FakeForward: Using Deepfake Technology for Feedforward Learning

Christopher Clarke
University of Bath
Bath, United Kingdom

Jingnan Xu
University of Bath
Bath, United Kingdom

Ye Zhu
University of Bath
Bath, United Kingdom

Karan Dharamshi
University of Bath
Bath, United Kingdom

Harry McGill
University of Bath
Bath, United Kingdom

Stephen Black
University of Bath
Bath, United Kingdom

Christof Lutteroth
University of Bath
Bath, United Kingdom

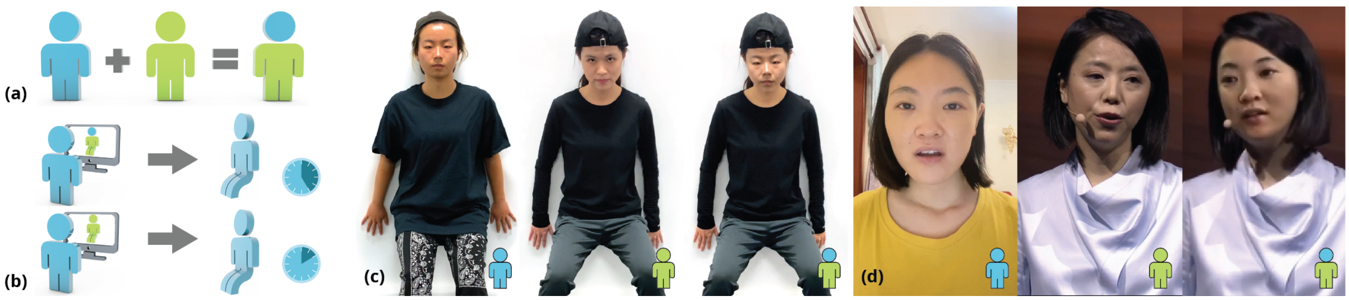

Figure 1: FakeForward combines deepfake technology with the theories of feedforward and video self-modelling to create effective, personalised training tools. A user's face (blue) is swapped with someone who is better at a skill than they are (green) creating a FakeForward version of themselves performing a skill at a higher level than they would usually (a). After watching FakeForward videos, users perform better compared to watching the original video for many physical exercises (b) including wall sits (c) and other strength endurance and cardio tasks. For public speaking, users are more confident and less anxious after having watched a FakeForward video (d).

ABSTRACT

Videos are commonly used to support learning of new skills, to improve existing skills, and as a source of motivation for training. Video self-modelling (VSM) is a learning technique that improves performance and motivation by showing a user a video of themselves performing a skill at a level they have not yet achieved. Traditional VSM is very data and labour intensive: a lot of video footage needs to be collected and manually edited in order to create an effective self-modelling video. We address this by presenting FakeForward – a method which uses deepfakes to create self-modelling videos from videos of other people. FakeForward turns videos of better-performing people into effective, personalised training tools by replacing their face with the user's. We investigate how FakeForward can be effectively applied and demonstrate its efficacy in rapidly improving performance in physical exercises, as well as confidence and perceived competence in public speaking.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

Feedforward, Deepfake, Videos, Training, Skill Acquisition, Fitness, Physical Exercise, Public Speaking

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581100

ACM Reference Format:
Christopher Clarke, Jingnan Xu, Ye Zhu, Karan Dharamshi, Harry McGill, Stephen Black, and Christof Lutteroth. 2023. FakeForward: Using Deepfake Technology for Feedforward Learning. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3544548.3581100

1 INTRODUCTION

For decades, videos have been used as a tool for learning, training, and motivation [15, 23]. The low barrier to entry of video creation results in a wide range of instructors offering guidance on a plethora of skills which are easily accessible on modern streaming platforms [50, 82, 101, 136]. Besides helping people to acquire new skills, video streaming sites also provide a low-cost platform to help people stay fit and healthy through physical exercise training videos. This has been especially important during the COVID-19 pandemic, which has exacerbated an already existing global issue of physical inactivity and sedentary behaviour [62, 108].

Videos are a particularly effective tool for learning and training because they facilitate *modelling* [118], i.e. learners can watch a model performer execute the desired skill. Video modelling is a form of *feedforward learning*, a psychological technique where people learn from a demonstration of a skill or behaviour at a level they have not yet acquired [44–46]. Video modelling does not only support the acquisition of a skill [92, 129, 146], it can also increase a user's engagement in a task by enhancing their motivation and self-efficacy [123, 131, 143, 144]. However, the physical similarity between the model performer and the observer is a crucial aspect for the efficacy of video modelling – the more alike the model is to the observer, the more likely the observer is to learn effectively from the video [123, 125]. Video self-modelling (VSM) is a special case of video modelling that avoids this disparity: the observer learns and watches themselves as the model performer [20, 44]. VSM has been successfully applied across a range of users in a number of applications including physical skills (e.g. cycling [74], swimming [29, 131], football [100, 133]), academic and vocational issues (e.g. language and cognitive skills [67, 114, 119, 125]), and social adjustment (e.g. children with behavioural issues [43, 44, 95, 117] and suppressing inappropriate sexual behaviour [49]).

VSM relies on footage showing how the learner engages in behaviour that they may not typically achieve in real life. This has traditionally been achieved by collecting footage of a learner and then manually editing the footage to extract pieces of desired skills from the user's repertoire and placing them in different contexts or in different orders to create the illusion of a better performance [44]. This does not allow for complex behaviours to be self-modelled that are not already in the user's repertoire, and the technological and time complexity of editing the videos are a current limitation in terms of wide-spread adoption of VSM [20, 31]. However, advances in artificial intelligence have led to the development of automated video editing techniques, including "deepfake" algorithms, which replace one person's likeness with another's using deep learning techniques [71, 152]. To the best of our knowledge, deepfakes have never been utilised for feedforward learning, which raises the question whether deepfake videos can be used to address the challenge of creating effective self-modelling videos.

Photo manipulation and face swapping date back to the 19th century [61]. Traditionally the process of manipulating faces in an image was laborious and required highly skilled digital artists. By contrast, deepfake algorithms automate the process and are advanced enough that validating their authenticity is a major concern and active research topic [65, 90], raising a number of ethical issues around consent, deception and misattribution [38]. Despite this, deepfake technology is widely available both as commercial [55] and as open-source software [53]. Due to its availability and ease of use anyone can create convincing deepfakes, which has led to highly controversial creations and negative use cases including blackmail, revenge pornography, and disinformation [1, 64, 140]. Positive applications of deepfake technology are relatively scarce and in some cases speculative [56]. In this paper, we investigate how deepfake technology can be embraced by individuals to create effective, personalised training and development tools.

We present *FakeForward*, a novel method for VSM using deepfake technology (Figure 1). FakeForward leverages deepfake technology to superimpose the user's face onto a video of someone who has greater mastery at a skill or activity. Not only does this reduce the complexity of creating the VSM video by removing any requirements for manual editing; it also enables the user to see themselves execute a skill or activity that they would not yet have achieved or mastered. This raises a number of interesting research questions surrounding the use of deepfakes for VSM:

- **RQ1** How can deepfake technology be used for VSM?
- **RQ2** What areas of training and development can be supported with FakeForward?
- **RQ3** How effective is FakeForward in helping users to develop skills?

To address these research questions, we developed guidelines for creating FakeForward videos (Section 3), and validated our approach in two user studies. Study 1 (Section 4) investigates what type of physical skills FakeForward may be applicable to including both strength and cardiovascular skills. Due to the wide variety of strength exercises available, we chose different exercises that cover different muscles groups (upper body, lower body, and core) [17] and type of motor function (static and dynamic) [60]. Study 2 (Section 5) focuses on psychological skills, assessing FakeForward's effectiveness for public speaking. In summary, we make the following novel contributions:

(1) FakeForward – a novel method of using deepfake technology for effective video self-modelling.
(2) Technical and ethical protocols for developing FakeForward videos, based on our experiences throughout this project.
(3) Two studies demonstrating the efficacy of FakeForward which showcase how it can be a powerful support tool when developing and practising a variety of physical and psychological skills.

2 BACKGROUND AND RELATED WORK

Prior research has explored how interactive technologies can support learning and skill acquisition. A common approach is to sense the user's movements using camera-based technologies [99] or wearable devices [86] and provide visual real-time feedback overlaid onto a video [137], augmented mirrors [2], in virtual reality [150], or through tactile sensors embedded in a robotic wearable system [86]. However, these systems depend on curated media and are often aimed at learning a specific skill, for example ballet [139] or weightlifting [80]. Video games are another popular medium which are effective because they make the learning process more enjoyable through gamification [41, 105, 113]. Compared to these approaches, FakeForward is more accessible as it does not rely on real-time processing and removes the need for complex bespoke content authoring such as gamification. Instead, we present a protocol which provides the opportunity for users to generate FakeForward videos on a wide variety of skills using open-source software.

Videos are a compelling medium for learning a new skill, as they require minimal hardware and video content is easily accessible from a wide variety of sources. In addition to modelling, one of the main ways in which videos are used to enhance motor learning is the use of *augmented feedback* to understand "what was done" [118, 130]. By recording the user, feedback can be used to analyse how well they completed a task [148]. Proposed improvements to using videos for developing skills include automatic editing to remove repetitive actions or mistakes from amateur instructional videos [27], enhanced video interfaces based on data-driven insights [77], video playback which adapts to the user's input [30], and inputs such as voice [25] or gestures [66] allowing the user to focus on the task at hand. FakeForward has the potential to complement these approaches, leveraging deepfake processing to generate feedforward videos that can better support skills development.

In contrast to feedback, which uses current information about a system and its environment, feedforward control uses predictions about the future state of a system and takes corrective action before the controlled variable deviates [128]. Feedforward control has been applied extensively in the fields of engineering (e.g. [128]) and cognitive psychology (see Basso and Belardinelli [10] for an extensive overview), playing an important role in psychological models [40] and descriptions of the visual and motor systems [35, 81]. Building upon this core principle, feedforward has successfully been used as a process of learning by constructing an image of success that illustrates a desired behaviour an individual has not yet achieved [45]. For feedforward learning to be successful, the depicted behaviour must not be beyond the individual's capability, and many works support the theory that rapid learning occurs because the individual already possesses component skills to achieve the desired outcome [44, 45].

Feedforward, in the context of learning, has foundations in Bandura's work on social learning which suggests the majority of human behaviour and skill development occurs through observing models [4, 5]. Prior work suggests the feedforward effect is continuous, and the degree to which a person identifies with the model is an important contributor to the magnitude of the effect [45, 46]. Evidence suggests this may be a result of superior mirror neuron response, which fire when performing and observing an action [45]. Self-modelling maximises model similarity, and involves an individual observing themselves performing a task or behaviour successfully. Peer modelling (i.e. observing someone else) shows improved results when the models have similar competence to the observer [123, 124, 126, 144], and research into virtual reality (VR) exergaming has shown self-identification is also possible with a digital avatar [79]. In addition to physical similarity, the behavioural similarity of the model is important [46, 123]. Prior work has shown how feedforward increases performance when self-identification takes on a purely behavioural form, using past and future performances in the form of video game "ghosts" [8, 98].

Video self-modelling (VSM) combines feedforward and self modelling using video as the medium to develop and teach skills [20, 44, 45, 47, 96]. An individual watches a video of themselves (or someone resembling them) engaged in a task, and they replicate the behaviours and successful actions performed in the video [44]. Viewing oneself – or just identifying with a model – stimulates the individual, commands attention, and increases motivation, and self-efficacy, which in turn allows for more focused learning [3, 9, 18, 44, 69, 96]. Many studies have successfully evaluated the effectiveness of feedforward and VSM for skills development. There have been positive applications of VSM for physical activities and sports, often as a supplement to training, including swimming [29, 131], gymnastics [132, 142], football [100, 133], and power-lifting [57]. Beyond physical activities, VSM has successfully been applied to assisting with communication [14, 75], academic and vocational issues [67, 114, 119, 125], and motor and functional skills [94]. VSM has also been prominent in teaching behaviours and skills to children with physical disabilities [48] and people with developmental and behavioural problems [18, 44, 49, 76, 117, 151] including extensive research on its application to those with autism spectrum disorder [13, 19, 21, 22, 26, 32, 63, 145].

Prior work has explored how someone's perception of their ability to perform a skill or activity may lead to the performance benefits observed with feedforward interventions. The principle of self-efficacy – an individual's belief in their capacity to perform a task, which is directly linked to the task outcome – was derived from social cognitive theory [3] and has been hypothesised to play an important role, with success raising self-belief and failure lowering it [5]. However, VSM applied in the context of physical activities and sports resulted in no significant differences when measuring self-efficacy, despite improving performance in all cases [29, 131, 132]. Similar to self-efficacy, perceived competence considers the extent to which an individual believes they can learn and execute a given task. Conceived from self-determination theory [122], perceived competence can enhance intrinsic motivation according to cognitive evaluation theory [121]. Feedforward interventions have been shown to increase perceived competence and intrinsic motivation in VR exergaming applications [54, 70, 98]. However, contrasting results have been found in VSM: Clark and Ste-Marie [29] found VSM increased intrinsic motivation, whereas Ste-Marie et al. [132] found no significant difference. Therefore, we investigate how FakeForward affects the participants' self-efficacy, perceived competence and intrinsic motivation, in addition to their performance.

VSM videos are edited to show the individual performing at a more advanced level than they would usually achieve. This is traditionally achieved by editing videos of an individual by removing error or non-desirable behaviour from the task [29], showing specific behaviour in other environments [44], or by recombining existing component skills [132, 142]. It is also important that the videos are long enough to keep the individual engaged for observational learning to occur and for the representation of the desired behaviour in memory [5, 31, 68]. Existing methods for creating VSM videos are limited as the individual may not be able to correctly perform the desired task, or it could take a long time to collect the required footage. This can result in many videos to manually edit, which takes a long time. FakeForward overcomes these limitations as the initial model in the video is not the user themselves. Instead, deepfake technology superimposes the user's likeness onto someone else who may be more capable at performing the skill at a higher level, creating a unique opportunity to show behaviour that is beyond the user's current capabilities but still achievable.

Deepfake research predominantly falls into three categories: developing the underlying algorithms [24, 135, 138], developing approaches to detect fake media as it becomes increasingly more difficult to distinguish what is real and what is fake [84, 85, 90], and exploring the ethical issues due to the malicious nature of many deepfake videos created and lack of regulations [97, 107]. Deepfakes have become infamous for nefarious purposes such as blackmail, revenge pornography, and disinformation [1, 64, 140]. They also have the potential to replace computer-generated imagery (CGI) in the film industry, which has been contentiously used to bring actors back to life, again raising the question of how individuals can be protected from abuse [149]. However, deepfakes also have the potential for positive uses. For example, this was recently demonstrated by a campaign in which deepfakes were used to break down language barriers by creating multilingual videos of David Beckham announcing the "*Malaria Must Die*" petition [39].

3 FAKEFORWARD

In this section we discuss both ethical and technical protocols for the development of FakeForward videos (RQ1). We refer to FakeForward as a general concept involving the use of deepfake-based software to create video self-models. In this paper, we focus on face swapping using deepfake techniques, for which open source software is openly available for tech savvy users. Additionally, commercial applications provide easy-to-use interfaces which simplify the process, making the technology accessible for mainstream use.

3.1 Ethical Protocol

Due to the widespread misuse of deepfake technology, the ethical implications of deploying the FakeForward concept should be carefully considered.

Be transparent about potential benefits. Clear instructions about the potential benefits should be communicated to users, noting that not everyone may experience the same level of improvements, or none at all, by watching their FakeForward videos.

Verify that users are creating self-models. The underpinning concept of FakeForward is to create self-models of a user for private consumption. Therefore, there is no need for the generated FakeForward videos to be publicly available, and verification processes should be in place to ensure that users are indeed creating self-model videos rather than videos of others. This is not simply a verification that the user is a unique individual and legitimate user, it also entails validation that the source material used to create FakeForward videos is of the user themselves. This protects individuals from those who may wish to create FakeForward videos of others in order to cause distress, embarrassment, or to undermine someone's self confidence rather than building it up.

Protect individual's personal data. The source material required to generate FakeForward videos is classed as personal data, requiring the appropriate levels of security when storing to ensure others cannot access and use an individual's appearance inappropriately (this is often also a legal requirement, e.g., in the European Union's General Data Protection Regulation). Watermarks or other suitable identifiers (e.g., relevant meta data) should be included in FakeForward videos to provide traceability and foster accountability.

Moderate peer model content to prevent harm. The content of skills and activities available for users to generate FakeForward videos needs to be carefully moderated to protect users from themselves. In this paper, FakeForward utilises positive peer models to elicit increases in performance and confidence. However, work on peer learning has shown that negative peer models can encourage negative behaviour, e.g. in the contexts of childhood education [7], harassment and gender violence [104], and self injury [115]. Some individuals may seek out content to engage in negative behaviours such as suicide and self-harm, including the consumption of video material that depicts these acts [51]. When sourcing video material for use with FakeForward, only skills and activities which contribute to an individual's positive development and which do not cause harm are appropriate.

Ensure peer models are good quality. Peer model videos should be checked for quality to maximise potential performance benefits from FakeForward. This includes both technical aspects of the video (detailed in the technical protocol), and also the correct form of the skills presented.

Respect peer models. When collecting original peer model videos, voluntary informed consent should be sought from peer models for their videos to be used for FakeForward. Similarly, when using existing video resources of peer models the content should be used in a manner that respects the original purpose for which the material was intended and copyright should be respected.

Foster inclusion and fair participation of all people. Gaining a wide variety of peer models for a given activity helps to foster a sense of inclusion and fair participation by a wide range of individuals. Having peer models that look similar to the user not only helps foster a sense of inclusion, but also improves the quality of the generated self-models. Gaining a wide variety of peer models that represent realistic and attainable performance improvements could be problematic depending on external video resources available. One way to overcome this is to collect original video material, making sure to include a diverse and inclusive sample of peer models. FakeForward is based on observing realistic and attainable performance improvements, therefore it is important to also cover a range of different abilities.

3.2 Technical Protocol

At the beginning of this project, we defined a technical protocol for the application of FakeForward as a support tool for training and development. This protocol was inspired by existing literature and developed in earlier stages of the research during pilot experiments (prior to the studies reported) where we tried to maximise the quality of the deepfakes using different versions of recording and processing the video. While the procedure was overall consistent throughout the main studies, some of its details were refined as we gained more experience. It specifies how baseline videos of a participant performing a task should be recorded in order to yield adequate FakeForward videos when using that participant as a model. The protocol addresses the following requirements relating to the task, the participant, the way the participant should perform the task, and the way the participant should be recorded.

Participants should look as similar as possible. According to feedforward theory, the more similar a model is to a participant, the greater the feedforward effect will be [44, 45]. At first we do not know who will be the best model for which participant, therefore it is advisable to reduce the variety in visual appearance of participants (apart from the face) in order to facilitate a good model fit. To achieve this, we asked participants to wear uniform clothes (a black T-shirt, jeans and light coloured trainers), put their hair to the back of their head with a hair band in case of long hair, and wear a black baseball cap front-to-back to cover their hair. We found this effectively reduces the visual features that need to be matched during model selection, making it much easier to find an adequate model for each participant.

Participants should be clearly shown how a task is performed. Although this may sound obvious, it is surprising how easy it is for participants to deviate from the way we want a task to be performed. A good way to show each participant how to perform a task, in a manner consistent across participants, was to show them the same video of the task. Then, *participants should practice the task*, demonstrating the task to the experimenter in order to ascertain that they have understood it and giving the experimenter the opportunity to correct them. In the case of strength or endurance related tasks, task practice should focus on correct technique and incorporate rests, in order to not exhaust the participant.

Tasks should be feasible and adjustable to the participants' ability. Sometimes it is hard for every participant to do a task in exactly the same manner. For example, for some participants the threshold of some tasks such as press-ups on the toes is too high. In that case, well-defined task variations should be offered, such as press-ups on the knees, as feedforward is only known to work if a task is generally within the capabilities of a person [44, 45]. Similarly, participants who cannot perform a task at all or are suffering from health conditions affected by the task, e.g. some physical injury, should be screened out early on.

Tasks should be performed in a slow, steady and controlled manner where possible. If a task such as press-ups is performed quickly, this makes it harder for participants to see what is happening in the video. Furthermore, if participants perform a task at drastically different speeds, this makes it harder to select a believable model for a participant. Finally, sharp movements in the video cause blur and lead to problems in the deepfake processing, reducing the visual quality of the FakeForward video. Slow but steady movements mitigate these problems, so we made our participants aware of this.

Tasks should be performed at least twice. When recording the peer model videos, we found it useful for participants to perform and record each task at least twice because for some tasks there is an element of chance and a single performance may not have been representative of a participant's ability. Furthermore, having more than one peer model recording increases the variety in the model footage available and provides redundancy against problems with the video (e.g. glitches, mistakes and problems with visual quality).

Participants' faces should stay visible and face the camera as much as possible. Deepfake face swapping only works well if a participant's face is clearly visible in the recording, both when building a model of a participant's face and when replacing the face with another. Furthermore, feedforward works best when users clearly see themselves in the video [44, 45]. Therefore, we chose camera angles for each task that had a good view of the face and asked participants to face the camera as much as possible. For example, when doing press-ups participants were asked to look forward towards the camera rather than down onto the floor.

All videos should have the same or similar visual background. Seeing a background in their FakeForward videos that is similar to the environment a participant initially performed their tasks in supports the illusion that the FakeForward video is really theirs. If there are obvious differences in the background then this is generally noticed by participants and can break the illusion.

All videos should be recorded under similar, well-lit light conditions. We recommend to use only artificial light in order to create a lighting environment that is consistent across all videos. Ideally, the environment should be lit in a way that keeps facial features clearly visible and reduces shadows on the face. Variations in lighting conditions and shadows can reduce realism in the generated deepfake videos as it makes it harder for the deepfake processing pipeline to match the lighting of the replaced and replacing faces.

4 STUDY 1: STRENGTH & CARDIO

Our first study was designed to address all three research questions in order to provide a first overview of the FakeForward method. First, we defined, refined and tested a protocol for the application of FakeForward for skills development (RQ1). To determine which types of skills FakeForward is effective (RQ2), we first considered strength endurance and cardio exercises, based on prior work in VSM demonstrating potential performance benefits for physical skills. We chose six exercises based on muscle group (upper body, lower body, and core) [17] and type of motor function (static and dynamic) [60]. In particular, we investigate both static and dynamic lower body and core exercises, and a dynamic upper body exercise. We do not include a static upper body exercise due to their rarity and instead investigate FakeForward's effect on cardiovascular exercise. To test the efficacy of FakeForward (RQ3), we compared task performance after watching a FakeForward video vs. after watching the corresponding peer model video.

4.1 Methodology

We used a within-participant experimental design in which participants took part in three sessions. In the first session, a participant's baseline (B) performance in all tasks was measured and recorded on video. At the end of the session they selected a peer model based on images of individuals who outperformed them in the baseline session. Then, the face of the model in the peer model video was replaced with the face of the participant to form the participant's FakeForward video for the task. The second and third sessions involved two training conditions which were counterbalanced. In the peer model video (P) condition, participants were shown the original video of the peer model they had selected and were then asked to perform the task as well as possible. In the FakeForward (F) condition, participants were shown the FakeForward video and then asked to perform the task as well as possible. The only difference between the conditions was the face shown in the video (the face of the original peer model vs. the participant's deepfaked face).

4.1.1 *Peer Models.*

Eight participants (3 female, 5 male) between the ages of 21 and 45 were recruited to act as peer models and were recruited through word of mouth. Participants were invited to the research laboratory where they were recorded performing the six physical exercises to the best of their abilities. We specifically recruited a diverse range of people to maximise the likelihood that there would always be a suitable peer model for the participants in the study. None of the peer models took part in the study.

4.1.2 Tasks. The first task was wall sits, i.e. sitting in the air with the back against a wall and the legs shoulder-width apart, keeping the knees bent at 90° (Figure 2a). This was chosen as a typical static strength endurance exercise for the lower body. The second task was body-weight squats, i.e. bending the knees 90° with the feet shoulder-width apart and the chest facing forward, and then coming up again (Figure 2b). This was chosen as a typical dynamic strength endurance exercise for the lower body. The third task was a plank, supporting the body on the elbows and the feet while holding the torso in a straight line (Figure 2c). This was chosen as a typical static core strength endurance task, as it engages the whole core musculature in keeping the body straight. The fourth task was press-ups, with the hands placed slightly wider than the shoulders and the back kept straight (Figure 2d). This was chosen as a typical dynamic upper body strength endurance task where the chest is lowered until the elbow joints are bent at right angles and then raised again until the elbow joints are straight. Participants performed these either on their toes or on their knees, according to their level of fitness and preference. The fifth task was sit-ups, i.e. lying with the back on the floor with knees bent, placing hands across shoulders and bringing torso to a vertical position (Figure 2e). This was chosen as a typical dynamic strength endurance exercise for the core. The sixth task was squat jumps, i.e. jumping in the air from a squat position (Figure 2f). This was chosen as a typical cardiovascular endurance exercise, requiring power from the lower body and core.

4.1.3 Outcome Measures. Performance was measured as either duration or number of correct repetitions. For the squats, press-ups, sit-ups and squat jumps tasks, we measured performance as the number of full repetitions a participant was able to perform without resting in between the repetitions. For the wall-sit and plank tasks, we measured performance as the time a participant was able to hold the required position.

Motivation to perform the task was measured using the Intrinsic Motivation Inventory (IMI) [120], which is a well-used and validated scale measuring intrinsic motivation in the context of an activity [28, 93]. The IMI is comprised of several subscales, which can be selected to fit a task; we included the Interest/Enjoyment subscale, which is the main measure of intrinsic motivation, and the Perceived Competence subscale, which measures how competent a participant feels they are at performing a task. All subscales are Likert scales rated from 1 (lowest) to 7 (highest).

We measured self-efficacy using an Exercise Self Efficacy (ESE) questionnaire [59] which has been shown to be a valid and reliable measure for use with a diverse population [147]. Self-efficacy is a key construct in Bandura's social cognitive theory, and it has been hypothesised that the feedforward effect arises in part due to increased self-efficacy [44]. All items in the ESE are 7-point Likert items ranging from "I cannot do it at all" to "I am certain that I can do it". We used the IMI and ESE scores to understand if any underlying cognitive processes led to changes in performance between the experimental conditions.

The video identification questionnaire (VIQ) used to measure identification, i.e. how much participants identify with the person shown in a video, was based on a variation of Van Looy et al.'s well-used and validated player identification questionnaire [141]. While the questionnaire was originally designed to measure identification with an avatar in an online role-playing game, the items measuring similarity and wishful identification can be straightforwardly applied to a video with only minor changes by referring to the video instead of an online character. We used the resulting identification score to determine whether participants are able to identify with their FakeForward video.

Flow is a psychological state associated with high performance and enjoyment which occurs when there is a balance between perceived challenge and skill in an activity. Increased flow has been shown to be an important factor in an athlete's peak performance [72] and has been linked to sports confidence, a similar principle to self-efficacy [112]. We therefore investigated whether flow may have contributed to the performance of participants. We measured flow using the Flow Scale Questionnaire (FSQ) of the Positive Psychology Lab [89], which has been used and validated for various activities. The FSQ has two subscales: Absorption, which measures how much a participant is mentally absorbed in an activity, and Balance, which measures how well the skills of a participant are balanced with the challenge they are given. The subscales are Likert scales rated from 1 (lowest) to 5 (highest).

4.1.4 Procedure. All three sessions were conducted in the same research laboratory in which the peer models were recorded. The first session began by obtaining informed consent and screening participants by ensuring they were able to perform the six tasks without inherent problems or health concerns. Participants completed the Physical Activity Readiness Questionnaire (PAR-Q), developed by the British Columbia Ministry of Health and the Multidisciplinary Board on Exercise, and we excluded participants who were unable to perform the exercises because of an injury. Participants were not informed about our hypotheses. We followed the FakeForward protocol from Section 3, e.g. we asked participants to wear uniform clothing and had some common items of clothing such as a black T-shirt and a black baseball cap ready for them to use. After completing a demographics questionnaire, participants were guided one-by-one through the six tasks (wall-sits, squats, plank, press-ups, sit-ups, and squat jumps) in the baseline condition (B). First, we showed them a brief tutorial video illustrating the task, sourced from YouTube. They then answered the VIQ, IMI, and ESE questionnaires to better understand immediate psychological effects of the videos. Then we gave them a couple of minutes time to practice the task's technique, until they felt they were ready. We guided and corrected their technique where necessary. Then they were asked to perform the exercise to the best of their ability, i.e. as long as possible (for wall-sits and plank) or with the most repetitions (for squats, press-ups, sit-ups, and squat jumps), while using the correct technique. For each task, participants were video-recorded and the experimenter recorded the performance in real-time in order to present the list of better performing peer models at the end of the first session. After performing a task, participants completed the

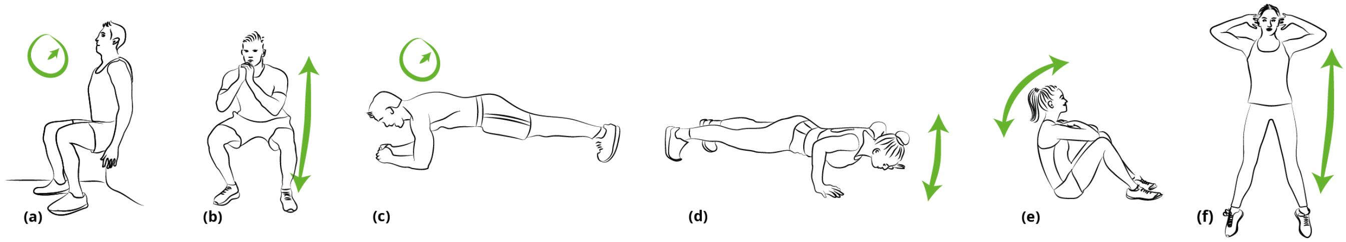

Figure 2: Study One investigated the efficacy of FakeForward for six different exercises including (a) wall-sits, (b) squats, (c) plank, (d) press-ups, (e) sit-ups, and (f) squat jumps.

post-task questionnaires which included the IMI, ESE, and FSQ scales. The first session took on average about an hour.

After participants completed the first session they selected a peer model that they felt was most visually similar in appearance to themselves. Participants were presented with images of the peer models who outperformed them in the baseline session based on the real-time performance measure recorded by the experimenter (i.e., number of repetitions or duration of task). All participants were presented with at least one peer model of the same identified gender and similar ethnicity. After selection, the face of the model in the model's video was replaced by the face of the participant to form the participant's FakeForward video for the task.

We used the DeepFaceLab [37, 111] deepfake software to replace the face in each model video with the face of the corresponding participant, using high quality settings on a PC with a fast GPU (NVIDIA RTX GeForce 2080, 3070, or 3090). The recorded videos were first converted into images, prior to the face extraction process which took approximately two hours, followed by training for 500,000 iterations which took approximately six hours. Finally, we performed merging of the deepfakes which sometimes involved manual adjustments (e.g., erode mask, blur mask, motion blur, sharpen tool, colour and face scale).

About a week later participants attended the second session, which was was one of two experimental conditions: in the peer model video condition (P) participants were shown the original videos of the peer models they selected, and in the FakeForward condition (F) they were shown their FakeForward videos (i.e., their face on the peer model). The order of the two conditions was counterbalanced across participants to mitigate order effects such as training and fatigue. In each condition, participants were guided through the six tasks as they were in the baseline condition, and told to perform each task as best as they could, i.e., the longest (for wall-sits and plank) or with the most repetitions (for squats, press-ups, sit-ups, and squat jumps) while using the correct technique. After watching a video for a task, participants completed post-video questionnaires (VIQ, IMI, and ESE) before performing the physical exercise. During the tasks, participants were reminded and corrected by the experimenter to maintain the same form as shown in the original instruction video (e.g., hands on the wall for wall-sits), to remove the possibility that form which deviated from the instructional video would allow participants to achieve better results (e.g., hands resting on the legs for the wall-sits). After the exercise, they answered post-task questionnaires (IMI, ESE, and FSQ) using the same procedure as the first session. Between exercises, participants were given several minutes rest to mitigate fatigue.

The third session was exactly the same as the second session but with the opposite experimental condition. The third session was carried out at least one day after the second session to allow at least one day's recovery from carrying out the exercises, and to minimise the impact that the second session had on the performance in the third session. After completing the third session, participants were interviewed about their experience using four open-ended questions ("*Did you prefer the original peer model videos or the deepfake videos?*", "*Do you think one of the videos, the original peer model video or the deepfake video, helped you to perform better?*", "*Are there any specific exercises where you think your performance was affected by the video?*", "*In general, how do you feel about the deepfake videos you saw?*"). Both the second and third sessions took about one hour to complete each.

4.1.5 Hypotheses. We posed the following *a-priori* hypotheses based on predictions from feedforward theory:

- **H1.1** Participants hold the wall sit for longer after watching the FakeForward video compared to the peer model video.
- **H1.2** Participants do more squats after watching the FakeForward video compared to the peer model video.
- **H1.3** Participants hold the plank for longer after watching the FakeForward video compared to the peer model video.
- **H1.4** Participants do more press-ups after watching the FakeForward video compared to the peer model video.
- **H1.5** Participants do more sit-ups after watching the FakeForward video compared to the peer model video.
- **H1.6** Participants do more squat jumps after watching the FakeForward video compared to the peer model video.

4.1.6 Participants. A sample of 24 participants (15 male, 9 female), aged 19–34 (M = 24.38, SD = 3.08) were recruited through word of mouth, mailing lists, and social media. An additional five participants were recruited to take part, however four withdrew from the study and one participant outperformed all peer models in the baseline condition and therefore we excluded them. The participants for the peer models and for the study were independent of each other (i.e., none of the peer models participated in the study). When selecting peer models, all participants selected peer models with the same gender and similar ethnicity to themselves with the exception of two participants, who selected a peer model with the same gender but different ethnicity.

4.2 Results

The performance results are illustrated in Figure 3. All statistics are calculated using the JASP statistical software package [87]. The hypotheses were tested using paired, one-tailed t-tests where possible. If a Shapiro-Wilk test indicated that the assumption of normality was violated, a non-parametric Wilcoxon signed-rank test was used instead. Table 1 summarises the performance results for the six exercises. In the absence of a hypothesis a two-tailed t-test or Wilcoxon signed-rank test was used, as appropriate. For the Student t-test, effect size is given as Cohen's d, while for Wilcoxon tests it is given by the matched rank biserial correlation. We tested for order effects using a mixed ANOVA with video (P and F) as the within-subjects factor and order (P-F and F-P) as the between-subjects factor, but found that none of these were significant for *Order* or for the interaction of *Video* $\times$ *Order*, meaning the order of videos in the second and third sessions did not affect the results. A power analysis indicates that we were able to detect medium effects (Cohen's d ≥ 0.523) according to our hypotheses at $\alpha = .05$ with a power of 0.8. The error bars in Figure 3 show 95% confidence intervals of the mean which are calculated using normalized values to better represent the experimental manipulation [102].

Manipulation Check. Identification questionnaire scores for the peer videos were significantly greater than for the original YouTube videos in all exercises ($p < .002^{**}$), with medium to large effect size (≥ 0.674), indicating that the peer model videos were a better representation of the user than the generic YouTube videos selected. Identification questionnaire scores for the FakeForward videos were significantly greater than for the peer videos of the models in all but two of the exercises. FakeForward resulted in greater reported similarity for squats ($t(23) = 2.55, p = .009^{**}$), planks ($t(23) = 2.32, p = .015^{*}$), press ups ($t(23) = 2.40, p = .012^{*}$), and sit ups ($W(23) = 21.0, p = .003^{**}$) with medium effect size ($\geq .474$). However, despite the effect size in the correct direction, both wall sits ($W(23) = 57.5, p = .115$) and squat jumps ($W(23) = 60.5, p = .085$) were not significant with effect sizes of $r = 0.327$ and $r = 0.363$ respectively.

4.2.1 Wall Sits. Participants performed significantly longer wall sits in F (M=79.5, SD=38.5) compared to P (M=71.0, SD=32.3) ($t(23) = 2.13, p = .022^{*}$), with a medium effect size ($d = 0.436$); therefore we accept H1.1. Participants also performed significantly longer wall sits in P compared to B (M=51.2, SD=25.6) ($t(23) = 4.47, p < .001^{***}$) with a large effect size ($d = 0.913$). For wall sits, participants' exercise self efficacy was significantly higher after watching the FakeForward video (M=4.45, SD=1.45) compared with the peer model video (M=4.10, SD=1.58) ($t(23) = 2.17, p = .041^{*}$) with a medium effect size ($d = 0.442$). None of the other psychometric tests yielded significant differences between F and P conditions.

4.2.2 Squats. Participants performed significantly more squats in F (M=48.6, SD=25.5) compared to P (M=43.8, SD=30.3) ($W(23) = 82.0, p = .046^{*}$), with a medium effect size ($r = 0.406$); therefore we accept H1.2. Participants also performed significantly more squats in P compared to B (M=33.1, SD=25.5) ($W(23) = 34.0, p = .002^{**}$) with a medium to large effect size ($r = 0.706$). Responses for the FSQ balance sub-scale were significantly different, with signifcantly higher scores in the F condition (M=3.94, SD=0.773) compared with P (M=3.64, SD=0.944) ($t(23) = 2.69, p = .013^{*}$) with a medium effect size ($d = 0.549$). None of the other psychometric tests yielded significant differences between F and P conditions.

Table 1: Statistical results of performance comparisons between P and F conditions across the six physical exercises, demonstrating how FakeForward leads to significant performance improvements in four of the six exercises. *Note.* For the Student t-test, effect size is given by Cohen's *d*. For the Wilcoxon test, effect size is given by the matched rank biserial correlation.

Measure	Test	Statistic	df	p	Effect Size
Wall sits	Student	2.134	23	0.022^{*}	0.436
Squats	Wilcoxon	82.000	23	0.046^{*}	0.406
Plank	Student	1.171	23	0.127	0.239
Press ups	Student	1.458	23	0.079	0.298
Sit ups	Student	2.130	23	0.022^{*}	0.435
Squat jumps	Wilcoxon	65.000	23	0.024^{*}	0.486

4.2.3 Plank. There were no significant differences with the amount of time participants could hold the plank position between F (M=78.8, SD=36.9) and P (M=73.8, SD=33.1) ($t(23) = 1.17, p = .127$), with a small effect size ($d = 0.239$); therefore we reject H1.3. However, participants could hold the plank position for a significantly longer amount of time in P compared to B (M=61.5, SD=33.3) ($t(23) = 3.16, p = .002^{**}$) with a medium effect size ($d = 0.645$). Participants scored significantly higher on the IMI competency sub-scale after having watched the FakeForward video (M=4.24, SD=1.83) compared with the peer model video (M=4.06, SD=1.86) ($t(23) = 2.18, p = .039^{*}$) with a medium effect size ($d = 0.446$). None of the other psychometric tests yielded significant differences between F and P conditions.

4.2.4 Press-ups. There was not a significant difference between the number of press ups participants could do in the F (M=21.4, SD=9.86) compared with P (M=19.3, SD=10.0) ($t(23) = 1.46, p = .079$), with a small effect size ($d = 0.298$); therefore we reject H1.4. However, participants could perform significantly more press ups in P compared to B (M=16.1, SD=9.78) ($t(23) = 2.91, p = .004^{**}$) with a medium effect size ($d = 0.593$). None of the psychometric tests yielded significant differences between the two conditions.

4.2.5 Sit-ups. Participants performed significantly more sit ups in F (M=21.0, SD=9.38) compared to P (M=18.5, SD=8.82) ($t(23) = 2.13, p = .022^{*}$), with a medium effect size ($d = 0.435$); therefore we accept H1.5. Additionally, participants performed significantly more sit ups in P compared to B (M=13.5, SD=7.18) ($W(23) = 2.00, p < .001^{***}$) with a large effect size ($r = 0.987$). None of the psychometric tests yielded significant differences between the two conditions.

4.2.6 Squat Jumps. Participants performed significantly more squat jumps in F (M=22.7, SD=7.81) compared to P (M=20.4, SD=9.06) ($W(23) = 65.0, p = .024^{*}$), with a medium effect size ($r = 0.486$); therefore we accept H1.6. Participants also performed significantly more squat jumps in P compared to B (M=16.0, SD=6.45) ($W(23) =$

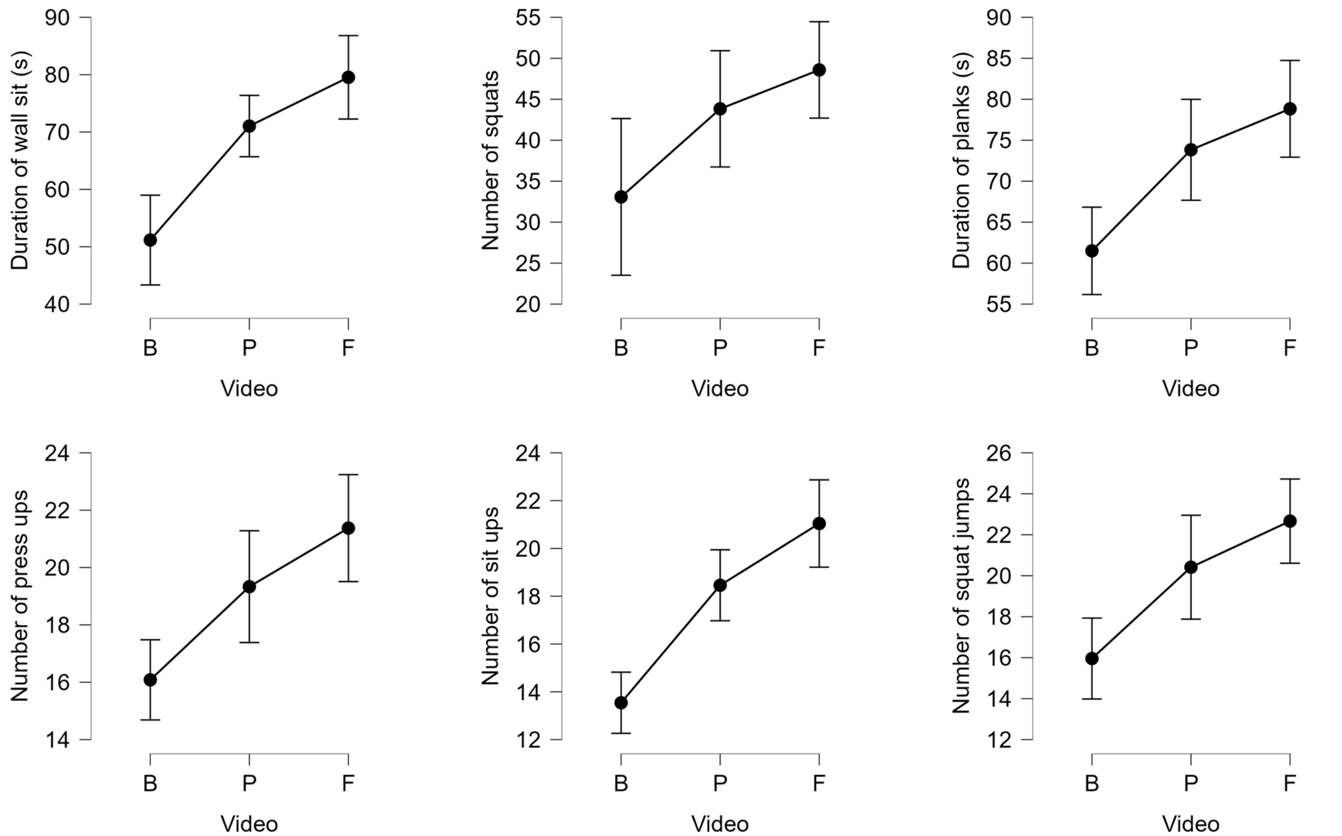

Figure 3: Study 1 performance results for baseline (B), peer model (P) and FakeForward (F) videos. The error bars show 95% confidence intervals of the mean which are calculated using normalized values to better represent the experimental manipulation.

40.5, p = .002**) with a medium to large effect size (r = 0.707). None of the psychometric tests yielded significant differences between the two conditions.

4.2.7 Qualitative Results. To gain further insight into the effects of FakeForward we conducted a reflexive thematic analysis, where common themes and sub-themes were generated inductively from the data [16]. We discuss the overarching themes and sub-themes using participant quotes as illustrative examples; text in square brackets is used to add context to a quote to make it easier to understand.

The majority of participants (N=17) reported that watching the FakeForward videos helped them to perform better, with only three participants thinking the peer models helped their performance, and the remaining four having no preference. The thought process as to why FakeForward might increase performance is unclear ("gut feeling", "I felt more immersed"), and there may have been a novelty factor at play ("entertaining seeing the deepfake, [...] quite fun to watch"). Interestingly though, a participant's memory of their performance did not necessarily correlate with what they achieved (e.g, "only the plank video worked on me" – for this participant this was not the case and the participant's performance was *worse* for the plank in the FakeForward condition).

Watching the FakeForward videos can motivate participants to perform better. Participants report how watching the FakeForward videos increased their motivation and provided a realistic future image of themselves ("*this is what I would look like if I keep working on it*", "*like a fitter and better version of me was doing it [...] motivated me in the sense that it gave me a realistic view of what I could achieve*", "*It makes me feel I can do it*", "*I think "oh maybe I can do it"*", "*I think it's helpful for me to imagine that I can do the same thing as the person in the videos*").

However, not all participants preferred the FakeForward videos. Despite 18 participants thinking it improved their performance, only 12 preferred watching the FakeForward videos. Six preferred the peer model videos and six had no preference. Participants did not like the FakeForward videos, raising issues similar to the uncanny valley [103] ("*it felt a little unnatural*", "*gave me a little bit of uneasy feeling*", "*I thought it's kind of weird*") and some participants preferred the 'realism' of the peer models ("*brings me a realistic feeling*", "*feels more natural*"). However, participants actively preferred watching the peer videos because of a sense of competition ("*seeing another guy made me compete with him*"), or because of social connections ("*I know the peer model as a friend too, so it helped me perform better*"). Interestingly, competitiveness also came up in the context of the FakeForward video ("*Still gave me the competitiveness I felt with the normal peer model videos*").

Future iterations of FakeForward should concentrate on more than just face swapping. Participants commented on how the model lacked behavioural elements, such as facial expressions ("*you really really need to elaborate on the facial expressions*"). In addition, the dissimilarity of body characteristics was also raised as an issue ("*the body structure was very different*", "*does not feel real because of the physical attributes [...] having a very different body structure compared to the peer model*", "*did not have the same physical characteristics*", "*strange seeing my face on someone else's body*"). However, for some participants the change in body characteristics appeared to be a positive factor ("*seeing my face pictured onto a fitter person made me realise I could also reach there*").

4.3 Summary

This study enabled us to develop and test a protocol for the use of FakeForward (Section 3), addressing RQ1. The importance of redundant video material emerged when some parts of videos could not be recognised by the DeepFaceLab software. Sharp movements and unfavourable face angles, e.g. during press-ups, made it difficult to extract clear facial images at times, and some participants noted how doing these exercises with their faces visible was physically uncomfortable. These challenges reduced the quality of some of the FakeForward videos; however, identification questionnaire scores and participant comments indicate that overall, an increased level of self similarity and recognition was achieved. Interestingly, similarity was only significantly different in four of the six physical exercises, with wall sits and squat jumps not being significantly different despite the effect being in the right direction. Notwithstanding this, performance was significantly better for both of these exercises. Combined with the lack of significant results in the psychometric tests, this may indicate there are deeper cognitive processes involved in the efficacy of FakeForward.

The study provides first insights into the areas of training that can be supported with FakeForward and its efficacy, addressing RQ2 and RQ3: significant medium effects for wall sits (H1.1), squats (H1.2), sit ups (H1.5), and squat jumps (H1.6) support its efficacy in developing strength endurance. Quick improvements in neuromuscular coordination and control are possible for strength-related tasks [91, 109], and such quick improvements have also been reported in the feedforward literature [44, 45]. The improvements could in principle be due to psychological factors such as intrinsic motivation and flow; however, we were unable to detect any consistent effects of FakeForward on these variables. FakeForward also worked for participants who preferred the original videos. This supports the mechanism proposed in the feedforward literature [44, 45], which predicts neuromuscular improvements through mirror neuron activity.

However, we did not observe performance benefits for press-ups or plank. Although intended to engage different muscle groups (dynamic upper and static lower respectively), the form of these two physical exercises is similar which may provide insights into why we did not see significant performance gains. The specific nature of the muscles used in these exercises may offer an explanation: the upper body muscles used in press-ups as well as planks are all comparatively small, so their smaller muscle energy reservoirs are more easily depleted [52]. These are muscles groups that participants may not commonly engage relative to the legs and core, which are used everyday for walking, climbing stairs, and stabilisation of the core. Alternatively, the way in which these exercises were recorded may have affected the performance results. We recorded videos low to the ground so that participants could see their (or the peer model's) face in the recording; however, it is uncommon for these exercises to be seen from such a perspective, which may have impeded identification and mirror neuron activity. Finally, it is worth nothing that in both cases the effect is in the right direction ($F > P$), however the effect sizes are small, meaning that the study would have been under-powered to detect such small changes and further work with higher statistical power is required.

5 STUDY 2: PUBLIC SPEAKING

The aim of Study 2 is to further determine which types of skills FakeForward is effective (RQ2) by considering a task that focuses on psychological rather than physical skills: public speaking combines multiple important psychological functions including cognition and language, self-confidence, emotional regulation and motivation [88]. Furthermore, public speaking is a skill many people struggle with due to a very high prevalence of public-speaking anxiety [116]. It would therefore be particularly useful if FakeForward could help people feel more comfortable during public speaking. To test the efficacy of FakeForward (RQ3), we compared psychological measures after watching a FakeForward video vs. after watching the corresponding original video.

5.1 Methodology

We used a within-participant experimental design in which participants took part in two sessions. In Session 1, a participant was recorded giving a talk to the experimenter and baseline (B) measurements were collected. Between the sessions, the participant was assigned a model in the form of a popular short TED talk video, with a presenter that was similar in appearance (match of gender and skin colour, as well as other visual characteristics as much as possible). TED talks were chosen because of the overall good reputation of their presentation quality and their success in disseminating ideas [134]. The face of the presenter was then replaced by the face of the participant to form the participant's FakeForward video. In Session 2, participants were shown the original video (O) and then briefly prepared and gave a talk about a topic, and also shown their FakeForward (F) video followed by them giving another talk about a different topic. The order of the two conditions (O and F) and the topics given to participants in each condition were counterbalanced to mitigate order effects and avoid confounding factors.

5.1.1 Task. The task was overall the same in all conditions (B, O and F): Participants were given up to 10 minutes to prepare a 2-minute talk about a topic. Then they gave the talk while facing the experimenter. In condition B, participants were given a list of topics to choose one they were comfortable with, e.g. questions such as "Can artificial intelligence replace human beings?" and "Should animals have the same rights as humans?". Alternatively, they had the option to choose their own topic. For conditions O and F their topics were given as "Can money buy happiness?" and "Does learning require interest?", counterbalanced across the conditions to avoid confounding factors.

5.1.2 Outcome Measures. The Personal Report of Confidence as a Speaker (PRCS) scale [110] is a validated measure of public speaking anxiety, which is used widely in treatment and research [83]. It comprises 30 items with yes/no responses that evaluate the emotional and behavioural responses of public speaking, e.g. "I look forward to the opportunity to speak in front of the public" and "I am always afraid to forget the content of my speech". The PRCS is scored from 0 to 30, with higher scores indicating higher anxiety.

Speaking performance was measured using the Public Speaking Evaluation rubric of the University of Vermont [106], which is used to evaluate public speeches of many students every year. The rubric is based on nine categories of well-accepted public speaking criteria such as an engaging introduction, clarity, memorability and focus [88]. Each criterion was rated on a weighted 5-point scale by the same researcher, yielding an overall score between 0 and 40.

Similar to the previous studies, we also used the Interest/Enjoyment and Perceived Competence subscales of the Intrinsic Motivation Inventory (IMI) [120] and the Physical Similarity, Wishful Identification and Liking subscales of the polythetic identification model [42]. As in Study 2, we applied the IMI scales to measure Interest/Enjoyment and Perceived Competence both right after watching a video and after giving a talk.

5.1.3 Procedure. Due to COVID-19 restrictions, this study was conducted remotely through video calls. Similar to the previous studies, we used session 1 to record a video of the participant giving a talk about a topic of their choice, based on the FakeForward protocol from Section 3, and collected baseline (B) measurements for PRCS, IMI and speaking performance. Session 1 took on average about 30 minutes. Then we selected a model for each participant from 17 popular TED videos, matching gender and skin colour, as well as other salient visual characteristics such as hair colour and hair style, to achieve a visual similarity as close as possible. From each of these TED videos, we extracted a 40-second clip that was a) self-contained, i.e. could be understood without watching the rest of the video, b) consisted predominantly of close-up shots of the face of the speaker, and c) was found by the research team to demonstrate good oratory skills, e.g. being articulate, motivational and generally impressive. These clips were our original videos (O), and we then used the FaceSwap software [53] to replace the face in each model video with the face of the corresponding participant, resulting in their FakeForward video (F). About a week later in Session 2, participants ran through our two experimental conditions O and F in counterbalanced order. In each condition, they were first shown their respective video and answered the identification and IMI questionnaires (about how they felt when watching the video). Then they prepared and gave their own talk, followed by the IMI (about how they felt when giving the talk) and PRCS questionnaries. Each talk was observed and scored by the same researcher using the Public Speaking Evaluation rubric. Finally, we asked participants about their preference and for any general comments ("*Did you prefer the original peer model videos or the deepfake videos?*", "*Please can you explain your answer?*", and "*Do you have any general comments?*"). Session 2 took on average about 45 minutes.

5.1.4 Hypotheses. We posed the following *a-priori* hypotheses based on predictions from feedforward theory:

H2.1 Participants have higher speaking performance ratings when giving their talk after watching the FakeForward video (F) compared to the original (O).

H2.2 Participants have higher perceived competence when giving their talk after watching the FakeForward video compared to the original.

H2.3 Participants have lower public speaking anxiety after watching the FakeForward video compared to the original.

5.1.5 Participants. We recruited 20 participants (8 female, 12 male), aged 20–32 (M = 25.10, SD = 3.27) through social networks and word of mouth. According to a 7-point Likert scale question about their public speaking experience (1 signifying no experience and 7 a lot of experience), the participants were overall moderately experienced ($M = 3.53, SD = 2.30$) with a wide range of levels of experience.

5.2 Results

The performance results are illustrated in Figure 4. We used the same statistical methodology as in Study 1. A power analysis indicates that we were able to detect medium effects (Cohen's d ≥ 0.577) at $\alpha = .05$ with a power of 0.8.

Manipulation Check. Identification questionnaire scores for the FakeForward videos were significantly greater than for the original videos of the models: Physical Similarity was greater ($t(19) = 7.660, p < .001^{***}$) with a very large effect size (Cohen's d = 1.713), Wishful Identification was greater ($t(19) = 2.762, p = .006^{**}$) with a medium effect size (Cohen's d = 0.618), and Liking was greater ($t(19) = 7.743, p < .001^{***}$) with a very large effect size (Cohen's d = 1.731), indicating that the FakeForward videos were successful in eliciting self recognition, identification and positive affect.

5.2.1 Speaking Performance Rating. Participants' speaking performance ratings of their talk after the FakeForward video (F) were significantly higher compared to their talk after the original video (O) ($t(19) = 2.173, p < .021^{*}$), with a medium effect size (Cohen's d = 0.486); therefore we accept H3.1. The difference in speaking performance ratings between the baseline talk (B) and their talk after the original video was not significant ($t(19) = 1.571, p = .133$).

5.2.2 IMI Perceived Competence. Participants' Perceived Competence while giving a talk after the FakeForward video (F) was significantly higher compared to their talk after the original video (O) ($t(19) = 4.271, p < .001^{***}$), with a large effect size (Cohen's d = 0.955); therefore we accept H3.2. Furthermore, participants' Perceived Competence while watching the FakeForward video (F) was significantly higher compared to watching the original TED video (O) ($t(19) = 5.752, p < .001^{***}$), with a large effect size (Cohen's d = 1.286). Participants' Perceived Competence while giving a talk after watching the original video was not significantly different from the baseline (B) ($t(19) = 2.026, p = .057$).

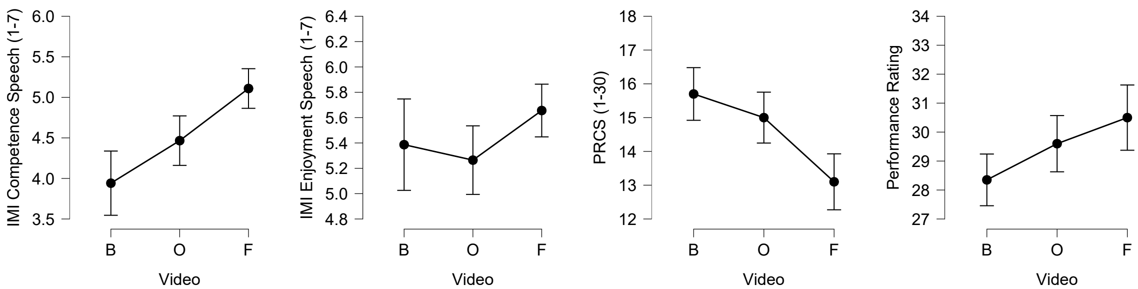

Figure 4: Study 2 IMI Perceived Competence and Interest/Enjoyment after giving a talk, PRCS scores, and speaking performance ratings for baseline (B), original (O) and FakeForward (F) videos. The error bars show 95% confidence intervals of the mean.

5.2.3 *IMI Interest/Enjoyment.*

Participants had a significantly higher Interest/Enjoyment when watching the FakeForward video ($t(19) = 6.710, p < .001^{***}$) compared to watching the original video, with a large effect size (Cohen's d = 1.500). The main effect of the conditions (B, O and F) on Interest/Enjoyment when giving a talk was not significant ($F(1.532, 29.102) = 2.138, p = .146$).

5.2.4 *Personal Report of Confidence as a Speaker (PRCS).*

Participants' PRCS scores after watching their FakeForward video and giving a talk (F) were significantly lower compared to after watching the original TED video and giving a talk (O) ($t(19) = 3.535, p = .001^{**}$), with a large effect size (Cohen's d = 0.790), indicating that participants felt less anxious and more confident after the FakeForward video; therefore we accept H3.3. Participants' PRCS scores after watching the original video and giving a talk were not significantly different compared to baseline (B) ($t(19) = 1.391, p = .180$).

5.2.5 *Qualitative Results.*

Although it was obvious to all participants that the person in the FakeForward video was not themselves ("*it is very obvious that the TED video is not me*"), they noticed the similarity ("*I do know that the person is not me but it looks like me with a new hair cut.*", "*when I saw my face it was kinda weird*"). One participant who watched the FakeForward video first, appeared to see more similarity when watching the original video later (*"After watching the deepfake video, I felt that when I watched the original video later, I had ignored his face in the first 5-10 seconds, which made me feel that we look a bit similar with each other."*). Most did not notice that they sounded different from themselves when they first saw the FakeForward video (*"Oh! I did not notice until you asked me. Indeed, we sound different, but I did not notice this when watching the video. It might be due to my excitement that my attention was focused on the face and also it is quite a short video"*). However, all participants asked if they could replay the video one or two more times after they had completed the study. With repeated viewing, participants generally became aware of the different voice.

All participants preferred watching the FakeForward video over the original. Most mentioned how the FakeForward video made them feel better and more confident (*"because it is using my face and somehow I literally feel that I am able to give such a good performance in front of the public. It gives me a lot of strength."*, *"It's hard to explain, but when I saw myself standing there and speaking, I kinda feel proud of myself"*). Two participants described the comparative ordinariness of the original video (*"not bad but there's nothing different"*, *"just like watching a TED video"*), and the training opportunities they saw in FakeForward (*"I'll use the fake video as tutorial."*, *"I am actually looking forward to public speaking now, I would like to see if I can be as calm as in the video. I have never thought about it like this before."*, *"Although I am still afraid of public speaking, the deepfake video makes me feel that speaking is actually not that scary. I think my second speech (after watching deepfake) performance should be the best. That video made me feel like if I could be so calm of facing a hundred people"*).

5.3 Summary

Study 2 indicates that FakeForward can also be effective for tasks focusing on psychological skills, such as public speaking. This is consistent with research on traditional feedforward through VSM [44, 45] and can be explained through increased mirror neuron activity and other psychological effects such as self-efficacy and social learning [6]. Interestingly, the fact that it was not the participant's voice in the FakeForward video did not prevent people from experiencing the positive effects of watching their visual self-models. This could be due to a common perception that one's own voice sounds unfamiliar when recorded [78], i.e., an unfamiliar voice may not have significantly affected identification with the self-model.

Limitations. Visual quality of the videos varied and data collection had to be performed remotely due to COVID-19. The Public Speaking Evaluation rubric has not been formally validated and performance was only scored by a single researcher, so the performance results should be considered with care; however, they are backed up by validated measures (PRCS and IMI). We did not include a larger audience, which may have affected the results in all conditions, due to COVID-19 protocols being in place at the time of the study. In future, investigating how FakeForward affects public speaking in front of a larger audience may provide further insights due to the increased social anxiety that presenters may face.

6 DISCUSSION

To address how deepfake technology can be used for creating video self-models (RQ1), we presented ethical and technical protocols for generating FakeForward videos. The technical protocol was used, refined and validated in Study 1 based on video footage from participants, while Study 2 extended the use to existing professionally recorded videos. The quality of the FakeForward videos varied between participants and some noted that the person was not them in the video, especially in tasks involving more movement. Despite this, the protocol resulted in FakeForward videos which elicited significantly greater self recognition and identification compared to videos of peer models – an important factor for the success of VSM and feedforward interventions [45].

The current protocol requires a source video of the user featuring similar pose angles as featured in the desired video – although not necessarily performing the activity. Advances in the generalisation of deepfake algorithms may remove or reduce this requirement, allowing users to match themselves to a wider variety of target videos with a single source video. In future, it is also possible that self-identification, and the effect of feedforward, increase as deepfake technology improves. Similarly, the use of more advanced deepfake algorithms, such as body swapping [24], have the potential to remove the requirement for the peer model to have a similar physical appearance to the user.

The principle of feedforward is based on a user's level of identification with the model in the video, and is currently applicable for videos where the face is visible to leverage face swapping capabilities of deepfakes. This raises questions about what other methods of increasing identification could contribute to increased performance benefits. For example, video tutorials of manual tasks with close up of the hands are beyond the remit of face swapping technologies, however, the hands are unique and can be used for identification [73] with prior work in VR demonstrating how some individual's sense of presence is affected by the visual representation of their hands [127]. Similarly, in Study 2 some participants commented how the voice of the speaker was not their own. Audio deepfakes may be able to increase identification with the self-model and potentially unlock more benefits.

We sought to understand what areas of training and training FakeForward can support (RQ2), as well as the efficacy of FakeForward in helping users to develop their skills (RQ3). For physical exercises we observed performance gains for the lower and core muscle groups and for both types of motor functions (static and dynamic). More specifically, we observed that FakeForward provided significant performance improvements for wall sits (static lower), sit-ups (dynamic core), squats (dynamic lower), and we also saw performance improvements in the cardiovascular exercise (squat jumps). Our results suggest that FakeForward may also be effective for other similar physical exercises. For two physical exercises FakeForward did not result in significant performance increases. Both plank (static core) and press-ups (dynamic upper) showed only non-significant, small effect sizes (Cohen's d = 0.239 and 0.298). Our study did not have enough power to detect this, so further work with higher statistical power is needed.

Despite positive qualitative comments, psychological variables such as intrinsic motivation, exercise self-efficacy, perceived competence, and flow did not explain the improvements either. Feedforward theory suggests that sub-conscious processes, such as "mirror neurons" and other associated circuitry, are activated more strongly by self-models, which can lead to the associated performance gains [45]. Mirror neurons have been speculated as an approach to offsetting fatigue in dance science research [11, 33], and these sub-conscious processes may have enabled participants to unlock extra performance by affecting central muscle fatigue, which occurs due to the inability of motor neurons to drive muscle fibres during voluntary muscle contractions [36, 58]. Interestingly, recent work has discussed how mirror neurons have been hypothesised to increase fatigue and decrease performance when a person observes someone fatiguing or in pain due to perceptual empathic response [12] – the opposite effect of feedforward. The underlying reasons for if and how mirror neurons affect performance and fatigue, and their exact role in the feedforward effect, are unclear and still speculative [34].

Beyond physical skills, Study 2 illustrated how FakeForward can be effectively applied to tasks which focus on psychological ability such as public speaking. The positive effects on enjoyment and perceived competence here contrast our findings for physical exercises, in which these psychological variables were unaffected. A possible explanation is that these psychological factors are inherent parts of public speaking: they are expressed in the FakeForward videos through a speaker's behaviour, language, charisma and confidence – unlike the videos showing physical activities – and are consequently 'trained'. The results suggest that FakeForward may also hold promise for other psychological training scenarios where VSM has been successful, e.g. to address behavioural and educational challenges [13, 14, 44].

FakeForward has demonstrated performance benefits for a variety of physical tasks, but future work is required to identify the extent to which it can be successfully applied beyond the activities studied in this paper. More specifically, the physical exercises in this study focused on FakeForward as a support tool for training, helping some participants increase their performance. However, VSM has traditionally been used as a learning support tool and understanding if, and how well, FakeForward can support learning and development of psychomotor skills remains an open question. Additional understanding is required about the duration of the feedforward effect, i.e., how soon should one perform the task after having watched the FakeForward video, or how the effect varies over a prolonged period which is especially important when understanding learning curves. Similarly, the participants recruited for the study were not professional athletes or public speakers, raising the question about how FakeForward differs across different skill levels – from beginners learning a new task to professional athletes.

When applying FakeForward in practice on a larger scale, it may be challenging to find suitable peer models for every user because peer models need to be similar in appearance as well as showing a better performance – but one that is realistically attainable. To overcome this, FakeForward users could be encouraged to share their own videos for others to use as peer models. This could lead to the creation of public peer model repositories, covering different people, skills and levels of performance. In order to use such a repository effectively, users would need to receive guidance about how to choose a peer model and in particular, which level of performance in a peer model would be most beneficial for them. Ideally, the skill level for a peer model could be automatically detected, allowing a system to suggest optimal peer models for users which maximise the feedforward effect.

6.1 Limitations

Creating the FakeForward videos was computationally intensive: PCs with high-end GPUs took a minimum of 24 hours to generate satisfactory deepfake videos. Approaches using cloud computing, specialised hardware and more efficient algorithms may be able to mitigate this and make FakeForward more immediately accessible to the wider public. Our study participants were mainly in their twenties, which limits the generalisability of our results. However, our results are in line with the traditional VSM literature [20, 44, 45, 47, 96], which has confirmed similar results for different age groups. For the physical exercises, all of the data was collected in a controlled environment in a research laboratory. The ultimate goal with FakeForward is for the technology to assist people with physical activities in their own home, and whether these effects manifest for physical exercises through remote applications is an open research question. However, as part of our data collection had to be conducted remotely due to COVID-19 restrictions we demonstrate the potential of FakeForward to be effective in remote applications, making its benefits available through the Internet.

6.2 Impact

Decades of research have shown how VSM can enhance learning and skill development. FakeForward demonstrates how deepfake technologies can simplify and 'democratise' the creation of VSM videos, bringing the advantages of VSM closer to the general public. FakeForward has shown success in supporting exercise performance and hence physical health, and in improving confidence and perceived competence, which are linked to mental health [116]. This raises hopes that it could be applied to a wider variety of skills and to affect positive behavioural and psychological changes. For example, it could be applied in interventions designed for specific user groups, such as children with autism spectrum disorder, which have been shown to benefit from VSM [45].

7 CONCLUSIONS

We have introduced FakeForward, a novel method using deepfake technology to generate VSM videos for skills development. Technical and ethical protocols provide guidance on how to generate successful FakeForward videos, and two user studies revealed that FakeForward can rapidly improve a variety of skills. We conclude that:

(1) FakeForward is effective at improving performance for many physical skills including strength endurance and cardio tasks.
(2) FakeForward is effective at improving performance for some tasks focused on psychological abilities such as public speaking, improving self-confidence and reducing anxiety.

ACKNOWLEDGMENTS

This work was 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] Marwan Albahar and Jameel Almalki. 2019. Deepfakes: Threats and countermeasures systematic review. *Journal of Theoretical and Applied Information Technology* 97, 22 (2019), 3242–3250.

[2] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 311–320.

[3] Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.

[4] Albert Bandura. 1986. *Social foundations of thought and action: A social cognitive theory*. New Jersey: Prentice Hall.

[5] Albert Bandura. 1997. The anatomy of stages of change. *American journal of health promotion: AJHP* 12, 1 (1997), 8–10.

[6] Albert Bandura, WH Freeman, and Richard Lightsey. 1999. Self-efficacy: The exercise of control.

[7] Albert Bandura, Dorothea Ross, and Sheila A Ross. 1961. Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology* 63, 3 (1961), 575.

[8] Soumya C Barathi, Daniel J Finnegan, Matthew Farrow, Alexander Whaley, Pippa Heath, Jude Buckley, Peter W Dowrick, Burkhard C Wuensche, James LJ Bilzon, Eamonn O'Neill, et al. 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*. 1–14.

[9] Norman J Barry and Phyllis B Overmann. 1977. Comparison of the effectiveness of adult and peer models with EMR children. *American Journal of Mental Deficiency* (1977).

[10] Demis Basso and Marta Olivetti Belardinelli. 2006. The role of the feedforward paradigm in cognitive psychology. *Cognitive Processing* 7, 2 (2006), 73–88.

[11] Glenna Batson. 2013. Exercise-induced central fatigue: A review of the literature with implications for dance science research. *Journal of Dance Medicine & Science* 17, 2 (2013), 53–62.

[12] David G Behm and Tori B Carter. 2021. Empathetic Factors and Influences on Physical Performance: A Topical Review. *Frontiers in Psychology* 12 (2021).

[13] Scott Bellini and Jennifer Akullian. 2007. A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional children* 73, 3 (2007), 264–287.

[14] Janine Bolger. 2014. Video self-modelling and its impact on the development of communication skills within social work education. *Journal of Social Work* 14, 2 (2014), 196–212.

[15] Cynthia J Brame. 2016. Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education* 15, 4 (2016), es6.

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

[17] Douglas Brooks. 2001. *Effective strength training: Analysis and technique for upper-body, lower-body, and trunk exercises.* Human Kinetics.

[18] Tom Buggey. 2005. Video self-modeling applications with students with autism spectrum disorder in a small private school setting. *Focus on autism and other developmental disabilities* 20, 1 (2005), 52–63.

[19] Tom Buggey, Grace Hoomes, Mary Elizabeth Sherberger, and Sarah Williams. 2011. Facilitating social initiations of preschoolers with autism spectrum disorders using video self-modeling. *Focus on Autism and Other Developmental Disabilities* 26, 1 (2011), 25–36.

[20] Tom Buggey and Lindsey Ogle. 2012. Video self-modeling. *Psychology in the Schools* 49, 1 (2012), 52–70.

[21] Tom Buggey, Kristina Toombs, Pia Gardener, and Michele Cervetti. 1999. Training responding behaviors in students with autism: Using videotaped self-modeling. *Journal of Positive Behavior Interventions* 1, 4 (1999), 205–214.

[22] Cami E Burton, Darlene H Anderson, Mary Anne Prater, and Tina T Dyches. 2013. Video self-modeling on an iPad to teach functional math skills to adolescents with autism and intellectual disability. *Focus on Autism and Other Developmental Disabilities* 28, 2 (2013), 67–77.

[23] Katherine S Cennamo. 1993. Learning from video: Factors influencing learners' preconceptions and invested mental effort. *Educational Technology Research and Development* 41, 3 (1993), 33–45.

[24] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A Efros. 2019. Everybody dance now. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 5933–5942.

[25] Minsuk Chang, Anh Truong, Oliver Wang, Maneesh Agrawala, and Juho Kim. 2019. How to design voice based navigation for how-to videos. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[26] Marjorie H Charlop and Janice P Milstein. 1989. Teaching autistic children conversational speech using video modeling. *Journal of applied behavior analysis* 22, 3 (1989), 275–285.

[27] Pei-Yu Chi, Joyce Liu, Jason Linder, Mira Dontcheva, Wilmot Li, and Bjoern Hartmann. 2013. Democut: generating concise instructional videos for physical demonstrations. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 141–150.

[28] Rachel B Clancy, Matthew P Herring, and Mark J Campbell. 2017. Motivation Measures in Sport: A Critical Review and Bibliometric Analysis. *Frontiers in*

Psychology 8 (2017).
[29] Shannon E Clark and Diane M Ste-Marie. 2007. The impact of self-as-a-model interventions on children's self-regulation of learning and swimming performance. *Journal of sports sciences* 25, 5 (2007), 577–586.
[30] Christopher Clarke, Doga Cavdir, Patrick Chiu, Laurent Denoue, and Don Kimber. 2020. Reactive Video: Adaptive Video Playback Based on User Motion for Supporting Physical Activity. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 196–208.
[31] Melissa A Collier-Meek, Lindsay M Fallon, Austin H Johnson, Lisa MH Sanetti, and Marisa A Delcampo. 2012. Constructing self-modeling videos: Procedures and technology. *Psychology in the Schools* 49, 1 (2012), 3–14.
[32] Catherine Coyle and Peter Cole. 2004. A videotaped self-modelling and self-monitoring treatment program to decrease off-task behaviour in children with autism. *Journal of Intellectual and Developmental disability* 29, 1 (2004), 3–16.
[33] Emily S Cross, Antonia F de C Hamilton, and Scott T Grafton. 2006. Building a motor simulation de novo: observation of dance by dancers. *Neuroimage* 31, 3 (2006), 1257–1267.
[34] Alessandro D'Ausilio, Eleonora Bartoli, and Laura Maffongelli. 2015. Grasping synergies: a motor-control approach to the mirror neuron mechanism. *Physics of life reviews* 12 (2015), 91–103.
[35] Paul R Davidson and Daniel M Wolpert. 2005. Widespread access to predictive models in the motor system: a short review. *Journal of neural engineering* 2, 3 (2005), S313.
[36] J Mark Davis and Stephen P Bailey. 1997. Possible mechanisms of central nervous system fatigue during exercise. *Medicine and science in sports and exercise* 29, 1 (1997), 45–57.
[37] DeepFaceLab. 2022. *DeepFaceLab Github page*. Retrieved September 3, 2022 from https://github.com/iperov/DeepFaceLab
[38] Nicholas Diakopoulos and Deborah Johnson. 2021. Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society* 23, 7 (2021), 2072–2098.
[39] Malaria Must Die. 2019. *David Beckham speaks nine languages to launch Malaria Must Die Voice Petition*. Retrieved August 20, 2021 from https://www.youtube.com/watch?v=QiiSAvKJIHo
[40] Michael Domjan, Brian Cusato, and Ronald Villarreal. 2000. Pavlovian feed-forward mechanisms in the control of social behavior. *Behavioral and Brain Sciences* 23, 2 (2000), 235–249.
[41] Mary Jo Dondlinger. 2007. Educational video game design: A review of the literature. *Journal of applied educational technology* 4, 1 (2007), 21–31.
[42] 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.
[43] PW Dowrick and M Hood. 1978. Transfer of talking behaviors across settings using faked films. In *Proceedings of New Zealand conference for research in applied behavior analysis*. 78–87.
[44] Peter W Dowrick. 1999. A Review of Self Modeling and Related Interventions. *Applied and Preventive Psychology* 8, 1 (1999), 23–39.
[45] Peter W Dowrick. 2012. Self model theory: Learning from the future. *Wiley Interdisciplinary Reviews: Cognitive Science* 3, 2 (2012), 215–230.
[46] Peter W Dowrick. 2012. Self Modeling: Expanding the Theories of Learning. *Psychology in the Schools* 49, 1 (2012), 30–41.
[47] Peter W Dowrick and Simon J Biggs. 1984. *Using video: Psychological and social applications*. Wiley.
[48] Peter W Dowrick and Cynthia Dove. 1980. The use of self-modeling to improve the swimming performance of spina bifida children. *Journal of Applied Behavior Analysis* 13, 1 (1980), 51–56.
[49] Peter W Dowrick and Karen M Ward. 1997. Video feedforward in the support of a man with intellectual disability and inappropriate sexual behaviour. *Journal of Intellectual and Developmental Disability* 22, 3 (1997), 147–160.
[50] Peter Duffy. 2008. Using Youtube: Strategies for using new media in teaching and learning. In *Enhancing learning through technology: research on emerging technologies and pedagogies*. World Scientific, 31–43.
[51] Michele P Dyson, Lisa Hartling, Jocelyn Shulhan, Annabritt Chisholm, Andrea Milne, Purnima Sundar, Shannon D Scott, and Amanda S Newton. 2016. A systematic review of social media use to discuss and view deliberate self-harm acts. *PloS one* 11, 5 (2016), e0155813.
[52] ROGER M Enoka and Douglas G Stuart. 1992. Neurobiology of muscle fatigue. *Journal of applied physiology* 72, 5 (1992), 1631–1648.
[53] Faceswap. 2021. *Faceswap - Faceswap is the leading free and Open Source multi-platform Deepfakes software*. Retrieved August 20, 2021 from https://faceswap.dev/
[54] Matthew Farrow, Christof Lutteroth, Peter C Rouse, and James LJ Bilzon. 2019. Virtual-reality exergaming improves performance during high-intensity interval training. *European journal of sport science* 19, 6 (2019), 719–727.
[55] Forbes. 2019. *Chinese Deepfake App ZAO Goes Viral, Privacy Of Millions 'At Risk'*. Retrieved August 20, 2021 from https://www.forbes.com/sites/zakdoffman/2019/09/02/chinese-best-ever-deepfake-app-zao-sparks-huge-faceapp-like-privacy-storm/
[56] Forbes. 2020. *Why Deepfakes Are a Net Positive for Humanity*. Retrieved August 20, 2021 from https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/
[57] Ian M Franks and LAWRENCE J Maile. 1991. The use of video in sport skill acquisition. *Practical guide to using video in the behavioral sciences* (1991), 231–243.
[58] Simon C Gandevia. 2001. Spinal and supraspinal factors in human muscle fatigue. *Physiological reviews* (2001).
[59] Anne W Garcia and Abby C King. 1991. Predicting long-term adherence to aerobic exercise: A comparison of two models. *Journal of Sport and Exercise Psychology* 13, 4 (1991), 394–410.
[60] Thomas Graven-Nielsen, Peter Svensson, and Lars Arendt-Nielsen. 1997. Effects of experimental muscle pain on muscle activity and co-ordination during static and dynamic motor function. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control* 105, 2 (1997), 156–164.
[61] David Güera and Edward J Delp. 2018. Deepfake video detection using recurrent neural networks. In *2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS)*. IEEE, 1–6.
[62] Grenita Hall, Deepika R Laddu, Shane A Phillips, Carl J Lavie, and Ross Arena. 2021. A tale of two pandemics: How will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another? *Progress in cardiovascular diseases* 64 (2021), 108.
[63] Thomas G Haring, Craig H Kennedy, Mary J Adams, and Valerie Pitts-Conway. 1987. Teaching generalization of purchasing skills across community settings to autistic youth using videotape modeling. *Journal of applied behavior analysis* 20, 1 (1987), 89–96.
[64] Douglas Harris. 2018. Deepfakes: False pornography is here and the law cannot protect you. *Duke L. & Tech. Rev.* 17 (2018), 99.
[65] Haya R Hasan and Khaled Salah. 2019. Combating deepfake videos using blockchain and smart contracts. *Ieee Access* 7 (2019), 41596–41606.
[66] Seongkook Heo, Michelle Annett, Benjamin J Lafreniere, Tovi Grossman, and George W Fitzmaurice. 2017. No Need to Stop What You're Doing: Exploring No-Handed Smartwatch Interaction.. In *Graphics Interface*. 107–114.
[67] 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.
[68] Nicola J Hodges and Ian M Franks. 2002. Modelling coaching practice: the role of instruction and demonstration. *Journal of sports sciences* 20, 10 (2002), 793–811.
[69] Ray E Hosford. 1981. Self-as-a-model: A cognitive social learning technique. *The Counseling Psychologist* 9, 1 (1981), 45–62.
[70] Christos Ioannou, Patrick Archard, Eamonn O'neill, and Christof Lutteroth. 2019. Virtual performance augmentation in an immersive jump & run exergame. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–15.
[71] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1125–1134.
[72] Susan A Jackson, Patrick R Thomas, Herbert W Marsh, and Christopher J Smethurst. 2001. Relationships between flow, self-concept, psychological skills, and performance. *Journal of applied sport psychology* 13, 2 (2001), 129–153.
[73] Gaurav Jaswal, Amit Kaul, and Ravinder Nath. 2016. Knuckle print biometrics and fusion schemes–overview, challenges, and solutions. *ACM Computing Surveys (CSUR)* 49, 2 (2016), 1–46.
[74] Cameron T Jennings, Peter Reaburn, and Steven B Rynne. 2013. The effect of a self-modelling video intervention on motor skill acquisition and retention of a novice track cyclist's standing start performance. *International Journal of Sports Science & Coaching* 8, 3 (2013), 467–480.
[75] Thomas J Kehle, Steven V Owen, and Elizabeth T Cressy. 1990. The use of self-modeling as an intervention in school psychology: A case study of an elective mute. *School Psychology Review* 19, 1 (1990), 115–121.
[76] Lee Kern-Dunlap, Glen Dunlap, Shelley Clarke, Karen E Childs, Ronnie L White, and Mary Pat Stewart. 1992. Effects of a videotape feedback package on the peer interactions of children with serious behavioral and emotional challenges. *Journal of Applied Behavior Analysis* 25, 2 (1992), 355–364.
[77] Juho Kim, Philip J Guo, Carrie J Cai, Shang-Wen Li, Krzysztof Z Gajos, and Robert C Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 563–572.
[78] Marino Kimura and Yuko Yotsumoto. 2018. Auditory traits of" own voice". *PloS one* 13, 6 (2018), e0199443.
[79] 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. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.
[80] Yousef Kowsar, Masud Moshtaghi, Eduardo Velloso, Lars Kulik, and Christopher Leckie. 2016. Detecting unseen anomalies in weight training exercises. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. 517–526.

[81] Victor AF Lamme and Pieter R Roelfsema. 2000. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences* 23, 11 (2000), 571–579.

[82] Patricia G Lange. 2019. Informal learning on YouTube. *The international encyclopedia of media literacy* (2019), 1–11.

[83] Mark R Leary. 2013. Social Anxiety, Shyness, and related constructs. *Measures of personality and social psychological attitudes: Measures of social psychological attitudes* 1 (2013), 161–176.

[84] Yuezun Li and Siwei Lyu. 2018. Exposing deepfake videos by detecting face warping artifacts. *arXiv preprint arXiv:1811.00656* (2018).

[85] Yuezun Li, Xin Yang, Pu Sun, Honggang Qi, and Siwei Lyu. 2020. Celeb-df: A large-scale challenging dataset for deepfake forensics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3207–3216.

[86] Jeff Lieberman and Cynthia Breazeal. 2007. TIKL: Development of a wearable vibrotactile feedback suit for improved human motor learning. *IEEE Transactions on Robotics* 23, 5 (2007), 919–926.

[87] Jonathon Love, Ravi Selker, Maarten Marsman, Tahira Jamil, Damian Dropmann, Josine Verhagen, Alexander Ly, Quentin F Gronau, Martin Šmíra, Sacha Epskamp, et al. 2019. JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software* 88 (2019), 1–17.

[88] Stephen E Lucas. 2009. *The art of public speaking 11th edition*. Boston. Mc. Graw Hill.

[89] Tímea Magyaródi, Henriett Nagy, Péter Soltész, Tamás Mózes, and Attila Oláh. 2013. Psychometric Properties of a Newly Established Flow State Questionnaire. *The Journal of Happiness & Well-Being* 1, 2 (2013), 85–96.

[90] Marie-Helen Maras and Alex Alexandrou. 2019. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence & Proof* 23, 3 (2019), 255–262.

[91] Diogo L Marques, Henrique P Neiva, Luis B Fail, Maria H Gil, and Mario C Marques. 2019. Acute effects of low and high-volume resistance training on hemodynamic, metabolic and neuromuscular parameters in older adults. *Experimental gerontology* 125 (2019), 110685.

[92] Kathleen A Martin, Sandra E Moritz, and Craig R Hall. 1999. Imagery use in sport: A literature review and applied model. *The sport psychologist* 13, 3 (1999), 245–268.

[93] 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.

[94] M McGraw-Hunter, GD Faw, and PK Davis. 2006. The use of video self-modelling and feedback to teach cooking skills to individuals with traumatic brain injury: A pilot study. *Brain Injury* 20, 10 (2006), 1061–1068.

[95] Stephen S Meharg and Lee E Lipsker. 1992. Parent training using videotape self-modeling. *Child & Family Behavior Therapy* 13, 4 (1992), 1–27.

[96] Stephen S Meharg and Mitchel A Woltersdorf. 1990. Therapeutic use of videotape self-modeling: A review. *Advances in Behaviour Research and Therapy* 12, 2 (1990), 85–99.

[97] Edvinas Meskys, Aidas Liaudanskas, Julija Kalpokiene, and Paulius Jurcys. 2020. Regulating deep fakes: legal and ethical considerations. *Journal of Intellectual Property Law & Practice* 15, 1 (2020), 24–31.

[98] Alexander Michael and Christof Lutteroth. 2020. Race Yourselves: A Longitudinal Exploration of Self-Competition Between Past, Present, and Future Performances in a VR Exergame. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.

[99] Microsoft. 2020. *Kinect for Windows*. Retrieved August 20, 2021 from https://developer.microsoft.com/en-us/windows/kinect/

[100] Simon Middlemas and Chris Harwood. 2020. A pre-match video self-modeling intervention in elite youth football. *Journal of Applied Sport Psychology* 32, 5 (2020), 450–475.

[101] Sedigheh Moghavvemi, Ainin Sulaiman, Noor Ismawati Jaafar, and Nafisa Kasem. 2018. Social media as a complementary learning tool for teaching and learning: The case of youtube. *The International Journal of Management Education* 16, 1 (2018), 37–42.

[102] Richard D Morey et al. 2008. Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology* 4, 2 (2008), 61–64.

[103] Masahiro Mori. 1970. On the Uncanny Valley. *Energy* 7, 4 (1970). https://ci.nii.ac.jp/naid/10027463083/

[104] Solène Neyret, Xavi Navarro, Alejandro Beacco, Ramon Oliva, Pierre Bourdin, Jose Valenzuela, Itxaso Barberia, and Mel Slater. 2020. An embodied perspective as a victim of sexual harassment in virtual reality reduces action conformity in a later milgram obedience scenario. *Scientific reports* 10, 1 (2020), 1–18.

[105] Nintendo. 2020. *Wii Fit*. Retrieved August 20, 2021 from https://www.nintendo.co.uk/Games/Wii/Wii-Fit-283894.html

[106] The University of Vermont. [n.d.]. *Public Speaking Evaluation/Criteria*. https://www.uvm.edu/sites/default/files/media/eval-public-speaking.pdf

[107] Carl Öhman. 2020. Introducing the pervert's dilemma: a contribution to the critique of Deepfake Pornography. *Ethics and Information Technology* 22, 2 (2020), 133–140.

[108] World Health Organization. 2020. *Physical activity*. Retrieved August 20, 2021 from https://www.who.int/news-room/fact-sheets/detail/physical-activity

[109] Armin H Paravlic, Maamer Slimani, David Tod, Uros Marusic, Zoran Milanovic, and Rado Pisot. 2018. Effects and dose–response relationships of motor imagery practice on strength development in healthy adult populations: a systematic review and meta-analysis. *Sports Medicine* 48, 5 (2018), 1165–1187.

[110] Gordon L Paul. 1966. *Insight vs. desensitization in psychotherapy: An experiment in anxiety reduction*. Stanford University Press.

[111] Ivan Perov, Daiheng Gao, Nikolay Chervoniy, Kunlin Liu, Sugasa Marangonda, Chris Umé, Mr Dpfks, Carl Shift Facenheim, Luis RP, Jian Jiang, et al. 2020. DeepFaceLab: Integrated, flexible and extensible face-swapping framework. *arXiv preprint arXiv:2005.05535* (2020).

[112] Timothy R Pineau, Carol R Glass, Keith A Kaufman, and Darren R Bernal. 2014. Self-and team-efficacy beliefs of rowers and their relation to mindfulness and flow. *Journal of Clinical Sport Psychology* 8, 2 (2014), 142–158.

[113] Jan L Plass, Bruce D Homer, and Charles K Kinzer. 2014. Playful learning: An integrated design framework. *White paper* 2 (2014), 2014.

[114] 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.

[115] Mitchell J Prinstein, Nicole Heilbron, John D Guerry, Joseph C Franklin, Diana Rancourt, Valerie Simon, and Anthony Spirito. 2010. Peer influence and non-suicidal self injury: Longitudinal results in community and clinically-referred adolescent samples. *Journal of abnormal child psychology* 38, 5 (2010), 669–682.

[116] Charles B Pull. 2012. Current status of knowledge on public-speaking anxiety. *Current opinion in psychiatry* 25, 1 (2012), 32–38.

[117] Helen Regan and Julia Howe. 2017. Video self-modelling: an intervention for children with behavioural difficulties. *Educational Psychology in Practice* 33, 1 (2017), 93–102.

[118] Jennifer R Richardson and Timothy D Lee. 1999. The effects of proactive and retroactive demonstrations on learning signed letters. *Acta Psychologica* 101, 1 (1999), 79–90.

[119] 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.

[120] Richard M Ryan. 1982. Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. *Journal of Personality and Social Psychology* 43, 3 (1982), 450–461.

[121] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.

[122] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[123] Dale H Schunk. 1987. Peer models and children's behavioral change. *Review of educational research* 57, 2 (1987), 149–174.

[124] Dale H Schunk and Antoinette R Hanson. 1989. Influence of peer-model attributes on children's beliefs and learning. *Journal of Educational Psychology* 81, 3 (1989), 431.

[125] Dale H Schunk and Antoinette R Hanson. 1989. Self-modeling and children's cognitive skill learning. *Journal of Educational Psychology* 81, 2 (1989), 155.

[126] Dale H Schunk, Antoinette R Hanson, and Paula D Cox. 1987. Peer-model attributes and children's achievement behaviors. *Journal of educational psychology* 79, 1 (1987), 54.

[127] Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017. " 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*. 1577–1582.

[128] Dale E Seborg, Thomas F Edgar, Duncan A Mellichamp, and Francis J Doyle III. 2016. *Process dynamics and control*. John Wiley & Sons.

[129] Ben Sidaway and M Jerome Hand. 1993. Frequency of modeling effects on the acquisition and retention of a motor skill. *Research quarterly for exercise and sport* 64, 1 (1993), 122–126.

[130] 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.

[131] 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.

[132] Diane M Ste-Marie, Kelly Vertes, Amanda M Rymal, and Rose Martini. 2011. Feedforward self-modeling enhances skill acquisition in children learning trampoline skills. *Frontiers in psychology* 2 (2011), 155.

[133] Kylie A Steel, Roger D Adams, Susan E Coulson, Colleen G Canning, and Holly J Hawtin. 2013. Video self-model training of punt kicking. *International Journal of Sport and Health Science* (2013), 201233.

[134] Cassidy R Sugimoto, Mike Thelwall, Vincent Larivière, Andrew Tsou, Philippe Mongeon, and Benoit Macaluso. 2013. Scientists popularizing science: characteristics and impact of TED talk presenters. *PloS one* 8, 4 (2013), e62403.

[135] Supasorn Suwajanakorn, Steven M Seitz, and Ira Kemelmacher-Shlizerman. 2017. Synthesizing obama: learning lip sync from audio. *ACM Transactions on Graphics (ToG)* 36, 4 (2017), 1–13.

[136] Elaine Tan. 2013. Informal learning on YouTube: Exploring digital literacy in independent online learning. *Learning, Media and Technology* 38, 4 (2013), 463–477.

[137] Richard Tang, Xing-Dong Yang, Scott Bateman, Joaquim Jorge, and Anthony Tang. 2015. Physio@ Home: Exploring visual guidance and feedback techniques for physiotherapy exercises. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 4123–4132.

[138] Justus Thies, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. 2016. Face2face: Real-time face capture and reenactment of rgb videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2387–2395.

[139] Milka Trajkova and Mexhid Ferati. 2015. Usability evaluation of kinect-based system for ballet movements. In *International Conference of Design, User Experience, and Usability*. Springer, 464–472.

[140] Cristian Vaccari and Andrew Chadwick. 2020. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society* 6, 1 (2020), 2056305120903408.

[141] Jan Van Looy, Cédric Courtois, Melanie De Vocht, and Lieven De Marez. 2012. Player identification in online games: Validation of a scale for measuring identification in MMOGs. *Media Psychology* 15, 2 (2012), 197–221.

[142] Kelly A Vertes and Diane M Ste-Marie. 2013. Trampolinists' self-controlled use of a feedforward self-modeling video in competition. *Journal of Applied Sport Psychology* 25, 4 (2013), 463–477.

[143] Maureen R Weiss, Vicki Ebbeck, and Diane M Wiese-Bjornstal. 1993. Developmental and psychological factors related to children's observational learning of physical skills. *Pediatric Exercise Science* 5, 4 (1993), 301–317.

[144] Maureen R Weiss, Penny McCullagh, Alan L Smith, and Anthony R Berlant. 1998. Observational learning and the fearful child: Influence of peer models on swimming skill performance and psychological responses. *Research Quarterly for Exercise and Sport* 69, 4 (1998), 380–394.

[145] Barbara Yingling Wert and John T Neisworth. 2003. Effects of video self-modeling on spontaneous requesting in children with autism. *Journal of Positive Behavior Interventions* 5, 1 (2003), 30–34.

[146] HTA Whiting, MJ Bijlard, and BPLM Den Brinker. 1987. The effect of the availability of a dynamic model on the acquisition of a complex cyclical action. *The Quarterly Journal of Experimental Psychology Section A* 39, 1 (1987), 43–59.

[147] Sara Wilcox, Patricia A Sharpe, Brent Hutto, and Michelle L Granner. 2005. Psychometric properties of the self-efficacy for exercise questionnaire in a diverse sample of men and women. *Journal of Physical Activity and Health* 2, 3 (2005), 285–297.

[148] Carolee J Winstein and Richard A Schmidt. 1990. Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16, 4 (1990), 677.

[149] Wired. 2021. *The messy legal scrap to bring celebrities back from the dead.* Retrieved August 20, 2021 from https://www.wired.co.uk/article/james-dean-dead-actors-rights

[150] Ungyeon Yang and Gerard Jounghyun Kim. 2002. Implementation and evaluation of "just follow me": An immersive, VR-based, motion-training system. *Presence: Teleoperators & Virtual Environments* 11, 3 (2002), 304–323.

[151] Cheryl A Young-Pelton and Samantha L Bushman. 2015. Using video self-modelling to increase active learning responses during small-group reading instruction for primary school pupils with social emotional and mental health difficulties. *Emotional and Behavioural Difficulties* 20, 3 (2015), 277–288.

[152] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*. 2223–2232.