

The Effect of an Adaptive Simulated Inner Voice on User's Eye-gaze Behaviour, Ownership Perception and Plausibility Judgement in Virtual Reality

Ding, Ding; Neerincx, Mark A.; Brinkman, Willem Paul

DOI

[10.1093/iwcomp/iwab008](https://doi.org/10.1093/iwcomp/iwab008)

Publication date

2020

Document Version

Final published version

Published in

Interacting with Computers

Citation (APA)

Ding, D., Neerincx, M. A., & Brinkman, W. P. (2020). The Effect of an Adaptive Simulated Inner Voice on User's Eye-gaze Behaviour, Ownership Perception and Plausibility Judgement in Virtual Reality. *Interacting with Computers*, 32(5-6), 510-523. <https://doi.org/10.1093/iwcomp/iwab008>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

The Effect of an Adaptive Simulated Inner Voice on User's Eye-gaze Behaviour, Ownership Perception and Plausibility Judgement in Virtual Reality

DING DING^{1,2,*}, MARK A. NEERINCX^{2,3} AND WILLEM-PAUL BRINKMAN²

¹*Southeast University, School of Computer Science and Engineering, Nanjing, 210018, China*

²*Delft University of Technology, Department of Intelligent Systems, Delft, 2628 XE, The Netherlands*

³*TNO, Perceptual and Cognitive Systems, Soesterberg, 3769 DE, The Netherlands*

*Corresponding author: dingding-1@seu.edu.cn

Virtual cognitions (VCs) are a stream of simulated thoughts people hear while emerged in a virtual environment, e.g. by hearing a simulated inner voice presented as a voice over. They can enhance people's self-efficacy and knowledge about, for example, social interactions as previous studies have shown. Ownership and plausibility of these VCs are regarded as important for their effect, and enhancing both might, therefore, be beneficial. A potential strategy for achieving this is the synchronization of the VCs with people's eye fixation using eye-tracking technology embedded in a head-mounted display. Hence, this paper tests this idea in the context of a pre-therapy for spider and snake phobia to examine the ability to guide people's eye fixation. An experiment with 24 participants was conducted using a within-subjects design. Each participant was exposed to two conditions: one where the VCs were adapted to eye gaze of the participant and the other where they were not adapted, i.e. the control condition. The findings of a Bayesian analysis suggest that credibly more ownership was reported and more eye-gaze shift behaviour was observed in the eye-gaze-adapted condition than in the control condition. Compared to the alternative of no or negative mediation, the findings also give some more credibility to the hypothesis that ownership, at least partly, positively mediates the effect eye-gaze-adapted VCs have on eye-gaze shift behaviour. Only weak support was found for plausibility as a mediator. These findings help improve insight into how VCs affect people.

RESEARCH HIGHLIGHTS

- Eye-gaze-adaptive VCs have a positive effect on people's perceived ownership of the VCs.
- Eye-gaze-adaptive VCs have a positive effect on the numbers of eye-gaze shifts after hearing an instructional VC.
- The effect of eye-gaze-adaptive VCs on user's eye-gaze behaviour is partially mediated by the user's ownership of VCs.

Keywords: eye tracking; eye-gaze adaptive; virtual cognitions; virtual reality exposure; spider phobia; snake phobia

Handling Editor: Rod McCall

Received 30 April 2020; Revised 30 December 2020; Accepted 26 February 2021

1. INTRODUCTION

Imagine that you are immersed in a virtual reality (VR) environment from a first-person perspective. You are sitting in a virtual room. In front of you, a virtual spider crawls over the floor. You hear a voice-over that aims to mimic your inner voice. It speaks about the spider as part of an internal monologue of thoughts. Would these thoughts you hear change your beliefs about spiders? What if the content of these thoughts refer to what your eyes are fixated on, so instead of the spider, the couch? Would this content adaptation make these thoughts feel more plausible and even feel more like your own thoughts? Besides, would it make these thoughts, or in other words virtual cognitions (VCs), more effective in influencing your behaviour in VR? This paper presents a study that examines this idea to establish more understanding of how VCs work, which can help to make them more effective as a persuasive strategy for behaviour change.

VCs are an extension of the more classical VR experience. In the past decades, research on VR focused on the development and improvement of the virtual environment, such as the realism of virtual objects (Hoffman, 1998), the fidelity of the lighting model (Slater *et al.*, 2009), the virtual depth-of-field blur (Hillaire *et al.*, 2008) and the spatialization of the audio-rendering model (Naef *et al.*, 2002, Zotkin *et al.*, 2004). Besides the external environment surrounding the user in virtual reality, recently the virtual representation of the user in the virtual environment, i.e. the virtual body, has gained more attention. For example, creating a photo-realistic rendering of users' own physical image into VR (Achenbach *et al.*, 2017, Feng *et al.*, 2017), or the synchronization of virtual body movement with user's actual physical movement (Banakou *et al.*, 2013, Slater & Sanchez-Vives, 2014). The latter is often studied in the context of the virtual body illusion or body transfer (Maselli & Slater, 2013, Slater *et al.*, 2010). Here, users experience that virtual body as their own body. This illusion is regularly enhanced by using real-time mirror reflections of motor action (De la Peña *et al.*, 2010, Gonzalez-Franco *et al.*, 2010, Normand *et al.*, 2011) in a virtual mirror. Where virtual bodies are used to give users the experience of another body, VCs are used to give users the experience of another cognition, i.e. thoughts.

VCs are a stream of simulated thoughts people perceive while emerged in a virtual environment, e.g. by hearing a simulated *inner voice*. They can provide people with information about a topic or reflections on the current situation or offer motivational encouragement. The inner voice, regarded as 'verbal sets, instructions to oneself, or a verbal interpretation of sensation and perceptions' (Sokolov, 2012), plays a crucial role in conscious thoughts. They are believed to serve various cognitive functions such as self-awareness (Morin & Everett, 1990), self-reflection (Morin & Hamper, 2012) and, importantly, self-control (Tullett & Inzlicht, 2010). People tend to follow the instruction from their inner voice (Morin,

2009, Tullett & Inzlicht, 2010). With the help of a simulated inner voice, individuals might be able to perceive not only the external virtual environment, the behaviour and the events via an avatar, but also the internal thoughts. Together, VCs might work as a new medium to change people's beliefs and behavior.

VR, in the form of therapy (Anderson *et al.*, 2013, Bouchard *et al.*, 2017) or training (Broekens *et al.*, 2012), can contribute to change in people's beliefs and behaviour. Cognitive behavioral therapy (CBT) is an effective psychotherapeutic approach in treating phobias (Paquette *et al.*, 2003, Schienle *et al.*, 2009). It intends to help people with anxiety disorder gain control over their dysfunctional cognitions and behaviors via gradual exposure to the anxiety-provoking stimuli. Exposure in VR has also extensively been studied. A recent meta-analysis (Carl *et al.*, 2019) of 30 controlled trials of VR exposure therapy for anxiety-related disorders concluded that this therapy is an effective and equal medium for exposure therapy. Typically in such therapies, patients are exposed to objects or situations they fear. This can range from seeing a virtual spider (Garcia-Palacios *et al.*, 2002), snake (Bouchard *et al.*, 2008) or audience (Pertaud *et al.*, 2002), to situations, such as, closely confined space (Botella *et al.*, 2000), standing on a high building (Emmelkamp *et al.*, 2002), sitting in an aeroplane (Rothbaum *et al.*, 2000) or talking with a virtual character (Grillon *et al.*, 2006). VCs have the promise to extend these exposures by directly advancing users' cognitive understanding of the situation. The first support for this tenet can be found in the work of Kang *et al.* (2019). The authors had users passively experience giving a public talk in VR. Instead of actually asking people to give a physical talk in front of a virtual audience, the authors had them listening to a pre-recorded presentation from a first-person perspective. The authors found that how well users identified with the virtual person coincided with how well they changed their self-efficacy on giving this public talk in real life. The second support comes from the work of Ding *et al.* (2020). They used VCs to train people on negotiation. By comparing the trainees with an inactive control group, they found that trainees had improved their self-efficacy, knowledge about negotiation and their negotiation satisfaction and result. As these effects were still observed multiple weeks after the training, it warrants further research into understanding how VCs work and how its effect can be extended.

Although some studies (Krijn *et al.*, 2004, Price & Anderson, 2007) failed to find a direct relation between the sense of presence in the virtual world and the outcome of a VR therapy, others (Price *et al.*, 2011) found some support for such association. Still, some indirect support exists. For example, Meuret *et al.* (2012) found that fear experienced during exposure is a weak predictor for therapy outcome. Also, Krijn *et al.* (2004) reported that compared to patients that finished the therapy successfully, dropouts reported experiencing less presence and less anxiety in the first half of the first session. Next, a meta-analysis (Ling *et al.*, 2014) of 33 studies found a medium effect size for the correlation between the sense of presence and self-reported

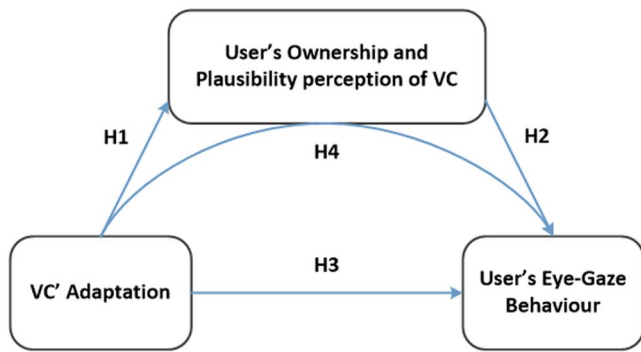


FIGURE 1. Hypothesized relationship between adaptation, ownership, plausibility and eye-gaze behavior.

anxiety during VR exposure. Understanding factors that can enhance presence, therefore, could be of interest. One such factor is the alignment of sensory cues in the virtual environment. For example, does what you see match with what your hear or smell? Access to a multitude of these cues enhances people's sense of presence in VR (Davis *et al.*, 1999, Dinh *et al.*, 1999). Once exposed in VR, a VC is a cue besides the very dominant visual sensory cue. Therefore, to adapt VCs to user's viewing behaviour, a system would need to know where the users' eyes are fixated on. Eye-tracking technology could provide this information. Whereas eye movement conveys information about people's emotional and mental states (Baron-Cohen *et al.*, 2001), it also gives some insight into people's attention because of its functional relation with eye movement (Corbetta, 1998). More specifically, an eye movement toward a location could induce a concurrent shift of attention toward that same location. Also, what people perceive triggers and shapes their thoughts (Allen *et al.*, 2013, Cho *et al.*, 2014, Stroop, 1935). For example, a room with or without a picture of John F. Kennedy influences whether people will mention this US president when asked to name a world famous politician (Qu *et al.*, 2013). This puts forward the concept of *adaptive VCs*, i.e. VCs with content that matches the object of people's eye fixation. People should find them more plausible as they seem to be triggered by the things in the virtual environment.

2. THEORY AND HYPOTHESES

Figure 1 illustrates the key concepts in the relationship between adaptive VCs and eye-gaze behaviour, which is explained by the extent people perceive them as plausible and as their own. Changes in eye-gaze behavior become especially noticeable when the VCs include instructions where to focus on. In fact, anyone who has taken a tour at a museum with a handheld audio guide device has probably experienced this. When the recorded commentary spoke about a specific physical aspect of the artefact, you might have found yourself inclined to look at it.

First of all, Fig. 1 shows the direct effect between adaptive VCs and people's eye-gaze behaviour (H3). Although providing VCs could bring benefits for users, it also requires one more source to attend, thus brings new challenges for users to deal with selective attention and divided attention. Selective attention is the ability to select from many factors or stimuli and to focus cognitive resources on information that is relevant to our goals or tasks (Gazzaley & Nobre, 2012). On the other hand, divided attention is the ability to integrate simultaneously multiple stimuli and allocate cognitive resources between different sets of input (Hahn *et al.*, 2008, Iacoboni, 2005). Suppose stimuli perceived in a virtual environment are not congruent, in that case, people might be forced to select and divide attention across multiple information sources, causing additional cognitive load that interferes with their experience (Wickens *et al.*, 2015). Whether people select, focus and process the content of VCs determines whether and to what extent they affect people's behaviour. If VCs are coherent with other stimuli, people might be able to pay more attention to their content and, consequently, follow up instructions embedded in VCs, i.e. instructional VC.

Next, Fig. 1 hypothesized that the direct effect is mediated by the sense of *ownership and plausibility* of the VCs (H4). Plausibility illusion is one of two factors that affect how people experience presence in VR (Slater, 2009). This illusion is about what is perceived as really happening. When it comes to VC, it refers to the illusion that the VC is being perceived as an actual cognition, i.e. thoughts in someone's own head. The other factor is place illusion, referring to the sense of being there (Slater, 2009). This factor seems less relevant to VCs as they are less about being in a specific place but more about the person itself. Instead, the concept of ownership is more suitable. It refers to 'the sense that I am the one who is undergoing an experience' (Gallagher, 2000). The sense of ownership is closely linked to the sense of agency, which received attention because of the work on the virtual body illusion (Cole *et al.*, 2009, Kilteni *et al.*, 2012, Nowak & Biocca, 2003, Perez-Marcos *et al.*, 2009), i.e. a virtual body that people experience as their own. The sense of agency refers to the feeling of control over one's actions (David *et al.*, 2008). For voluntary actions, the sense of ownership and sense of agency coincide but are distinct when the sensory experience is passive and involuntary. For example, when some external force passively moves one's arm, the sense of ownership can be still experienced but not the sense of agency (Gallagher, 2000, Shimada *et al.*, 2005). These two senses are also regarded as the constitutive components of the sense of embodiment (Giummarra *et al.*, 2008), which is also a key concept in VR. It refers to the sensations that emerge with being inside, having and controlling a virtual body in a virtual environment (Blanke & Metzinger, 2009, Kilteni *et al.*, 2012). Kilteni *et al.* (2012) indicate the sense of embodiment is grounded on the achievement of these two senses. However, little is known about the individual contribution of these two components to the sense of embodiment. When it comes to thoughts, it is hard to say that a person is in direct conscious

control of them; however, people would experience them as their own. In fact, when they do not feel this, people might experience what is known as thought insertion, a common symptom of psychosis and associated with schizophrenia, where people feel their thoughts are not one's own but being inserted into their mind (Bortolotti & Broome, 2009, Mullins & Spence, 2003). Framing the effect of inner voice on eye-gaze behaviour as a simple stimulus-response format helps in understanding why plausibility and ownership can mediate this effect. VCs are artificial stimuli. The more they resemble the actual stimuli, the more they can elicit the related response. Hence, plausibility and ownership are properties of this resemblance. When they are high, the resemblance is also high. In that case, people would be more inclined to follow instructions aimed to guide eye gaze embedded in VCs (Fig. 1; H2).

Adapting the VCs to the focus of attention (e.g. to user's eye gaze) is expected to bring about a higher sense of plausibility and ownership of the VC (Fig. 1; H1). Providing multiple types of sensory stimuli, such as visuals, audios and haptics, has the potential to enhance the sense of presence (Dinh *et al.*, 1999, Larsson *et al.*, 2007) and task performance. However, this only happens when the stimuli are consistent, as has been observed for visual cues and haptic stimuli (Carlin *et al.*, 1997) and for visual and auditory stimuli in virtual environments (Larsson *et al.*, 2001, 2007, Lindquist *et al.*, 2016). In contrast, when inconsistent, presence, memory recall, and task performance can deteriorate (Larsson *et al.*, 2001). Therefore, adjusting the VCs to match the visual stimuli has the potential of making the VCs aligned with a person's focus of attention and making them more plausible. Moreover, according to Shimada *et al.* (2005), the sense of ownership can be achieved by the integration of multiple sensory cues, such as the synchrony of visual and proprioceptive cues, or as Slater (2009) observed proprioception and visual exteroception. Hence, the adaptation has the potential of increasing the sense of ownership of the VCs.

Based on the above, four key hypotheses can be formulated as follows:

H1. Compared to non-adaptive, eye-gaze-adaptive VCs have a positive effect on people's perceived ownership (1-a) and plausibility (1-b) of the VCs.

H2. An individual's perceived plausibility and ownership of the VCs are positively associated with the number of eye-gaze shifts after hearing an instructional VC.

H3. Compared to non-adaptive, eye-gaze-adaptive VCs have a positive effect on the number of eye-gaze shifts after hearing an instructional VC.

H4. The effect of eye-gaze-adaptive VCs on the number of eye-gaze shifts after hearing an instructional VC is at least partially mediated by people's perceived ownership (4-a) and plausibility (4-b) of the VCs.

The remaining part of the paper presents an experiment to test these four hypotheses in the context of a pre-therapy for spider and snake phobia. But first, the paper describes the system we developed for this experiment.

3. SYSTEM

VR environments are being studied and applied in a broad set of application domains, such as for treatment (Howard, 2017, Valmaggia *et al.*, 2016), education (Freina & Ott, 2015, Jensen & Konradsen, 2018) and entertainment (De la Peña *et al.*, 2010, Yung & Khoo-Lattimore, 2019). One of its core benefits is the ability to experience real-world scenarios in a safe, controlled and engaging environment. People can transfer the VR experiences to their real world. For example, in VR therapy for anxiety disorders, people are exposed to virtual scenarios that resemble scenarios they fear in real life. After these exposures, they are more capable of coping with these real-life situations (Morina *et al.*, 2015). VR therapy is a cognitive behavioural therapy. The current emphasis, however, is mainly on the behavioural component. Extending the virtual exposure with VCs would allow addressing the cognitive component of CBT. Therefore, to test the four hypotheses in a potentially beneficial context, a VR system was developed that could precede a VR therapy for spider and snake phobia aiming to improve people's self-efficacy of coping with the presence of the feared animal. Being able to do that can be helpful, as Böhnlein *et al.* (2020) identify self-efficacy as predictor for the success of exposure therapy, including for VR (Côté & Bouchard, 2009).

Following the concept of VR exposure therapy, the following *usage scenario* was constructed. The exposure session consisted of three scenes. In the first scene, users sat alone at a fixed position in the virtual room, experiencing the virtual environment from a first-person perspective without having a virtual body in the virtual environment. Moreover, the system had no movement capturing and synchronization. Users heard VCs reflecting on the upcoming appearance of the feared animal, presenting some facts and information about these animals. Self-motivational cognitions accompanied these reflections. The second scene consisted of actual exposure to the feared animal. The feared animal moved slowly out from a sofa in the corner to the centre of the room (Fig. 2). Users could explore the virtual room and look at the animal by moving their head while being exposed to the VCs. Finally, the animal crawls out of the room and disappears from the users' sight. In the last scene, users were again alone in the virtual room like the first scene. They heard the VCs that reviewed the past exposure, their performance and also encouraged them affirmatively.

For the usage scenario, we designed VCs as follows. Users could hear three types of VCs, as shown in Table 1: (i) a factual statement about the feared animal, such as 'A spider is just a type of animal people are not so familiar with, but it does not mean you should feel afraid of it.'; (ii) self-motivational statements of handling the fear-animal, e.g. 'You are strong enough to handle that.'; and (iii) instructions to encourage specific-gaze behaviour, e.g. 'Try to look at the spider a bit longer.' Each type of VCs has two versions based on the



FIGURE 2. The screen shot of VR exposure from the perspective of users: spider condition (above) and snake condition (bottom).

condition: eye-gaze-adaptive VCs and non-eye-gaze-adaptive VCs. In the first and last scene, users only hear the generic VCs, i.e. non-eye-gaze-adaptive VCs, which include the first two types of VCs: factual and self-motivational VCs. Only in the second scene, which included the feared animal, did the system use adaptive cognitions that include all three types of VCs. Once immersed in a virtual room, the system used eye-tracking technology embedded in a head-mounted display (HMD) to register whether or not users were fixing their eyes on a virtual representation of the feared animal. Consequently, the system could trigger an animal-related VC.

For testing the hypotheses, two versions of the systems were created: one that had eye-gaze-adaptive VCs and one that only had generic VCs. In the eye-gaze-adaptive version, the VCs that users heard depended on where they were looking at. If users were looking at the spider, the VCs acknowledged that they were looking at it, described its movement and encouraged them to continue looking at it. In the non-adaptive version, no acknowledgment was included about whether they were looking at the spider or not. As the examples show in Table 1, the eye-gaze-adaptive version used a formulation that assumed a specific eye-gaze fixation. For this, the system used a minimum fixation threshold of at least 1 second. On the other hand, the generic version used a formulation that did not make assumptions about eye fixation.

TABLE 1. The number of times (N) the system used three types of VCs.

Type	State	Adaptive	Non-adaptive	N
Factual statements about the feared animal	When participants are looking at the feared animal	To some extent, the snake you are looking at seems ugly, but it's ok, you can face it.	To some extent, a snake seems ugly, but it's ok, you can face it.	8
	When participants are not looking at the feared animal	Avoid watching the snake, like what you are doing now, is not helping, it is avoidance.	Avoid watching a snake is not helping, it is avoidance.	13
Instructions to encourage specific gaze behaviour	Instruction of looking back to the feared animal	It's fine to feel a little bit fear but it can be overcome. Trust yourself. You are braver than you believe. Turn to look at the snake. Just do it.	It's fine to feel a little bit fear but it can be overcome. Trust yourself. You are braver than you believe. Take a chance to look at the snake is a cleverer way.	5
	Instruction of looking away from the feared animal	Do not keep an eye on the snake all the time like now! Just do what you want to do.	It's not necessary to keep an eye on the snake all the time, just do what you want to do.	5
Self-motivational statements of handling the fear animal		You are so brave that you can look at the snake like now.	You are so brave that you can look at a snake.	5

The instructions in the VCs also varied to address how people cope with animals they fear and therefore regulated their emotion. A deficit in this regulation has been suggested (Hermann *et al.*, 2009) as a cause for strong emotional reactions towards feared objects. Emotion regulatory strategies such as implementation intention, e.g. ‘if I see a weapon, then I will stay calm and relaxed’, can help in reducing the intensity of these emotional responses (Azbel-Jackson *et al.*, 2016). Instead of facing the feared object, people can also use attentional deployment (e.g. gaze fixation) as a strategy to regulate their emotion. For example, when they want to decrease an intense fear response towards a feared object, they can look away or look less at these emotion-relevant objects (Reekum *et al.*, 2007, Xing & Isaacowitz, 2006). People with an anxiety disorder, such as an intense spider fear, can use this strategy as well as a seemingly opposite strategy. For example, they might avoid looking at the animal (‘blunters’) or excessively look at the animal (‘monitors’) (Antony *et al.*, 2001, Pflugshaupt *et al.*, 2007). Therefore, at the start of the second scene, the instructional VC encouraged users to look at the animal, then later on in the scene, instructional VCs encouraged users to look away from the animal.

In the second scene of the exposure, the system used five sets of VCs to randomly sample, without replacement (Table 1). It used a set of five self-motivational statements. These statements always directly followed after an instructional VCs. Therefore, the system used similar set sizes for sampling an instructional VC, five for the eyes fixed on the animal state and five for the other state. A pilot study empirically determined the sizes of the factual statement sets. The pilot study showed that the four pilot participants (1 female, aged from 25 to 38, $M = 30.25$, $SD = 5.56$) triggered 13 VCs on average when their eyes were fixed on the animal and 18 when their eyes were not. Consequently, the system used for factual statements a set size of 8 ($13 - 5 = 8$) for the eyes fixed on the animal state and a set size of 13 ($18 - 5 = 13$) for the other state.

4. METHOD

To test the hypotheses, an empirical experiment with 24 participants was conducted. The experiment had a within-subject design, where all participants were exposed both the eye-gaze- and non-eye-gaze-adaptive VCs conditions. The order of the two conditions was randomized. Each participant was exposed to another animal in each condition. This reduced the risk of a potential carry-over effect associated with the experiences obtained with a specific animal. For each participant, the two animals were randomly assigned to the two conditions and the sequence of the animals displayed was also randomized. The study was approved by the university’s human research ethics committee (ID: 577) and pre-registered prior to accessing the data.

4.1. Participants

A total of 27 participants (15 males, 12 females) were recruited throughout the university campus via e-mail or personal approach. Three participants dropped off during the experiment owing to some personal arrangement reasons. The ages of the remaining 24 participants (14 males, 10 females) ranged from 22 to 36 ($M = 27.83$, $SD = 3.2$). They all spoke fluent Mandarin. Two participants had a bachelor’s degree; 16, a master’s degree; and 6, a Ph.D. degree. 29.1% of participants (7/24) had experienced VR before, the remaining had not.

Although the system was designed with the idea of a pre-treatment of spider or snake phobia, because of ethical reasons, participants with a high level of spider or snake fear were excluded from the study. Exclusion criteria were set to 23.76 and 24.44 for the spider questionnaire (SPQ) and the snake questionnaire (SNAQ) (Fredrikson, 1983), respectively. Fredrikson (1983) reported these as the mean score for phobia participants in their study. None of the participants crossed this threshold (SPQ range [2.0, 18.0], $M = 10.0$, $SD = 5.26$; SNAQ range [4.0, 22.0], $M = 12.75$, $SD = 6.24$).

4.2. Materials and measures

4.2.1. Primary measures

Plausibility of the VCs. To examine how plausible people found the VCs, we designed a 3-item plausibility questionnaire based on Rovira *et al.* (2009) who put forward three key features of a plausibility illusion: (1) *correlational*, the elements of the virtual environment react to the the actions or behaviour of the user; (2) *credibility*, the elements of the virtual environment or the events in the virtual environment are credible considering the physical reality and user’s prior knowledge and experience; and (3) *self-reference*, the elements of the virtual environment refer directly to the user. All the items were rated on a scale from -3 (very bad) to 3 (very good).

Sense of ownership of the VCs. To investigate the sense of ownership of the VCs, a dedicated questionnaire was created that was inspired by several existing questionnaires (Kalckert & Ehrsson, 2012, Olckers, 2013). All 5 items were rated on a scale from -3 (totally disagree) to 3 (totally agree).

Number of instructions followed up by eye-gaze shifts. The study used as a behavioural measure, the number of instructional VCs that upon hearing them resulted in people to shift their eye gaze to the target mentioned in the instruction. Take, for example, the situation that participants had their eyes fixed on a spider. Then, they heard an instruction to shift their eye fixation away. If they complied to this instruction by fixating their eyes not on the spider for at least 1 second, this instruction were counted as been followed up. Similarly, the instructional VC counted also in the opposite situation. For example, participants, who initially did not look at the spider, did so for at least 1 second, after hearing to look at the spider. Following Hermans

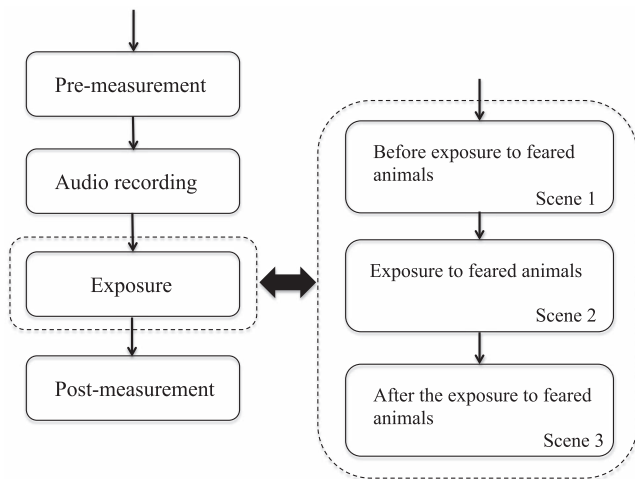


FIGURE 3. Experiment procedure.

et al. (1999), only instructional VCs were counted when eye fixation shift occurred within 3 seconds after hearing them.

The questionnaires used in this study are available online¹.

4.2.2. Secondary measures

Self-efficacy of handling spiders and snakes. Following the study of Bouchard *et al.* (2006), a one-item perceived self-efficacy measure was used. The question was formulated as: ‘On a scale of -5 to 5, to what extent do you feel that you can face a situation where you are in the presence of one or many spider(s)?’ Or snakes in the case of the snake exposure. The item was rated on a scale from -5 (highly certain cannot do) to 5 (highly certain can do).

4.3. Procedure and apparatus

As shown in Fig. 3, the study consisted of four phases: pre-measurement, audio recording, exposure and post-measurement. In the first phase, participants signed an informed consent form and completed self-efficacy questionnaire, fear with handling spiders (SPQ) and snakes (SNAQ) questionnaire and form collecting biographical data (e.g. age, gender, education, previous virtual experience). After this, participants proceeded to the audio recording phase. Here participants read out the complete set of 206 VCs, which were presented in a random order. They were recorded with a pair of in-ear binaural microphones (SP-TFB-2 Sound Professionals). VCs were written in Mandarin to match the participants’ mother tongue.

Ding *et al.*’s study about simulating the inner voice found that people’s sound perception differs between inner and outer

voice on several sound parameters (Ding *et al.* 2020). Furthermore, they found these difference to vary between individuals. Therefore, a similar procedure followed (Ding *et al.*, 2020). After recording, participants were asked to listen to a part of their voice recordings and instructed to set seven sound parameters such as pitch, speed and echo to make the recordings sound as their inner voice. These sound parameter settings were later used to create the VCs that participants heard in the exposure phase.

Two measures were in place that reduce potential carry-over effects from the recording to the exposure. First, the order of the individual VCs was randomized, both during recording and exposure. Second, at least one-day interval was taken between recording and exposure. In the exposure phase, participants, in both conditions, wore the FOVE 0 HMD, which is equipped with infrared eye-tracking technology. The specification provided by its manufacturer FOVE (2019) mentions a tracking accuracy of 1.15 degree median and frame rate of 120 fps. For the display resolution it states 2560×1440 pixels, with a field of view up to 100 degrees and a frame rate of 70fps. In addition, participants wore the same pair of in-ear binaural microphones to hear the VCs. Once participants wear the headset, they were asked to follow the instructions of the standard FOVE 0 HMD calibration procedure, which is displayed on the screen inside the headset. After the calibration, participants were exposed to the virtual environment.

The virtual room was created in Unity3D. As mentioned before, the exposure considered three scenes—before, during and after exposure to the feared animals (see in Fig. 3)—and only the second scene included the animal. In the eye-gaze-adaptive condition, participants were exposed to eye-gaze-adapted VCs, while in the non—eye-gaze-adaptive condition to randomly selected generic VC. After the third scene, participants completed the questionnaires about ownership and plausibility of the VCs, after which they again went through three VR scenes, but this time for the other experimental condition. During the exposure, the number of instructional VCs followed up by eye-gaze shift were counted automatically. The experiment ended with the post-measurement phase, where participants were asked to complete a questionnaire that assesses the plausibility and sense of ownership of the VCs and, again, their self-efficacy of handling spiders and snakes.

4.4. Data preparation and analysis

4.4.1. Data preparation

Cronbach’s alpha showed acceptable levels of reliability for plausibility (mean $\alpha = 0.79$; 0.77 and 0.81 for the first and second exposure, respectively) and the sense of ownership questionnaire (mean $\alpha = 0.85$; 0.81 and 0.89 for the first and second exposure, respectively). Hence, subsequent analyses used the mean value of the items.

¹ These files are stored for public access on a national database for research data with the 4TU Center for Research Data in the Netherlands. The DOI to this storage is <https://doi.org/10.4121/14339087>.

TABLE 2. Results Bayesian t-tests: eye-gaze-adaptive vs. non-eye-gaze-adaptive VCs, N = 24.

	Mean [HDI] (SD)		MD [HDI]	Effect size d [HDI]	MD* probability below and above
	Adaptive	Non-adaptive			
Number of eye-gaze shift	3.44 [2.60, 4.28] (0.42)	2.37 [1.65, 3.08] (0.36)	1.24 [0.43, 2.02]	0.72 [0.15, 1.41]	0.3% below 0, 99.7% above 0
Ownership	1.68 [1.33, 2.03] (0.18)	1.12 [0.67, 1.60] (0.24)	0.55 [0.31, 0.80]	0.99 [0.47, 1.50]	0% below 0, 100% above 0
Plausibility	1.41 [1.04, 1.76] (0.18)	1.35 [0.86, 1.82] (0.24)	0.080 [-0.15, 0.31]	0.15 [-0.26, 0.59]	24.7% below 0, 75.3% above 0

HDI, 95% high-density interval.

MD, mean difference.

4.4.2. Analysis

A Bayesian approach was used to examine the data. Compared to the frequentist approach, it has been associated with richer and more informative inferences, performing well with small sample sizes, no reliance on sampling distributions and *P*-value (Kruschke, 2014). By considering the posterior probability distribution of the parameter values supporting and those opposing the hypothesis, the analysis estimated the level of credibility that could be attributed to each hypothesis and its alternative. As an interpretation guideline of the posterior probability, Chechile (2020) labels a probability between 50% and 75% as ‘not worth betting on’, between 75% and 90% as ‘only a casual bet’, between 90% and 95% as ‘a promising but risky bet’ and between 95% and 99% as ‘good bet - too good to disregard’. Also, the analysis estimated the 95% credible interval (CI), which summarizes the central 95% mass of the marginal posterior distribution (Voorre & Bolger, 2018), or 95% highest density interval (HDI), which is the interval that spans 95% of distribution such that every point outside this interval has a lower credibility than points inside the interval (Kruschke, 2014).

All analyses were carried out with R version 3.4.2. The Bayesian First Aid package (Baath, 2014) was used to investigate the effect of adapting VCs to eye gaze on people’s perceived plausibility and ownership of the VCs and the numbers of instructional VCs that were followed up by an eye-gaze shift. This package is an extension of the Bayesian Estimation that Supersedes the T-test (BEST) package (Kruschke, 2013, Meredith & Kruschke, 2018), which works as a Bayesian estimation alternative to the *t*-test and calculated HDIs. The analysis used the default minimally informative priors from the package, namely standard deviations set to 1000 and exponentially distribution for the ν parameter set to 30.

The bmlm package (Voorre & Bolger, 2018) was used to examine the within-subjects mediation effect of the fourth hypothesis. The analysis used the uninformed default values as suggested by Voorre & Bolger (2018) as priors, namely regression parameters, the prior distributions were zero centered, with

the standard deviations of 1000; for the subject-level effects’ standard deviations a Cauchy distribution was used set to 50; for the correlation matrix a LKJ prior ν was used and set to 1. This analysis provides CIs.

All the experiment data and the R markdown file can be found online².

5. RESULTS

Table 2 presents the comparison results for the eye-gaze-adaptive and the control condition, i.e. the non-adaptive condition. The mean of the credible values shows an increase of 1.24 in the number of instructional VCs followed-up by an eye-gaze shift for the adaptive condition, with a 95% HDI ranging from 0.43 to 2.02, which does not include a zero difference as 99.7% of the credible values are above zero. Correspondingly, the data set gives only a 0.3% support to the alternative hypothesis, i.e. zero or a reduced eye-gaze shift. Table 2 gives similar information about the sense of ownership of the VCs. In the eye-gaze adaptive condition, an increase in ownership was observed, with a 0.55 mean difference and 95% HDI [0.31, 0.80] that excludes zero. Almost 100% of the credible values for the difference are above zero and, consequently, there is almost 0% support for the alternative hypothesis. However, the data gives less severe credibility to a difference in the plausibility of the VCs between the two conditions. A total of 75.3% of the credible values of the posterior probability distribution support an increase in plausibility, while 24.7% do not support an increase in plausibility.

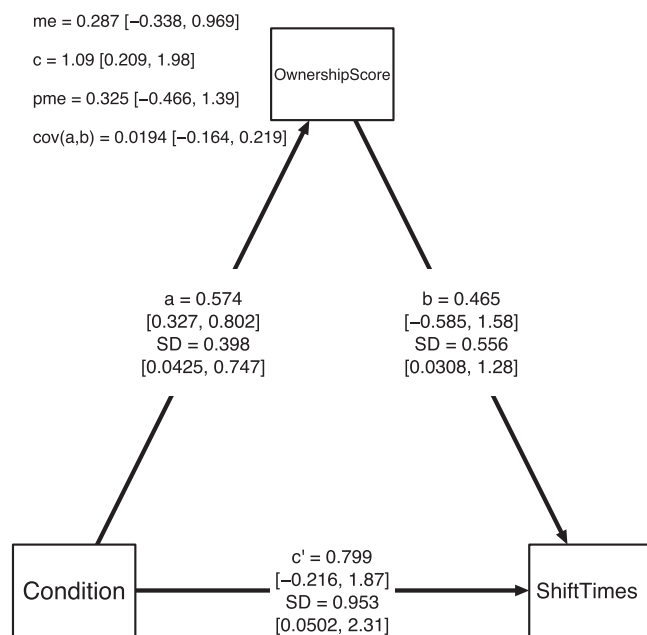
Table 3 shows the correlation between the ownership of VCs and the number of instructional VC followed up by eye-gaze shift behaviour. It shows some credibility for the claim of a positive association between them. Still, the 95% HDI [-0.15, 0.43] does not exclude zero or negative values. More precisely,

² These files are stored for public access on a national database for research data with the 4TU Center for Research Data in the Netherlands. The DOI to this storage is <https://doi.org/10.4121/14339087>.

TABLE 3. The result of correlation between parameters and the mediation effect of ownership and plausibility.

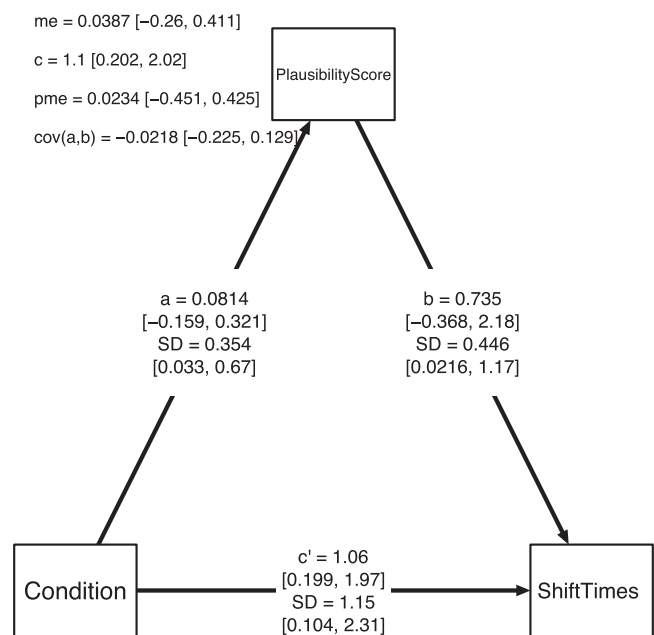
	Correlation		Mediation	
	Correlation [HDI]	Probability below and above	Mean [HDI] (SD)	Probability below and above
Ownership	0.14 [-0.15, 0.43]	18.0% below 0, 82.0% above 0	0.29 [-0.34, 0.97] (0.34)	19.5% below 0, 80.5% above 0
Plausibility	0.03 [-0.27, 0.32]	43.2% below 0, 56.8% above 0	0.04 [-0.26, 0.41] (0.16)	40.7% below 0, 59.3% above 0

HDI, 95% high-density interval.

**FIGURE 4.** Path diagram for the ownership score with point estimates (posterior means) of the parameters and associated 95% intervals (in square brackets below estimates).

82.0% of the credible values indicate a positive association, and 18.0% do not indicate a positive association. In addition to this, Table 3 also shows estimates for a mediation effect that ownership might have on the direct relationship between eye-gaze adaptiveness and the eye-gaze shift behaviour. As shown in Fig. 4, after taking indirect ownership relationship into account, the mean value for the direct effect changes from 1.09 to 0.80. Additionally, 80.5% of the credible values for an indirect effect change indicate a reduction, while 19.5% do not indicate a reduction.

The data gave relatively less credibility to a mediation effect for plausibility. As shown in Table 3, the 95% HDI for the two slopes did not exclude zero or negative associations. Only 56.8% of the credible values suggests a positive association between plausibility and eye-gaze behaviour. Likewise, as shown in Fig. 5, mean estimate for direct effect barely changed

**FIGURE 5.** Path diagram for the plausibility score with point estimates (posterior means) of the parameters and associated 95% intervals (in square brackets below estimates).

only from 1.10 to 1.06. Nevertheless, 59.3% of the posterior distribution support a reduction in the direct effect, while 40.7% do not support a reduction in the direct effect.

Analysis of the people's self-efficacy of handling spiders and snakes suggest an increase between pre- and post-measurement. For handling spiders, the mean increase was 1.6 with 95% HDI [0.7, 2.5], while for snakes, it was 2.7, 95% HDI [1.4, 4.0]. In both cases, almost 100% of the posterior distribution was above zero.

6. DISCUSSION

This study examined whether adapting VCs to individual's eye-gaze affects people's perceived ownership and plausibility of these VCs, thus influencing the numbers of instructional VC followed up by eye-gaze shift. The experiment shows

credible support for hypotheses 1-a and 3, i.e. that compared to non-adaptive, eye-gaze-adaptive VCs have a positive effect on people's perceived ownership of the VCs and the numbers of instructional VCs followed-up by eye-gaze shift. For plausibility (H1-b), the analysis provides relatively less strong credible support, only 75.3% in support and 24.7% opposing it. The finding shows a similar pattern for the relationship between, on the one hand, ownership and plausibility of the VCs and, on the other hand, the numbers of instructional VCs followed up by eye-gaze shift. While support in favour of a positive association for ownership was 82.0% (H2-a), for plausibility it was merely 56.8% (H2-b). The analysis also provided some credible support for the fourth hypotheses, the mediating effect of ownership (H4-a) and plausibility (H4-b) that could partly explain the increase in eye-gaze shifts when VCs become eye-gaze adaptive. Support for ownership mediation was 80.5%, while for plausibility a mere 59.3%. Although these findings are sometimes less strong, they seem overall in favour of all hypotheses, especially the ones related to ownership.

Like any empirical study, this study has some limitations that should be considered. First, a confounding variable for the difference between the adaptive and non-adaptive condition was the co-manipulation of cognitions that used information about the person's eye fixation and generic formulated VCs. It ensured that participants in the control condition were not confronted with conflicting VCs, e.g. while not looking at the snake hearing 'Stop looking out for the snake all the time. Turn your gaze away!'. Removing such potential artificial conflicts by offering generic cognitions resulted in plausible cognitions in both conditions, and hence likely led to less strong support for plausible related hypotheses. A second limitation is a potential error in eye-gaze shift counting caused by the accuracy of the eye-tracking device. For example, whereas people might have looked above the spider, it might have been classified as looking at the spider. Although this might not have systematically biased measurement towards one of the conditions, it could have resulted in some conflicting VCs in the adaptive eye-gaze conditions. The third limitation is the use of non-validated questionnaires. Instead, the plausibility and ownership questionnaire were specifically created for this study based on theoretical concepts and inspired by other related questionnaires. Still, despite this potential limitation, the findings support the superiority of the eye-gaze-adaptive condition. The fourth and final limitation was the use of non-phobic patients in this experiment because of ethical constraints. Therefore, caution should be taken when generalizing the findings to a phobic population. Still, the observed improvement in self-efficacy is promising.

This study can be extended in many directions. First, besides increasing the plausibility and ownership of VCs, other mechanisms could be explored that influence the effect of VCs. Potential mechanisms might be the perspective-taking to hear the VCs or the pronouns (e.g. you, I or one's name) to refer to the self within the VCs. For instance, [Osimo et al. \(2015\)](#)

argued that different perspective-taking to hear and observe can change individual's habitual ways of thinking about personal problems. [Myers et al. \(2014\)](#) found that when people were instructed to take a different perspective (i.e. imagine yourself or imagine someone else) their attitude towards helping others also changed. [Kross et al. \(2014\)](#) indicated that using a different perspective to refer to the self (i.e. I or one's own name) during introspection can influence people's ability to regulate their thoughts, feelings and behaviour. Similar effects have been observed in VR, when people experience from a first-person perspective the world of a person with another skin colour ([Banakou et al., 2016](#), [Peck et al., 2013](#)), a child ([Banakou et al., 2013](#), [Slater et al., 2010](#)) or of a superhero ([Rosenberg et al., 2013](#)). This has often been done by giving people a virtual body of a specific individual. Could something similar be accomplished with VCs? Until now, research has focused on exposing people to VCs that match their own inner voice. What would happen if people are exposed to cognitions that are clearly alien to their own? For example, [Banks et al. \(2004\)](#) exposed people to the world of someone with schizophrenia, including both audio and visual hallucinations. Combining VCs and virtual body, therefore, could potentially give people new experiences. For example, extending the work by [Slater et al. \(2010\)](#) by exposing adults to cognitions of a child in addition to giving them the body of a child.

Instead of a snake or spider phobia, future research could also explore the feasibility of VCs in other application areas. For example, targeting to social skills training, VCs can work as a medium to provide users with knowledge of social skills, encourage them and help them to reflect on their performance ([Ding et al., 2020](#)). For the treatment of autism, VCs could be studied for their ability to act as a guidance by helping people understanding social cues, similar to the idea of social story intervention ([Adjorlu et al., 2018](#)). Besides studying VCs, this study also shows the potential of using eye-tracking information in VR. Further research might, therefore, explore the experience when other elements in the virtual world react to people's eye gaze. For example, [Ginkel et al. \(2019\)](#) adapted the behaviours of a virtual audience depending on eye gaze of the user who gave a public talk.

7. CONCLUSION

This paper demonstrates that adapting VCs to individual's eye gaze is likely to have a guiding effect on people's eye-gaze behaviour in VR and to enhance people's perceived ownership of the VCs. Its main contribution can be summarized as providing initial empirical evidence about the effect of the eye-gaze-adaptive VCs technique— a technique with the potential of enhancing the effectiveness of VCs as a new training or intervention method. The more practical contribution of the study is the insight into the implementation of VCs in VR exposure, especially for a pre-therapy of the spider and snake

phobia. Moreover, we hope that the presented study serves as an inspiration for more research into the mechanism underlying the effect of VCs.

Acknowledgments

This work was supported by the China Scholarship Council (grant number: 201506090167).

REFERENCES

- Achenbach, J., Waltemate, T., Latoschik, M. E. and Botsch, M. (2017) Fast generation of realistic virtual humans. In *Proc. 23rd ACM symposium on virtual reality software and technology*, p. 12. ACM.
- Adjorlu, A., Hussain, A., M ødekj ær, C. and Austad, N. W. (2018) Head-mounted display-based virtual reality social story as a tool to teach social skills to children diagnosed with autism spectrum disorder. In *2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning Through Virtual & Augmented Reality (KELVAR)*. IEEE.
- Allen, A. K., Wilkins, K., Gazzaley, A. and Morsella, E. (2013) Conscious thoughts from reflex-like processes: A new experimental paradigm for consciousness research. *Conscious. Cogn.*, 22, 1318–1331.
- Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E. and Calamaras, M. R. (2013) Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. *J. Consult. Clin. Psychol.*, 81, 751.
- Antony, M. M., McCabe, R. E., Leeuw, I., Sano, N. and Swinson, R. P. (2001) Effect of distraction and coping style on in vivo exposure for specific phobia of spiders. *Behav. Res. Ther.*, 39, 1137–1150.
- Azbel-Jackson, L., Butler, L. T., Ellis, J. A. and Van Reekum, C. M. (2016) Stay calm! regulating emotional responses by implementation intentions: assessing the impact on physiological and subjective arousal. *Cogn. Emot.*, 30, 1107–1121.
- Baath, R. (2014) Bayesian first aid: a package that implements bayesian alternatives to the classical *.test functions in r. In *UseR 2014, the int. R User conf.*
- Banakou, D., Groten, R. and Slater, M. (2013) Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. National Academy of Sciences*, 110, 12846–12851.
- Banakou, D., Hanumanthu, P. D. and Slater, M. (2016) Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Front. Hum. Neurosci.*, 10, 601.
- Banks, J., Ericksson, G., Burrage, K., Yellowlees, P., Ivermee, S. and Tichon, J. (2004) Constructing the hallucinations of psychosis in virtual reality. *J. Netw. Comput. Appl.*, 27, 1–11.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. and Plumb, I. (2001) The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry*, 42, 241–251.
- Blanke, O. and Metzinger, T. (2009) Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.*, 13, 7–13.
- Böhnlein, J., Altegoer, L., Muck, N. K., Roesmann, K., Redlich, R., Dannowski, U. and Leehr, E. J. (2020) Factors influencing the success of exposure therapy for specific phobia: a systematic review. *Neurosci. Biobehav. Rev.*, 108, 796–820.
- Bortolotti, L. and Broome, M. (2009) A role for ownership and authorship in the analysis of thought insertion. *Phenomenol. Cogn. Sci.*, 8, 205–224.
- Botella, C., Baños, R. M., Villa, H., Perpiñá, C. and García-Palacios, A. (2000) Virtual reality in the treatment of claustrophobic fear: a controlled, multiple-baseline design. *Behav. Ther.*, 31, 583–595.
- Bouchard, S., Côté, S., St-Jacques, J., Robillard, G. and Renaud, P. (2006) Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games. *Technol Health Care*, 14, 19–27.
- Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, E., Forget, H., Loranger, C. and Roucaut, F. X. (2017) Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. *Br. J. Psychiatry*, 210, 276–283.
- Bouchard, S., St-Jacques, J., Robillard, G. and Renaud, P. (2008) Anxiety increases the feeling of presence in virtual reality. *Presence: Teleoperators and Virtual Environments*, 17, 376–391.
- Broekens, J., Harbers, M., Brinkman, W.-P., Jonker, C. M., Van den Bosch, K. and Meyer, J.-J. (2012) Virtual reality negotiation training increases negotiation knowledge and skill. In *Int. conf. on intelligent virtual agents*, pp. 218–230. Springer.
- Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J., Carlbring, P. and Powers, M. B. (2019) Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. *J. Anxiety Disord.*, 61, 27–36.
- Carlin, A. S., Hoffman, H. G. and Weghorst, S. (1997) Virtual reality and tactile augmentation in the treatment of spider phobia: a case report. *Behav. Res. Ther.*, 35, 153–158.
- Chechile, R. A. (2020) *Bayesian Statistics for Experimental Scientists: A General Introduction Using Distribution-Free Methods*. MIT Press.
- Cho, H., Godwin, C. A., Geisler, M. W. and Morsella, E. (2014) Internally generated conscious contents: interactions between sustained mental imagery and involuntary subvocalizations. *Front. Psychol.*, 5, 1445.
- Cole, J., Crowle, S., Austwick, G. and Henderson Slater, D. (2009) Exploratory findings with virtual reality for phantom limb pain; from stump motion to agency and analgesia. *Disabil. Rehabil.*, 31, 846–854.
- Corbetta, M. et al. (1998) A common network of functional areas for attention and eye movements. *Neuron*, 21, 761–773.
- Côté, S. and Bouchard, S. (2009) Cognitive mechanisms underlying virtual reality exposure. *Cyberpsycho. Behav.*, 12, 121–129.
- David, N., Newen, A. and Vokeley, K. (2008) The "sense of agency" and its underlying cognitive and neural mechanisms. *Conscious. Cogn.*, 17, 523–534.
- Davis, E. T., Scott, K., Pair, J., Hodges, L. F., and Oliverio, J. (1999). Can audio enhance visual perception and performance in a virtual environment? In *Proc. human factors and ergonomics society annual meeting*, volume 43, pp. 1197–1201. SAGE Publications: Los Angeles, CA.
- De la Peña, N., Weil, P., Llobera, J., Giannopoulos, E., Pomés, A., Spanlang, B., Friedman, D., Sanchez-Vives, M. V. and Slater, M.

- (2010) Immersive journalism: immersive virtual reality for the first-person experience of news. *Presence Teleop. Virt. Environ.*, 19, 291–301.
- Ding, D., Brinkman, W.-P. and Neerincx, M. A. (2020) Simulated thoughts in virtual reality for negotiation training enhance self-efficacy and knowledge. *Int. J. Hum. Comput. Stud.*, 39, 102400.
- Dinh, H. Q., Walker, N., Hodges, L. F., Song, C. and Kobayashi, A. (1999) Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proc. IEEE virtual reality (cat. no. 99CB36316)*, pp. 222–228. IEEE.
- Emmelkamp, P. M., Krijn, M., Hulsbosch, A., De Vries, S., Schuemie, M. J. and van der Mast, C. A. (2002) Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia. *Behav. Res. Ther.*, 40, 509–516.
- Feng, A., Rosenberg, E. S. and Shapiro, A. (2017) Just-in-time, viable, 3-D avatars from scans. *Comput. Animat. Virtual Worlds*, 28, e1769.
- FOVE (2019). FOVE0 Headset Specification. <https://fove-inc.com/product/>.
- Fredrikson, M. (1983) Reliability and validity of some specific fear questionnaires. *Scand. J. Psychol.*, 24, 331–334.
- Freina, L. and Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The int. scientific conf. eLearning and software for education*, vol. 1, p. 133. "Carol I" National Defence University.
- Gallagher, S. (2000) Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.*, 4, 14–21.
- Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness Iii, T. and Botella, C. (2002) Virtual reality in the treatment of spider phobia: a controlled study. *Behav. Res. Ther.*, 40, 983–993.
- Gazzaley, A. and Nobre, A. C. (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.*, 16, 129–135.
- Giummarra, M. J., Gibson, S. J., Georgiou-Karistianis, N. and Bradshaw, J. L. (2008) Mechanisms underlying embodiment, disembodiment and loss of embodiment. *Neurosci. Biobehav. Rev.*, 32, 143–160.
- Gonzalez-Franco, M., Perez-Marcos, D., Spanlang, B. and Slater, M. (2010) The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE virtual reality conf. (VR)*, pp. 111–114. IEEE.
- Grillon, H., Riquier, F., Herbelin, B. and Thalmann, D. (2006) Virtual reality as therapeutic tool in the confines of social anxiety disorder treatment. *Int. J. Disabil. Hum. Dev.*, 5, 243–250.
- Hahn, B., Wolkenberg, F. A., Ross, T. J., Myers, C. S., Heishman, S. J., Stein, D. J., Kurup, P. K. and Stein, E. A. (2008) Divided versus selective attention: Evidence for common processing mechanisms. *Brain Res.*, 1215, 137–146.
- Hermann, A., Schäfer, A., Walter, B., Stark, R., Vaitl, D. and Schienle, A. (2009) Emotion regulation in spider phobia: role of the medial prefrontal cortex. *Soc. Cogn. Affect. Neurosci.*, 4, 257–267.
- Hermans, D., Vansteenwegen, D. and Eelen, P. (1999) Eye movement registration as a continuous index of attention deployment: Data from a group of spider anxious students. *Cogn. Emot.*, 13, 419–434.
- Hillaire, S., Lécuyer, A., Cozot, R. and Casiez, G. (2008) Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE Comput. Graphics Appl.*, 28, 47–55.
- Hoffman, H. G. (1998) Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Proc. IEEE 1998 virtual reality annual int. symposium (cat. no. 98CB36180)*, pp. 59–63. IEEE.
- Howard, M. C. (2017) A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Comput. Hum. Behav.*, 70, 317–327.
- Iacoboni, M. (2005) Divided attention in the normal and the split brain: chronometry and imaging. In Itti, L., Rees, G., Tsotsos, J. (eds), *Neurobiology of Attention*, pp. 363–368. Elsevier, San Diego, CA.
- Jensen, L. and Konradsen, F. (2018) A review of the use of virtual reality head-mounted displays in education and training. *Educ. Inform. Technol.*, 23, 1515–1529.
- Kalckert, A. and Ehrsson, H. H. (2012) Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Front. Hum. Neurosci.*, 6, 40.
- Kang, N., Ding, D., Hartanto, D., Neerincx, M. A. and Brinkman, W.-P. (2019) Public speaking in virtual reality: audience design and speaker experiences, p. 89. Annual Review of Cybertherapy And Telemedicine.
- Kilteni, K., Groten, R. and Slater, M. (2012) The sense of embodiment in virtual reality. *Presence Teleop. Virt. Environ.*, 21, 373–387.
- Krijn, M., Emmelkamp, P., Biemond, M., de Ligny, C. D. W. R., Schuemie, M. J. and van der Mast, C. A. (2004) Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behav. Res. Ther.*, 42, 229–239.
- Kross, E., Bruehlman-Senecal, E., Park, J., Burson, A., Dougherty, A., Shablack, H., Bremner, R., Moser, J. and Ayduk, O. (2014) Self-talk as a regulatory mechanism: How you do it matters. *J. Pers. Soc. Psychol.*, 106, 304.
- Kruschke, J. (2014) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.
- Kruschke, J. K. (2013) Bayesian estimation supersedes the t test. *J. Exp. Psychol. General*, 142, 573.
- Larsson, P., Västfjäll, D., and Kleiner, M. (2001). *Ecological Acoustics and the Multi-modal Perception of Rooms: Real and Unreal Experiences of Auditory-Visual Virtual Environments*. Georgia Institute of Technology.
- Larsson, P., Västfjäll, D., Olsson, P. and Kleiner, M. (2007) When what you hear is what you see: presence and auditory-visual integration in virtual environments. In *Proc. 10th annual int. workshop on presence*, pp. 11–18.
- Lindquist, M., Lange, E. and Kang, J. (2016) From 3D landscape visualization to environmental simulation: the contribution of sound to the perception of virtual environments. *Landsc. Urban Plan.*, 148, 216–231.
- Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I. and Brinkman, W.-P. (2014) A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PLoS One*, 9, e96144.
- Maselli, A. and Slater, M. (2013) The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.*, 7, 83.
- Meredith, M. and Kruschke, J. (2018). *Bayesian Estimation Supersedes the t-test*.

- Meuret, A. E., Seidel, A., Rosenfield, B., Hofmann, S. G. and Rosenfield, D. (2012) Does fear reactivity during exposure predict panic symptom reduction? *J. Consult. Clin. Psychol.*, 80, 773.
- Morin, A. (2009). Inner speech and consciousness. *Encyclopedia of Consciousness*, 389–402
- Morin, A. and Everett, J. (1990) Inner speech as a mediator of self-awareness, self-consciousness, and self-knowledge: a hypothesis. *New Ideas Psychol.*, 8, 337–356.
- Morin, A. and Hamper, B. (2012) Self-reflection and the inner voice: activation of the left inferior frontal gyrus during perceptual and conceptual self-referential thinking. *Open Neuroimag. J.*, 6, 78.
- Morina, N., Ijntema, H., Meyerbröcker, K. and Emmelkamp, P. M. (2015) Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behav. Res. Ther.*, 74, 18–24.
- Mullins, S. and Spence, S. A. (2003) Re-examining thought insertion: Semi-structured literature review and conceptual analysis. *Br. J. Psychiatry*, 182, 293–298.
- Myers, M. W., Laurent, S. M. and Hodges, S. D. (2014) Perspective taking instructions and self-other overlap: different motives for helping. *Motiv. Emot.*, 38, 224–234.
- Naef, M., Staadt, O. and Gross, M. (2002) Spatialized audio rendering for immersive virtual environments. In *Proc. of the ACM symposium on virtual reality software and technology*, pp. 65–72. ACM.
- Normand, J.-M., Giannopoulos, E., Spanlang, B. and Slater, M. (2011) Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PLoS One*, 6, e16128.
- Nowak, K. L. and Biocca, F. (2003) The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence Teleop. Virt. Environ.*, 12, 481–494.
- Olickers, C. (2013) Psychological ownership: development of an instrument. *SA J. Ind. Psychol.*, 39, 1–13.
- Osimo, S. A., Pizarro, R., Spanlang, B. and Slater, M. (2015) Conversations between self and self as sigmund freud-a virtual body ownership paradigm for self counselling. *Sci. Rep.*, 5, 13899.
- Paquette, V., Lévesque, J., Mensour, B., Leroux, J.-M., Beaudoin, G., Bourgouin, P. and Beauregard, M. (2003) "Change the mind and you change the brain": effects of cognitive-behavioral therapy on the neural correlates of spider phobia. *Neuroimage*, 18, 401–409.
- Peck, T. C., Seinfeld, S., Aglioti, S. M. and Slater, M. (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cogn.*, 22, 779–787.
- Perez-Marcos, D., Slater, M. and Sanchez-Vives, M. V. (2009) Inducing a virtual hand ownership illusion through a brain-computer interface. *Neuroreport*, 20, 589–594.
- Pertaub, D.-P., Slater, M. and Barker, C. (2002) An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence Teleop. Virt. Environ.*, 11, 68–78.
- Pflugshaupt, T., Mosimann, U. P., Schmitt, W. J., von Wartburg, R., Wurtz, P., Lüthi, M., Nyffeler, T., Hess, C. W. and Müri, R. M. (2007) To look or not to look at threat?: Scanpath differences within a group of spider phobics. *J. Anxiety Disord.*, 21, 353–366.
- Price, M. and Anderson, P. (2007) The role of presence in virtual reality exposure therapy. *J. Anxiety Disord.*, 21, 742–751.
- Price, M., Mehta, N., Tone, E. B. and Anderson, P. L. (2011) Does engagement with exposure yield better outcomes? components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. *J. Anxiety Disord.*, 25, 763–770.
- Qu, C., Brinkman, W.-P., Wiggers, P. and Heynderickx, I. (2013) The effect of priming pictures and videos on a question-answer dialog scenario in a virtual environment. *Presence Teleop. Virt. Environ.*, 22, 91–109.
- Rosenberg, R. S., Baughman, S. L. and Bailenson, J. N. (2013) Virtual superheroes: using superpowers in virtual reality to encourage prosocial behavior. *PLoS One*, 8, e55003.
- Rothbaum, B. O., Hodges, L., Smith, S., Lee, J. H. and Price, L. (2000) A controlled study of virtual reality exposure therapy for the fear of flying. *J. Consult. Clin. Psychol.*, 68, 1020.
- Rovira, A., Swapp, D., Spanlang, B. and Slater, M. (2009) The use of virtual reality in the study of people's responses to violent incidents. *Front. Behav. Neurosci.*, 3, 59.
- Schienze, A., Schäfer, A., Stark, R. and Vaitl, D. (2009) Long-term effects of cognitive behavior therapy on brain activation in spider phobia. *Psychiatry Res. Neuroimaging*, 172, 99–102.
- Shimada, S., Hiraki, K. and Oda, I. (2005) The parietal role in the sense of self-ownership with temporal discrepancy between visual and proprioceptive feedbacks. *Neuroimage*, 24, 1225–1232.
- Slater, M. (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc.*, 364, 3549–3557.
- Slater, M., Khanna, P., Mortensen, J. and Yu, I. (2009) Visual realism enhances realistic response in an immersive virtual environment. *IEEE Comput. Graphics Appl.*, 29, 76–84.
- Slater, M. and Sanchez-Vives, M. V. (2014) Transcending the self in immersive virtual reality. *Computer*, 47, 24–30.
- Slater, M., Spanlang, B., Sanchez-Vives, M. V. and Blanke, O. (2010) First person experience of body transfer in virtual reality. *PLoS One*, 5, e10564.
- Sokolov, A. (2012) *Inner Speech and Thought*. Springer Science & Business Media, New York, USA.
- Stroop, J. R. (1935) Studies of interference in serial verbal reactions. *J. Exp. Psychol.*, 18, 643.
- Tullett, A. M. and Inzlicht, M. (2010) The voice of self-control: Blocking the inner voice increases impulsive responding. *Acta Psychol.*, 135, 252–256.
- Valmaggia, L. R., Latif, L., Kempton, M. J. and Rus-Calafell, M. (2016) Virtual reality in the psychological treatment for mental health problems: a systematic review of recent evidence. *Psychiatry Res.*, 236, 189–195.
- van Ginkel, S., Gulikers, J., Biemans, H., Noroozi, O., Roozen, M., Bos, T., van Tilborg, R., van Halteren, M. and Mulder, M. (2019) Fostering oral presentation competence through a virtual reality-based task for delivering feedback. *Comput. Educ.*, 134, 78–97.
- van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L. and Davidson, R. J. (2007) Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *Neuroimage*, 36, 1041–1055.
- Vuorre, M. and Bolger, N. (2018) Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behav. Res. Methods*, 50, 2125–2143.

- Wickens, C. D., Hollands, J. G., Banbury, S. and Parasuraman, R. (2015) *Engineering Psychology and Human Performance*. Psychology Press, New York, USA.
- Xing, C. and Isaacowitz, D. M. (2006) Aiming at happiness: How motivation affects attention to and memory for emotional images. *Motiv. Emot.*, 30, 243–250.
- Yung, R. and Khoo-Lattimore, C. (2019) New realities: a systematic literature review on virtual reality and augmented reality in tourism research. *Curr. Issues Tour.*, 22, 2056–2081.
- Zotkin, D. N., Duraiswami, R. and Davis, L. S. (2004) Rendering localized spatial audio in a virtual auditory space. *IEEE Trans. Multimed.*, 6, 553–564.