and were excluded. Each item is scored on a 7-point Likert scale, all items are summed to produce a score.

Finally, six open-ended questions were used to further understand how participants felt they performed (e.g. 'What did you think of your performance after the main training exercise?' *i.e. your final performance after learning from the avatar*), how they found learning from an avatar (e.g. 'How do you feel about learning from a digital model (avatar)?'), and what they liked/disliked about the experience (e.g. 'What, if anything, did you dislike about the dance training experience (learning from an avatar)?').

3.2 Procedure

Due to the COVID-19 pandemic this study was conducted via a Microsoft Teams meeting between the experimenter and each participant. Participants who signed up were randomly allocated to the self-avatar (n = 8) or generic avatar (n = 8) training condition. Those in the self-avatar condition completed the pre-study avatar customisation process. Those in the generic avatar condition were allocated a gender-matched generic avatar (Figure 1C/D). Each experiment session lasted 60 minutes. A Qualtrics survey was used to provide participants with information, instructions, and questionnaire measures. Initially participants completed a demographics questionnaire and a 5-minute warm-up video to stretch before moving on to the dance training. In part one, participants were given 5 minutes to practice with the 18-second human training video. After practicing, participants were required to perform what they had just learned. Microsoft Teams' meeting recording function was used to capture this performance, to avoid affecting participant behaviour the experimenter did not observe the live performance, the recording was reviewed and scored by the experimenter afterwards. Participants then returned to the online survey to answer questions about their experience of this training exercise. Part two followed the same procedure of 5-minute practice, recorded performance and questionnaire measures but with the 18-second avatar training video. Afterwards participants were fully debriefed, gave final consent for their data to be used, and could enter into a prize draw to win a £20 Amazon voucher.

3.3 Participants

Volunteers were recruited using a call for participants distributed via the university noticeboard, social media, and mailing lists. Eligibility criteria excluded any participants who were not aged 16+, physically fit to dance, and did not have access to a strong Internet connection from a computer/laptop equipped with a webcam. In total 16 participants were recruited (11 M, 5 F), aged 17 - 48 (M = 25.688, SD = 7.804).

4 RESULTS

4.1 Quantitative Results

The data met the assumptions for conducting independent and paired samples t-tests, the data were normally distributed and homogeneous as assessed by the Shapiro-Wilk ($p > .05$) and Levene's test ($p > .05$). Dance performance scores were not normally distributed ($p = .019$) so a non-parametric Wilcoxon signed-ranks test was used. Internal reliability tests suggests that the bespoke measure for perceived dance ability ($\alpha = 0.730$) and the adjusted Film IEQ ($\alpha = 0.856$) were highly reliable and in the acceptable range. Equality of groups tests revealed there were no significant differences in existing dance skill ($M_{diff} = 0.625; t(14) = 0.919, p = .373, d = 0.460$) or ability ($M_{diff} = 0.05; t(14) = 0.148, p = .884, d = 0.074$) between the two participant groups. There was also no significant difference in immersive tendencies ($M_{diff} = 3.625; t(14) = 0.696, p = .498, d = 0.348$). However, the self-avatar group ($M = 43.714, SD = 2.215$) significantly outperformed the generic-avatar group ($M = 37.625, SD = 5.502$) in the first dance learned from the human training video ($t(13) = 2.731, p = .017, d = 1.413$, Figure 2a). The avatar manipulation was successful with participants in the self-avatar condition reporting significantly greater avatar similarity ($M_{diff} = 2.075; t(14) = 2.910, p = .006^{**}, d = 1.455$), liking ($M_{diff} = 1.375; t(14) = 2.439, p = .014^{*}, d = 1.220$), and wishful identification ($M_{diff} = 1.708; t(14) = 2.243, p = .021^{*}, d = 1.121$) than those in the generic avatar condition (Figure 2d). Dance skills improved as a result of the training, with participants' reported dance ability being significantly higher after training than before ($M_{diff} = 1.55, t(15) = 4.810, p < .001^{**}, d = 1.203$, Figure 2b).

Human vs. Avatar Dance Training. There were no significant differences in performance scores (Human $Mdn = 41.000$, Avatar $Mdn = 42.000; T = 38.000, p = .624, r = 0.165$, Figure 2a), perceived competence (Human $M = 2.542, SD = 0.973$, Avatar $M = 2.927, SD = 0.901; t(15) = 1.572, p = .137, d = 0.393$, Figure 2c), or pressure/tension (Human $M = 3.663, SD = 1.188$, Avatar $M = 3.275, SD = 1.133; t(15) = 1.526, p = .148, d = 0.381$, Figure 2e) after learning from the human training video compared to the avatar training video. Interest/Enjoyment (Human $M = 4.446, SD = 1.221$, Avatar $M = 4.848, SD = 1.199; t(15) = 2.390, p = .030, d = 0.597$, Figure 2f) and Effort/Importance (Human $M = 5.075, SD = 0.786$, Avatar $M = 5.400, SD = 0.999; t(15) = 2.161, p = .047, d = 0.540$, Figure 2g) were significantly higher for the avatar training video compared to the human training video.

Self vs. Generic Avatars. Performance scores for the dance learned from the avatar was higher with the self-avatar ($M = 41.571, SD = 3.735$) than the generic avatar ($M = 40.375, SD = 5.502$, Figure 2a). Changes in perceived ability from baseline to post-training were greater for participants in the self-avatar condition ($M_{diff} = 1.500, SD = 2.507$) compared to the generic avatar condition ($M_{diff} = 1.00, SD = 2.777$, Figure 2b). However, these differences in performance ($t(13) = 0.485, p = .318, d = 0.251$) and perceived ability ($t(14) = 0.378, p = .356, d = 0.189$) were not statistically significant. There were no significant differences in perceived competence between the generic and self-avatar conditions ($M_{diff} = -0.021; t(14) = -0.045, p = .518, d = 0.022$, Figure 2c). Indicators of intrinsic motivation were higher in the self-avatar group, but not significantly (interest/enjoyment $M_{diff} = 0.518, t(14) = 0.856, p = .203, d = 0.428$, Figure 2f; pressure/tension $M_{diff} = 0.250, t(14) = 0.429, p = .337, d = 0.214$, Figure 2e; effort/importance $M_{diff} = 0.500, t(14) = 1.001, p = .167, d = 0.501$, Figure 2g). Immersion, indicated by the Film IEQ, was significantly greater in participants who observed a self-avatar ($M = 90.375, SD = 17.912$) compared to a generic avatar ($M = 74.000, SD = 14.794; t(14) = 1.994, p = .033, d = 0.997$, Figure 2h).

4.2 Qualitative Results

A reflexive thematic analysis was conducted to understand participant experiences and views of learning from the human dance training and avatar training videos. A data-driven inductive coding process identified a number of codes which were organised into

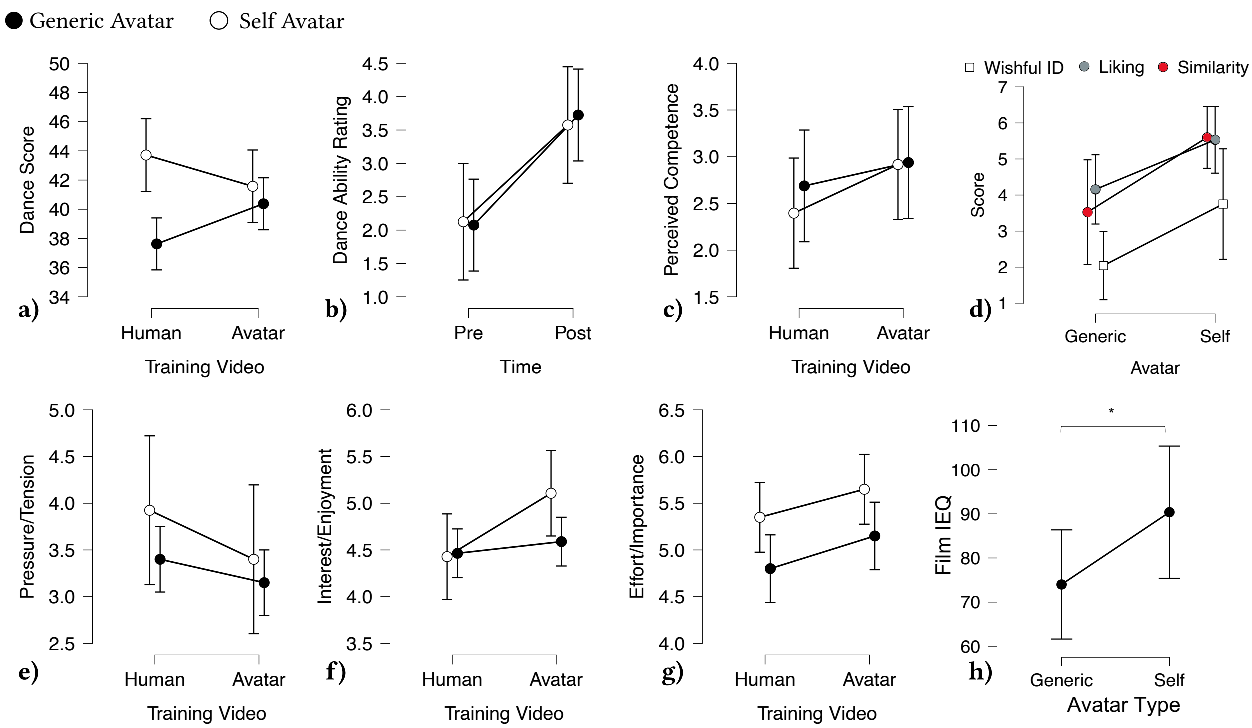

Figure 2: Effects of generic and self avatars on a) Dance performance, b) Dance Ability Rating, c) Perceived Competence, d) Avatar Identification, e) Pressure/Tension, f) Interest/Enjoyment, g) Effort/Importance and h) Film IEQ. Dots represent means and error bars show their 95% confidence intervals.

overarching themes and sub-themes [14]. Values coding was used to analyse how positive/negative participant attitudes were to the avatar training [54].

Human vs. Avatar Dance Training. **'Avatar dance training does not meet the user's expectations'**, with respect to verbal cues, movement style, and appearance. Participants noticed a lack of counting or verbal instruction from the avatar (*"lack of timing call outs")*, which they would expect during dance training (*"It was weird as there was no communication between myself and the avatar.")*. The avatar movement was perceived as being unnatural and restricted compared to the human training video (*"more 'boxy' and rigid movements so it feels less natural"; "The real person ... had more variation in the movements")*, and the avatar did not convey the Hip-Hop style as much as the human (*"the real person communicated the hip hop vibe better"; "she didn't seem as 'cool' as how I would normally perceive dancers"*. This lead to some suggestions for improving the user experience of learning from an avatar (*"(avatar) Required a voice to be more realistic"; "If the avatar could look more human-like or move more fluidly, that might improve it")*.

'Avatar dance training offers many benefits' to the user experience, being more simplified, encouraging concentration, and less nerve-racking than the human dance training. Participants felt that the avatar movement was easier to follow (*"It seemed easier to learn from the avatar"*) as it seemed more basic (*"avatar was easier to follow because it used more basic movements."*) and there was less happening (*"it felt like there was less going on so it was more easy than the other")*. The avatar dance training was more focused (*"I was more focused on the avatar than the YouTube video"*), users found they paid more attention than in the human dance training (*"It is a very good way to keep someone engaged, I was paying much more attention to my avatar than the dancer"*) and there were fewer distractions (*"it wasn't speaking so I could concentrate on the moves themselves"*). Users felt less self-conscious and nervous in the presence of an avatar compared to the human in the training video (*"I felt less self-conscious as I wasn't being watched by a real person"; "the human made everything more real, so there were more nerves."*) allowing them to be more relaxed during the training (*"I was slightly more relaxed so found it a little easier to learn, and enjoyed it more"*).

Self vs. Generic Avatars. Attitudes towards the avatar training experience were overall more positive for those in the self-avatar condition (16 statements coded as positive, 4 negative) compared to the generic avatar condition (9 statements coded as positive, 8 negative). **'Self-avatars make dance training relatable'**, the self is relatable (*"it (self-avatar) made the dance training feel more relatable"*), with customisation being a key component (*"I feel that if I had spent more time making the avatar look a lot like me, I would be more immersed in the task"*). Generic is not relatable, with generic avatars not producing the same level of connection/interest in the training experience (*"It (generic avatar) didn't connect as well with me, I don't think"*).

5 DISCUSSION

We examined the user experience of learning from a human training video compared to an avatar training video (RQ1), as well as comparing the effects of fully customised self-avatars and gender-matched generic avatars (RQ2) to investigate the application and effects of feedforward learning using a self-avatar. The results demonstrate that practically, learning from an avatar and a human in a video are both effective, however, avatars provide additional benefits. Our avatar design did not address all the user expectations, such as providing verbal dance instruction and conveying the Hip-Hop style, which could in principle be added with more design effort. Nevertheless, learning from an avatar elicited more enjoyment, concentration, and effort from the user compared to the human dance training. The avatars were perceived as providing a simplified model of movement to observe and learn the dance moves from. There was also a general consensus that learning from an avatar had a positive impact on the user experience, making users feel less self-conscious of their actions or judged by the dance instructor. Other research has also shown that people feel comfortable in the presence of avatars and indicates that this effect may be greater for self-avatars, with people feeling less embarrassed by their actions in front of a self-avatar compared with another [4].

As expected self-avatars lead to greater identification compared to generic avatars, making participants in the self-avatar group more susceptible to the effects of feedforward learning [20, 55] as the mirror neurons involved are more active in response to more familiar models [51]. We anticipated that generic avatars would result in much lower levels of identification and provide far less benefit to learning [28] than self-avatars. This effect was evident to a certain extent with participants in the self-avatar group generally demonstrating higher levels of intrinsic motivation and skill than those who learned with a generic avatar, however the difference in the feedforward effects observed between the two groups was not substantial.Despite baseline dance skill ratings and ability scores suggesting equality of groups, performance differences in the human dance training task could indicate those in the self-avatar group were more skilled, offering an alternative explanation for these differences. Furthermore, customisation of avatars is related to increased levels of effort and intrinsic motivation [10, 60] participants in the generic avatar group were assigned an avatar whilst participants in the self-avatar group personalised the avatar themselves, therefore the customisation process itself may play a role. Future research will have to show avatar effects in a within-subject design where the creation of self and other avatars involves the same extent of customisation. High impact features for identification with an avatar tend to be those which are most visible, such as skin-tone [11, 62] and gender, with gender-matched avatars leading to greater avatar self-relevance [48], opposed to finer details, e.g. facial features [24], which are harder to distinguish. In this study, participants observed either a highly similar self-avatar or a gender-matched white-skinned avatar. For Caucasian participants who made up the majority of the sample, it is possible that the generic avatar also resulted in reasonably high levels of identification and the associated benefits to learning [2].

Limitations. Not finding significant differences between the two avatar conditions is not surprising given the small sample size and hence low statistical power of this preliminary study. Additionally, it is possible that the simplicity of the task, learning three beginner Hip-Hop moves, may have been insufficient for any effects to become visible. Research has shown the effects of different avatar characteristics may only become visible when performing more complex tasks, such as Tai-Chi, opposed to those which are simple [61]. Furthermore it was difficult, particularly for a non-expert, to distinguish different levels of skill demonstrated in only three basic movements using the dance scoring rubric. Future research should seek to replicate this study with a larger sample size and more sensitive performance measures to validate the results and provide more concrete evidence of any avatar effects. Furthermore, the avatar training did not meet user expectations and so work is needed to improve the realism of the avatars to make them appear more like professional dancers. In this study, the experimenter's movement was recorded using Vive trackers to produce the animations; this motion capture technique likely impacted the user experience with participants finding the movement unnatural and lacking the Hip-Hop style compared to the human dancer. In further studies, professional dancers should be used to model the movements which are captured using professional motion capture equipment. Whilst participants learned different routines from the human and avatar training videos, the order of these conditions was not counterbalanced. Future research should mitigate order and sequence effects through counterbalancing.

Impact. This preliminary investigation into the use of self-avatars for feedforward learning is among the first to demonstrate the potential success of this type of training utilising avatars. Results of this study indicate that learning from an avatar can offer many benefits to the user experience over learning from a human in a video, and that these attentional and motivational benefits may be greater with a self-avatar than a generic avatar. It is evident that identification is a crucial factor involved in eliciting feedforward effects, with self-models producing the greatest benefits [58]. However avatar identification is seemingly strongly influenced by the most visible features [24, 60, 62], e.g. matching gender and skin-tone, to lead to some level of identification which benefits learning. Finding a balance between design effort and feedforward effects is an important consideration for future work, as it may not be necessary to produce highly detailed self-avatars which requires time and expertise [24]. Overall, avatars have great potential for feedforward learning, offering more flexibility than VSM, allowing the user to observe and learn from 'themselves' performing any skill [29] and determining the extent of avatar similarity required to promote significant learning effects whilst not requiring extensive design effort will be important in maximising this potential in the future.

ACKNOWLEDGMENTS

Isabel Fitton's research is funded by the UKRI EPSRC Centre for Doctoral Training in Digital Entertainment (CDE), EP/L016540/1 and industrial partner PwC. This work was also supported and partly funded by the Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA 2.0; EP/T022523/1) at the University of Bath.

REFERENCES

[1] Sun Joo (Grace) Ahn, Jesse Fox, and Jung Min Hahm. 2014. Using Virtual Doppelgängers to Increase Personal Relevance of Health Risk Communication. In *Intelligent Virtual Agents*, Timothy Bickmore, Stacy Marsella, and Candace Sidner (Eds.). Springer International Publishing, Cham, 1–12.

[2] Julie L. Andsager, Victoria Bemker, Hong-Lim Choi, and Vitalis Torwel. 2006. Perceived Similarity of Exemplar Traits and Behavior: Effects on Message Evaluation. *Communication Research* 33, 1 (2006), 3–18. https://doi.org/10.1177/0093650205283099 arXiv:https://doi.org/10.1177/0093650205283099

[3] Jeremy N. Bailenson, Andrew C. Beall, Jim Blascovich, Mike Raimundo, and Max Weisbuch. 2001. Intelligent Agents Who Wear Your Face: Users' Reactions to the Virtual Self. In *Intelligent Virtual Agents*, Angélica de Antonio, Ruth Aylett, and Daniel Ballin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 86–99.

[4] Jeremy N. Bailenson, Jim Blascovich, and Rosanna E. Guadagno. 2008. Self-Representations in Immersive Virtual Environments1. *Journal of Applied Social Psychology* 38, 11 (2008), 2673–2690. https://doi.org/10.1111/j.1559-1816.2008.00409.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1559-1816.2008.00409.x

[5] Albert Bandura. 1974. Behavior theory and the models of man. *American psychologist* 29, 12 (1974), 859.

[6] Albert Bandura and Richard H Walters. 1977. *Social learning theory*. Vol. 1. Englewood cliffs Prentice Hall, Oxford, England.

[7] Soumya C. Barathi, Daniel J. Finnegan, Matthew Farrow, Alexander Whaley, Pippa Heath, Jude Buckley, Peter W. Dowrick, Burkhard C. Wuensche, James L. J. Bilzon, Eamonn O'Neill, and Christof Lutteroth. 2018. Interactive Feedforward for Improving Performance and Maintaining Intrinsic Motivation in VR Exergaming. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, Montreal QC, Canada, 1–14. https://doi.org/10.1145/3173574.3173982

[8] L. Baudry, D. Leroy, and D. Chollet. 2006. The effect of combined self- and expert-modelling on the performance of the double leg circle on the pommel horse. *Journal of Sports Sciences* 24, 10 (2006), 1055–1063. https://doi.org/10.1080/02640410500432243 arXiv:https://doi.org/10.1080/02640410500432243 PMID: 17115520.

[9] Harmony Bench and Alexandra Harlig. 2021. This Is Where We Dance Now. *The International Journal of Screendance* 12 (2021), 9–12.

[10] Max V. Birk, Cheralyn Atkins, Jason T. Bowey, and Regan L. Mandryk. 2016. *Fostering Intrinsic Motivation through Avatar Identification in Digital Games*. Association for Computing Machinery, New York, NY, USA, 2982–2995. https://doi.org/10.1145/2858036.2858062

[11] Rainer Blum. 2016. *Investigation of the design of personal avatars in online shopping for effective virtual try-on applications*. Ph. D. Dissertation. Brunel University London.

[12] Bettina Bläsing, Beatriz Calvo-Merino, Emily S. Cross, Corinne Jola, Juliane Honisch, and Catherine J. Stevens. 2012. Neurocognitive control in dance perception and performance. *Acta Psychologica* 139, 2 (2012), 300–308. https://doi.org/10.1016/j.actpsy.2011.12.005

[13] Majdi Bouazizi, Driss Boudhiba, and Fairouz Azaiez. 2014. Effects of Learning by Video Modeling on Gymnastic Performances among Tunisian Students in the Second Year of Secondary Level. *IOSRJSPE* 1, 5 (2014), 05–08. https://doi.org/10.9790/6737-0150508

[14] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[15] Giovanni Buccino, Stefan Vogt, Afra Ritzl, Gereon R Fink, Karl Zilles, Hans-Joachim Freund, and Giacomo Rizzolatti. 2004. Neural Circuits Underlying Imitation Learning of Hand Actions: An Event-Related fMRI Study. *Neuron* 42, 2 (2004), 323–334. https://doi.org/10.1016/S0896-6273(04)00181-3

[16] Tom Buggey and Lindsey Ogle. 2012. Video self-modeling. *Psychology in the Schools* 49, 1 (2012), 52–70. https://doi.org/10.1002/pits.20618 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/pits.20618

[17] Jacky C.P. Chan, Howard Leung, Jeff K.T. Tang, and Taku Komura. 2011. A Virtual Reality Dance Training System Using Motion Capture Technology. *IEEE Transactions on Learning Technologies* 4, 2 (2011), 187–195. https://doi.org/10.1109/TLT.2010.27

[18] Jonathan Cohen. 2001. Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication and Society* 4, 3 (2001), 245–264. https://doi.org/10.1207/S15327825MCS0403_01

[19] Edward Downs, Nicholas D Bowman, and Jaime Banks. 2019. A polythetic model of player-avatar identification: Synthesizing multiple mechanisms. *Psychology of popular media culture.* 8, 3 (2019), 269–279.

[20] Peter W. Dowrick. 1999. A review of self modeling and related interventions. *Applied and Preventive Psychology* 8, 1 (1999), 23–39. https://doi.org/10.1016/S0962-1849(99)80009-2

[21] Peter W. Dowrick. 2012. Self modeling: Expanding the theories of learning. *Psychology in the Schools* 49, 1 (2012), 30–41. https://doi.org/10.1002/pits.20613 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/pits.20613

[22] Peter W. Dowrick, Weol Soon Kim-Rupnow, and Thomas J. Power. 2006. Video Feedforward for Reading. *The Journal of Special Education* 39, 4 (2006), 194–207. https://doi.org/10.1177/00224669060390040101 arXiv:https://doi.org/10.1177/00224669060390040101

[23] Peter W. Dowrick and John M. Raeburn. 1995. Self-modeling: Rapid skill training for children with physical disabilities. *J Dev Phys Disabil* 7, 1 (March 1995), 25–37. https://doi.org/10.1007/BF02578712

[24] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1151–1160. https://doi.org/10.1145/1518701.1518877

[25] Daniel L Eaves, Gavin Breslin, Paul Van Schaik, Emma Robinson, and Iain R Spears. 2011. The short-term effects of real-time virtual reality feedback on motor learning in dance. *Presence: Teleoperators and Virtual Environments* 20, 1 (2011), 62–77.

[26] Oliver Ehmer. 2021. Synchronization in demonstrations. Multimodal practices for instructing body knowledge. *Linguistics Vanguard* 7, s4 (2021), 20200038.

[27] Denny Fadli Fitrianto and Rizki Yantami Arumsari. 2016. Video Design As A Guide Of The Basic Movements In Breakdance In Indonesian Language. *Bandung Creative Movement (BCM)* 3, 1 (2016), 155–162.

[28] Jesse Fox, Jeremy Bailenson, and Joseph Binney. 2009. Virtual Experiences, Physical Behaviors: The Effect of Presence on Imitation of an Eating Avatar. *Presence: Teleoperators and Virtual Environments* 18, 4 (2009), 294–303. https://doi.org/10.1162/pres.18.4.294

[29] Jesse Fox, Jeremy N Bailenson, and Tony Ricciardi. 2012. Physiological Responses to Virtual Selves and Virtual others. *CyberTherapy & Rehabilitation* 5, 1 (2012), 93.

[30] Caryl H Hitchcock, Peter W Dowrick, and Mary Anne Prater. 2003. Video self-modeling intervention in school-based settings: A review. *Remedial and Special Education* 24, 1 (2003), 36–45.

[31] Mihran Kirakosian. 2016. 3 Simple Dance Moves for Beginners (Hip Hop Dance Moves Tutorial): Mihran Kirakosian. https://youtu.be/ujREEgxEP7g?t=291

[32] Martin Kocur, Melanie Kloss, Valentin Schwind, Christian Wolff, and Niels Henze. 2020. *Flexing Muscles in Virtual Reality: Effects of Avatars' Muscular Appearance on Physical Performance.* Association for Computing Machinery, New York, NY, USA, 193–205. https://doi.org/10.1145/3410404.3414261

[33] Susan R Koff. 2007. Psychomotor Domain: Dance. In *From Principles of Learning to Strategies for Instruction with Workbook Companion*. Springer, Boston, 307–321.

[34] Jordan Koulouris, Zoe Jeffery, James Best, Eamonn O'Neill, and Christof Lutteroth. 2020. *Me vs. Super(Wo)Man: Effects of Customization and Identification in a VR Exergame.* Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376661

[35] John W Krakauer, Alkis M Hadjiosif, Jing Xu, Aaron L Wong, and Adrian M Haith. 2019. Motor learning. *Compr Physiol* 9, 2 (2019), 613–663.

[36] Maïko Le Lay. 2021. Reflection on Teaching Dance on YouTube: Negotiating Between Maintaining a Culturally Relevant Pedagogy and Participating in the Commercialized Realities of Teaching Dance Online. *The International Journal of Screendance* 12 (2021), 264–268.

[37] Sohye Lim and Byron Reeves. 2009. Being in the game: Effects of avatar choice and point of view on psychophysiological responses during play. *Media psychology* 12, 4 (2009), 348–370.

[38] Edward McAuley, Terry Duncan, and Vance V. Tammen. 1989. Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport* 60, 1 (1989), 48–58. https://doi.org/10.1080/02701367.1989.10607413 PMID: 2489825.

[39] Mixamo. 2013. Adobe Fuse CC. https://www.adobe.com/wam/fuse.html

[40] Sadeghi Nahid, Nobakht Ramezani Zahra, and Attari Elham. 2013. Effects of video modeling on skill acquisition in learning the handball shoot. *European Journal of Experimental Biology* 3 (2013), 214–218.

[41] Jessica Navarro, Jorge Peña, Ausias Cebolla, and Rosa Baños. 2020. Can Avatar Appearance Influence Physical Activity? User-Avatar Similarity and Proteus Effects on Cardiac Frequency and Step Counts. *Health Communication* 37, 2 (2020), 222–229. https://doi.org/10.1080/10410236.2020.1834194 PMID: 33054371.

[42] Jorge Peña, Subuhi Khan, and Cassandra Alexopoulos. 2016. I am what I see: How avatar and opponent agent body size affects physical activity among men playing exergames. *Journal of Computer-Mediated Communication* 21, 3 (2016), 195–209.

[43] Jorge Peña and Eunice Kim. 2014. Increasing exergame physical activity through self and opponent avatar appearance. *Computers in Human Behavior* 41 (2014), 262–267. https://doi.org/10.1016/j.chb.2014.09.038

[44] Mary Anne Prater, Nari Carter, Caryl Hitchcock, and Peter Dowrick. 2012. Video self-modeling to improve academic performance: A literature review. *Psychology in the Schools* 49, 1 (2012), 71–81.

[45] Rimasari Pramesti Putri. 2021. Tiktok as an Online Learning Media During a Pandemic. In *Proceedings of the 6th International Conference on Education & Social Sciences (ICESS 2021)*. Atlantis Press, Paris, 282–287. https://doi.org/10.2991/assehr.k.210918.052

[46] Katherine M. Rahill and Marc M. Sebrechts. 2021. Effects of Avatar player-similarity and player-construction on gaming performance. *Computers in Human Behavior Reports* 4 (2021), 100131. https://doi.org/10.1016/j.chbr.2021.100131

[47] Rabindra Ratan and Béatrice S Hasler. 2010. Exploring self-presence in collaborative virtual teams. *PsychNology Journal* 8 (2010), 11–31.

[48] Rabindra A Ratan and Michael Dawson. 2016. When Mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research* 43, 8 (2016), 1065–1093.

[49] RCampus. 2014. iRubric: DANCE PERFORMANCE EVALUATION rubric - E8X3A9: RCampus. https://www.rcampus.com/rubricshowc.cfm?code=E8X3A9&sp=yes&

[50] Jacob M. Rigby, Duncan P Brumby, Sandy J. J. Gould, and Anna L Cox. 2019. Development of a Questionnaire to Measure Immersion in Video Media: The Film IEQ. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video* (Salford (Manchester), United Kingdom) *(TVX '19)*. Association for Computing Machinery, New York, NY, USA, 35–46. https://doi.org/10.1145/3317697.3323361

[51] Giacomo Rizzolatti and Corrado Sinigaglia. 2010. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci* 11, 4 (April 2010), 264–274. https://doi.org/10.1038/nrn2805

[52] Cathy Robson, Neville Blampied, and Lawrence Walker. 2015. Effects of feedforward video self-modelling on reading fluency and comprehension. *Behaviour Change* 32, 1 (2015), 46–58.

[53] Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (2006), 344–360. https://doi.org/10.1007/s11031-006-9051-8

[54] Johnny Saldaña. 2021. *The coding manual for qualitative researchers*. sage, London.

[55] Dale H. Schunk and Antoinette R. Hanson. 1989. Self-modeling and children's cognitive skill learning. *Journal of Educational Psychology* 81, 2 (1989), 155–163. https://doi.org/10.1037/0022-0663.81.2.155 Place: US Publisher: American Psychological Association.

[56] U.S. Figure Skating 2004. *Scoring System | U.S. Figure Skating*. U.S. Figure Skating. https://www.usfigureskating.org/about/scoring-system

[57] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review* 20, 1 (2013), 21–53.

[58] Joanna Starek and Penny McCullagh. 1999. The effect of self-modeling on the performance of beginning swimmers. *The Sport Psychologist* 13, 3 (1999), 269–287.

[59] Kylie A Steel, Roger D Adams, Susan E Coulson, Colleen G Canning, and Holly J Hawtin. 2013. Video Self-model Training of Punt Kicking. *Int. J. Sport Health Sci.* 11, 0 (2013), 49–53. https://doi.org/10.5432/ijshs.201233

[60] Selen Turkay and Charles K Kinzer. 2015. The effects of avatar-based customization on player identification. In *Gamification: Concepts, methodologies, tools, and applications*. IGI Global, USA, 247–272.

[61] JM Monica Van De Ridder, Karel M Stokking, William C McGaghie, and Olle Th J Ten Cate. 2008. What is feedback in clinical education? *Medical education* 42, 2 (2008), 189–197.

[62] Asimina Vasalou, Adam Joinson, Tanja Bänziger, Peter Goldie, and Jeremy Pitt. 2008. Avatars in social media: Balancing accuracy, playfulness and embodied messages. *International Journal of Human-Computer Studies* 66, 11 (2008), 801–811.

[63] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240.

[64] Nick Yee and Jeremy Bailenson. 2007. The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research* 33, 3 (2007), 271–290.