When Fear Overshadows Perceived Plausibility: The Influence of Incongruencies on Acrophobia in VR

Larissa Brübach* — HCI† and PIIS‡ Group, University of Würzburg, Germany

Deniz Celikhan§ — University of Würzburg, Germany

Lennard Rüffert§ — University of Würzburg, Germany

Franziska Westermeier — HCI and PIIS Group, University of Würzburg, Germany

Marc Erich Latoschik — HCI Group, University of Würzburg, Germany

Carolin Wienrich — PIIS Group, University of Würzburg, Germany

Figure 1: Fear of height inducing condition used in Experiment 1 (left) and Experiment 2 (right).

ABSTRACT

Virtual Reality Exposure Therapy (VRET) has become an effective, customizable, and affordable treatment for various psychological and physiological disorders. Specifically, it is used to treat specific anxiety disorders, such as acrophobia or arachnophobia, for decades. However, to ensure a positive outcome for patients, we must understand and control the effects potentially caused by the technology and medium of Virtual Reality (VR) itself. This article specifically investigates the impact of the Plausibility illusion (Psi), as one of the two theorized presence components, on the fear of heights. In two experiments, 30 participants each experienced two different heights with congruent and incongruent object behaviors in a 2 x 2 within-subject design. Results show that the strength of the congruence manipulation plays a significant role. Only when incongruencies are strong enough will they be recognized by users, specifically in high fear conditions, as triggered by exposure to increased heights. If incongruencies are too subtle, they seem to be overshadowed by the stronger fear reactions. Our evidence contributes to recent theories of VR effects and emphasizes the importance of understanding and controlling factors potentially assumed to be incidental, specifically during VRET designs. Incongruencies should be controlled so that they do not have an unwanted influence on the patient's fear response.

Index Terms: XR, VR, Plausibility, Congruence, Acrophobia, Virtual reality exposure therapy.

*e-mail: larissa.bruebach@uni-wuerzburg.de

†Human-Computer Interaction Group

‡Psychology of Intelligent Interactive Systems Group

§the authors have contributed equally to this work

1 INTRODUCTION

Virtual reality exposure therapy (VRET) has been researched and applied for nearly 30 years [22], offering numerous advantages in treating phobias. VR enables customizable therapy environments tailored to individual progress, providing tangible consequences and infinite scenarios enriched with additional elements. In addition, VR transcends the limitations of the real world [1], allowing therapies to adapt to patient needs while maintaining precise control over the environment. Physiological data, such as heart rate, can be monitored, and potential disruptions can be minimized or eliminated. Therapy can occur in a private, protected space rather than public settings, reducing embarrassment and enhancing patient comfort. VR also broadens access to therapy in rural areas with limited resources and specialists, enabling scalable solutions with minimal effort.

The effectiveness of VRET has been well-documented [1, 9, 12], with some studies indicating that it may surpass traditional real-life therapy [1]. VRET applications include treating anxiety disorders such as claustrophobia [9], acrophobia [12, 18], or arachnophobia [2, 23], by gradually introducing fear triggers and adjusting intensity to the patient's progress.

Given its promise, research has explored key factors influencing VRET success, including presence, immersion, and plausibility. However, while some studies link presence positively with fear responses [7, 25], others report no effect [16, 18, 27] or even a negative correlation [21, 35]. Presence also appears unrelated to physiological fear reactions [33]. Plausibility illusion, however, may play a more significant role than presence when it comes to peoples fear reactions [20]. Nonetheless, the influence of plausibility on fear seems to be underresearched [22].

With these mixed findings, the determinants of successful VR-based therapy remain uncertain to some notable extent, necessitating further exploration of factors like presence, immersion, and plausibility—particularly in light of growing interest in plausibility's role in XR [19, 32]. Thus, the present research investigates the relationship between perceived plausibility and fear of heights. In two experiments, we manipulated plausibility via object behavior congruence in low- and high-fear environments. Acrophobia, affecting about 1 in 20 adults [8], was chosen to evoke fear even in a non-clinical sample [12, 18]. Through targeted manipulation of plausibility and height, the present research contributes to a better understanding of the relationship between them and, thus, the VR-specific conditions for successful VRET.

2 Related Work

2.1 Virtual Reality Exposure Therapy

Literature reviews regarding VRET shows a great body of research in this field [1, 22, 24]. Its applications span diverse phobias, including acrophobia, social phobia, agoraphobia, claustrophobia, fear of animals, and many more. Studies have examined factors influencing VRET success, such as temporal effects [23] and pre-existing conditions [21]. However, studies on the influence of presence on fear account for most research to date (e.g. [7, 12, 16, 18, 20, 25, 27, 33]).

2.2 Presence and VRET

Presence, often central to VRET research, is widely defined using Slater's model [31], which comprises of the two orthogonal factors place illusion (PI) and plausibility illusion (Psi). PI is the feeling of "being there" and is used as an analogy to spatial presence. Immersion, as the technical properties of a VR application, is seen as the frame within which PI can occur. Psi is defined as the "credibility of events in comparison with what would be expected in reality in similar circumstances" [31, p. 3556]. This model by Slater [31] was later used as a basis by Skarbez et al. [28]. Among other things they introduced the social presence illusion. More importantly, they argue that there has to be a similar concept to immersion that influences Psi, which they specify as coherence. They define coherence as the level to which a "virtual scenario behaves in a reasonable and predictable way" [30, p. 44]. Skarbez [30] also argues that the perceived plausibility depends on users' expectations of a virtual environment. This model can be seen in figure 2.

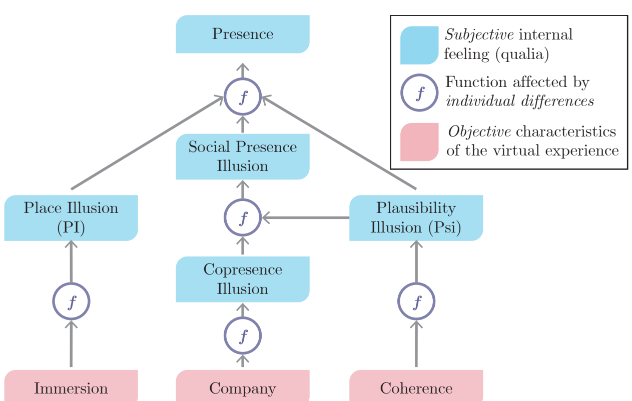


Figure 2: Slater's model, which describes the emergence of presence through place illusion and plausibility illusion [31] extended by Skarbez [29]. (layout redesigned by the authors)

PI is crucial for VRET as it enables users to feel "present" in phobia-inducing scenarios, such as high places for acrophobia treatment. As immersion is the frame within which PI can occur, it is often manipulated to investigate the influence of PI on fear [12, 18, 21, 23]. Studies typically compare highly immersive conditions (e.g., stereoscopic displays or CAVEs) with less immersive setups. It is assumed that the higher immersion should lead to a stronger feeling of PI and overall presence.

Peperkorn et al. [23] tested a stereoscopic (high immersion/presence) against a monoscopic (low immersion/presence) setting of a powerwall. They used arachnophobia, the fear of spiders, to induce fear. The participants were women who previously expressed that they experienced fear when confronted with spiders. They were repeatedly confronted with a large spider. The fear reactions were measured several times over the experiment. Results show that in the stereoscopic condition, fear reactions were stronger and presence ratings higher. After the first trial, presence seemed to directly influence fear, while over time, the two seemed to depend on each other mutually.

Krijn et al. [18] tested a head-mounted display (HMD, low presence) against a computer automatic virtual environment (CAVE, high presence). They used acrophobia, the fear of heights, to induce fear. Different virtual environments were tested, increasing the environment's height and, thus, fear each time. Results show that the CAVE environment did result in higher presence ratings. However, they found no significant differences between the two presence conditions regarding their effect on fear.

Gromer et al. [12] used the degree of sensory realism to manipulate presence. They tested high sensory realism against low sensory realism (reduced vertices, missing sounds). They also used acrophobia to induce fear. They tested a high-fear environment (virtual height) against a neutral control condition (forest environment). A stronger emotional involvement triggered by fear and a higher degree of sensory realism lead to a greater sense of presence. A higher presence rating in the control condition was a predictor for a higher fear response in the later high-fear condition. This was probably due to personal factors of the individuals. However, the presence manipulation did not increase or decrease fear.

2.3 Plausibility Illusion and Higher-Level Factors

Lin [20] shifted focus to Psi, exploring its role in fear elicitation. Instead of comparing low fear to high fear or low presence to high presence, they investigated the difference in the influence of PI and Psi elements on fear. This makes it one of the first studies to look specifically at the influence of Psi. They were interested in which elements contributed most to people's fear. PI elements in the game were defined as those related to the environment, e.g., ambient sounds or lighting conditions. Psi elements, conversely, are game or action-related things, such as control over what is happening, and thus, are more of a semantic nature. Another difference is that they did not use specific anxiety (like acrophobia), but rather general fear elicited from a VR horror game. Participants played the VR horror game and afterward rated different elements based on the impact it had on their fear. The results show that the Psi elements were more important for triggering fear than the PI elements. This underscores the importance of Psi in VR applications.

Ling et al. [22] similarly emphasized the underexplored role of Psi, noting that most presence research prioritizes PI. While previous studies and questionnaires often focussed on the PI component of Slater's presence definition, the general role of plausibility in VR recently gained increased attention [3, 4, 5, 6, 19, 32, 36, 37].

2.4 Congruence and Plausibility Model

According to the Congruence and Plausibility model (CaP) [19], plausibility arises from congruent information on the sensation, perception and cognition layer. The model does not assume plausibility to be orthogonal to the PI like in Slater's presence model [31], shifting the focus to cues from the three layers, which in turn holistically impact plausibility. For example, the model assumes Slater's PI to appear due to congruent spatial cues giving rise to a plausible spatial sensation. Overall, plausibility influences other qualia like placeness, body ownership, or co-presence. This new model can be seen in figure 3.

In this model, the often-used immersion manipulation is located on the perceptual layer. Previous work has shown that lower-level cues have a stronger effect on participants than higher-level cues [5]. However, higher-level cues must also be investigated to better understand their influence. In Lin [20], the Psi elements were of a semantic nature and often in the context of the task (control of what is happening). Brübach et al. [6] showed that incongruencies related to tasks may have a stronger effect than those not related to the task. This emphasizes the need for a multi-layered investigation.

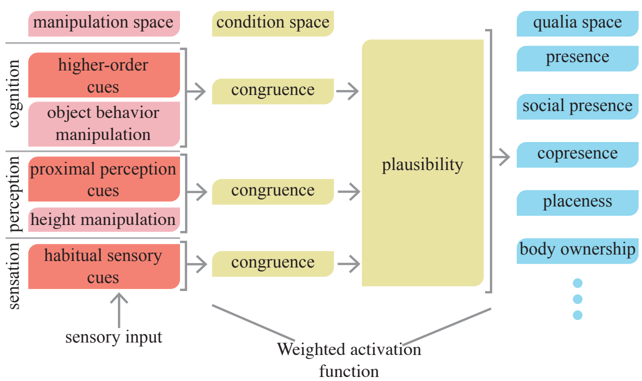


Figure 3: The Congruence and Plausibility model [19] and the manipulation location. (layout redesigned by the authors)

2.5 Summary and Present Work

While the influence of presence and immersion on the feeling of fear in VR has been investigated several times [1, 21], there is little research on the influence of plausibility [20, 22]. Previous research focused on manipulating immersion, a lower-level manipulation in the CaP model. We see a research gap in understanding the impact of higher-level manipulation, building on existing evidence that suggests such an effect [20].

For this purpose, we have developed two applications designed to trigger people's fear of heights. We did not manipulate immersion and kept the PI consistent for both conditions. In two experiments, we manipulated the object behavior congruence, influencing the perceived plausibility. In the first experiment, we manipulated the objects' gravity as in Brübach et al. [5], and in the second experiment, we manipulated the objects' size similar to Brübach et al. [3, 6]. In Experiment 1, we encountered challenges, and open questions arose from the results. To address this and further investigate the underlying phenomena, we designed and conducted Experiment 2. This is a semantic manipulation located on the cognitive layer of the CaP model. Incongruent object behavior should lead to a break in plausibility, which is defined as a significant effect between the manipulation and the control condition in the plausibility questionnaire. Participants interacted with these manipulated objects in a fear-inducing environment (top of a skyscraper or radio tower) or in a neutral environment (sidewalk or ground). This can be seen as a perceptual manipulation. We then measured both the perceived plausibility and their subjective feeling of fear. Our results contribute to a better understanding of the relationship between fear and perceived plausibility, showing that the intensity of the incongruence is an important factor for its influence on perceived fear. This shows that in the context of therapy, attention must be paid to how strong incongruences are and how they affect the interaction with the virtual environment.

Based on the results of Brübach et al. [5], we assume that the weightlessness (Experiment 1) or the change of size (Experiment 2) of objects is an effective incongruence manipulation and can affect the perceived plausibility. So, our first hypothesis is as follows:

- **H1** Incongruent object behavior, i.e., weightlessness (Ex1) or size changes (Ex2), will reduce perceived plausibility.

In the Lin [20] experiment, the Psi elements contributed more to participants' fear more than PI elements. This leads us to conclude that manipulating Psi elements, and thus reducing their influence, can reduce participants' fear reactions.

- **H2** Participants' fear will be lower with incongruent object behavior.

3 Experiment 1

3.1 Methods

Ethical approval was not required by the institution for this study. However, safety precautions identical to those approved for Experiment 2 were implemented, including conducting the experiment in an empty room to prevent injuries, closely monitoring participants, and providing detailed information about the environment beforehand.

3.1.1 Study Design

We used a randomized 2 x 2 within-subject design. The first factor was *height*. Participants were either down on the street sidewalk with the baskets in front of them (low-fear) or on top of a skyscraper where they had to cross a plank to reach them (high-fear). The second factor was the *congruence*. Participants had to sort colored bottles into the corresponding baskets. These bottles either behaved normally, i.e., they fell to the ground when participants let go of them (congruent), or they had a manipulated gravity, i.e., floating when they were let go (incongruent). This congruence manipulation should lead to a reduced perceived plausibility.

3.1.2 Application

We used a high-end computer with an Nvidia Geforce RTX 3080 GPU with 64 GB of RAM and an Intel i9-11900K CPU. The application was developed in the Unity Engine (v2021.3.11f1) using the Open XR Plugin from Microsoft Mixed Reality (v1.5.3) and the XR Interaction Toolkit (v2.0.4). The HP Reverb headset was used for development and user study.

We used the city center of an urban city as the environment. It contained several skyscrapers. We omitted moving objects or unrelated sounds to avoid distracting participants from the task. The bottles participants had to sort were dispensed by a tube on their left side. They fell into a bowl where participants were able to pick them up. Participants then had to walk a short distance to the baskets to sort the bottles into. In the high-fear condition, this involved walking on a plank over the edge of the building to reach the baskets. The plank had a length of approximately 1.8 m, and the building was 24 stories high. The environment can be seen in figure 4.

3.1.3 Measures

A variation of the Perceived Plausibility Questionnaire (PPQ) proposed by Brübach et al. [5] was used to measure the perceived plausibility. As in previous experiments [3, 4, 6], we replaced the term *objects* with *scenario* to better represent the perceived plausibility of the whole XR experience. The questionnaire has 13 items on a 7-point Likert scale with the endpoint *I do not agree at all (1)* and *I fully agree (7)*.

We used the visual height intolerance severity scale (vHISS) by Huppert et al. [14] to measure the participants' predisposition for acrophobia. This questionnaire measures the severity of visual height intolerance on a scale from 0 (least severely affected) to 13 (most severely affected). This is used as a control variable to ensure that no participant has clinical acrophobia. This was done as a precaution so that participants would not suffer any harm during the study.

(a) low, congruent (b) low, incongruent (c) high, congruent (d) high, incongruent

Figure 4: The four conditions of the experiment.

During the experiment, we used the Subjective Units of Distress Scale (SUDS) by Wolpe [38]. It measures feelings of anxiety on a one-item Likert scale from 0 (a state of absolute calmness) to 10 (the worst anxiety ever experienced). After each condition, participants were asked whether they had a heightened sense of stress from the height manipulation. The SUDS has been used in previous research for different anxiety disorders [23].

We used the Cybersickness in Virtual Reality Questionnaire (CSQ-VR) by Kourtesis et al. [17] to control for VR sickness. It measures sickness caused by VR with the subscales *Nausea*, *Vestibular*, and *Oculomotor*. Each subscale has two items with a 7-point Likert scale from 1 (absent feeling) to 7 (extreme feeling).

Additionally, the NASA-TLX by Hart et al. [13] was used to control for differences in the workload between the congruence conditions. It assesses mental, physical, and temporal demand and performance, effort, and frustrations participants feel during a task on a scale from 1 to 100.

3.1.4 Procedure

The study procedure can be seen in figure 5. The experiment took approximately 45 minutes. It started with the participants filling out the consent forms and answering demographical questions and questions about their previous VR experience. Additionally, they answered the vHISS questionnaire and the CSQ-VR. Participants were then informed that they could stop the experiment at any time with no consequences if they felt uncomfortable. The experiment started with a short tutorial to familiarize the participants with the application's controls. The participants then started with the first condition in VR. The VR part took about 5 minutes. Participants had to sort 10 bottles correctly. After the VR part, they answered the PPQ, the SUDS, the CSQ-VR, and the NASA-TLX. They repeated this four times. At the end, participants were told about the study's intention and had the option to ask questions.

3.1.5 Participants

A power analysis determined that the experiment required a minimum of 24 participants (effect size 0.25, estimated power 0.8). Thirty participants took part in the experiment. We had to exclude one participant from the data analysis due to technical problems. The pool was divided into eighteen female, ten male, and one non-binary participant. They were between 19 and 59 years old and the mean age was $M = 28.69$ ($SD = 12.48$). There were twenty-one students, five employees, two pupils, and one pensioner. Five participants had less than one hour of VR experience, eleven had between one to five hours of experience, three had more than ten hours of experience, and three had more than twenty hours of experience. Before the experiment started, the results of the vHISS questionnaire were analyzed. No participant showed a conspicuous disposition for acrophobia.

3.2 Results Experiment 1

We calculated a repeated measures ANOVA with a significance level of $p < .05$ for all variables. All means and standard deviations for each condition for Experiment 1 can be seen in table 1. We used Jasp 0.19 for the data analysis.

3.2.1 Control Variables

CSQ-VR We found a significant effect for the congruence factor in the CSQ-VR ($F(1,28) = 4.42, p = .045, \eta_p^2 = .03$). The incongruent behavior seemed to have decreased VR sickness. We found no significance for the height factor or an interaction effect between the two.

NASA-TLX The NASA-TLX also showed a significant main effect for the congruence factor ($F(1,28) = 80.61, p < .001, \eta_p^2 = .42$), and for the height factor ($F(1,28) = 48.90, p < .001, \eta_p^2 = .19$), and an interaction effect between the two ($F(1,28) = 16.39, p < .001, \eta_p^2 = .05$). Both the high-fear and the incongruent conditions caused a higher workload for participants.

Pairwise comparison using the Holm correction showed that the interaction effect is between the high-fear congruent condition and the low-fear congruent condition ($p_{holm} < .001$). There is also a significant effect between the high-fear incongruent condition and all other conditions: the low-fear congruent condition ($p_{holm} < .001$), the high-fear incongruent condition ($p_{holm} < .001$), and the low-fear incongruent condition ($p_{holm} < .001$). The high-fear incongruent condition has a significantly higher workload than all other conditions. Additionally, there is a significant effect between the low-fear incongruent condition and both the high-fear congruent condition ($p_{holm} = .033$) and the low-fear congruent condition ($p_{holm} = .018$). The low-fear incongruent condition has a higher workload than both the high-fear and the low-fear congruent condition.

The high-fear conditions have a higher workload than the low-fear conditions. The incongruent conditions have a higher workload than the congruent conditions. A combination of high-fear and incongruent has a significantly higher workload than all other conditions.

Results for each subscale can be seen in figure 8. The results of the post-hoc tests can bee seen in table 2.

vHISS No participant reached the threshold to consider "clinical acrophobia".

3.2.2 Subjective Units of Distress Scale

In the SUDS, there was a significant main effect for both congruence ($F(1,28) = 39.10, p < .001, \eta_p^2 = .42$) and height ($F(1,28) = 5.53, p = .026, \eta_p^2 = .03$). We found no significant interaction effect between the two.

3.2.3 Perceived Plausibility Questionnaire

We found significant main effects for both congruence ($F(1,28) = 90.60, p < .001, \eta_p^2 = .55$) and height ($F(1,28) = 20.61, p < .001, \eta_p^2 = .08$). We found no significant interaction effect between the two.

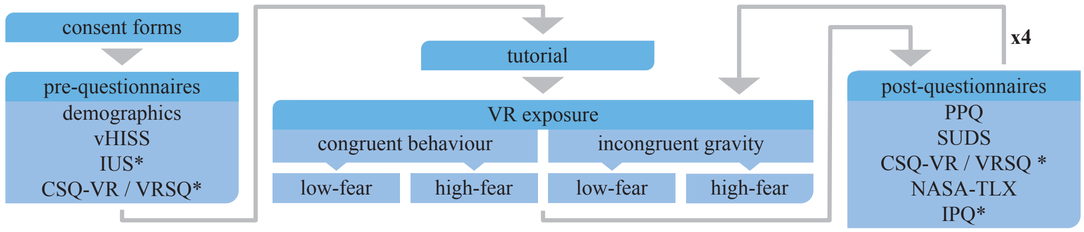


Figure 5: Experimental procedure consisting of pre-questionnaires, four VR exposures, and post-questionnaires after each condition. Questionnaires marked with * are only used in Experiment 2.

Height Congruence	Low-Fear Congruent		Low-Fear Incongruent		High-Fear Congruent		High-Fear Incongruent	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NASA-TLX	5.90	9.99	12.02	12.30	17.74	15.91	35.69	18.56
CSQ-VR	-0.35	2.06	-1.21	3.61	1.10	3.28	-0.14	2.52
SUDS	0.18	0.37	0.67	1.17	2.52	2.35	3.45	3.26
PPQ	5.95	0.73	3.33	1.55	4.75	1.39	2.83	0.95

Table 1: Means and standard deviations of all questionnaires for each condition in Experiment 1.

3.3 Discussion Experiment 1

As expected from previous work [5], we can accept H1: "Incongruent object behavior, i.e., weightlessness, will reduce perceived plausibility." The results from the PPQ show a significantly lower perceived plausibility in the incongruent conditions compared to the congruent ones, regardless of the height condition. Also, the high-fear condition was rated as less plausible than the low-fear condition. This might be due to the design of the environment. Participants had to walk over a plank, and the buckets were placed on a small platform off the building (see figure 1 left). Participants are more used to the low-fear condition where they are placed on a sidewalk, while the high-fear condition on the plank might seem unusual to them and, thus, less plausible.

From the SUDS questionnaire, we can see that our height manipulation worked. Participants had higher stress ratings, which indicates fear, in the high-fear condition than in the low-fear condition. Additionally, participants reported higher stress in the incongruent conditions. We must reject hypothesis H2, "Participants' fear will be lower with incongruent object behavior." The results show that participants reported higher stress and, thus, fear in the incongruent behavior. This is surprising given the findings of Lin et al. [20]. One explanation may be that the unpredictability of the object's behavior has caused additional stress. Dealing with unpredictable events in a frightening situation can add to stress and cause more fear [34]. Humans use prior knowledge and logical thinking to assess situations. If something in this situation now acts implausible, it is hard for humans to assess this behavior and situation, and this can, in turn, lead to fear. Additionally, the congruence manipulation directly influenced the sorting task participants had to complete. The incongruent scenario was more difficult, also reflected in the NASA-TLX. This may have, therefore, caused more pressure and stress for the participants and may, therefore, have used more cognitive resources. However, there was no interaction effect between the height and the congruence. Therefore, we cannot say for sure that incongruence and fear influence each other.

We did find significant effects in the control variable NASA-TLX. The results show that there is a significant interaction effect between the congruence and the height conditions. The results show that the high-fear condition had a higher workload. There was a significant interaction effect between the high-fear conditions and the congruence conditions. However, this effect was not significant for the low-fear condition. This indicates that the participants' experienced fear could explain some of this effect. Fear could have already occupied part of the mental capacity, which means that, in total, more mental capacity was needed to fulfill the task. Again, the unpredictability of the objects' behavior may have contributed to an increase in the mental load of the test subjects. The additional task of walking on the plank could have strengthened this effect. Also, sorting floating bottles is more complex than sorting bottles with normal behavior. Additionally, balancing over the plank might have also introduced an unintentional difficulty.

We also found a lower cybersickness in the incongruent conditions compared to the congruent condition. One explanation could be that as participants felt less immersed due to the implausible behavior, making them less sick. However, the absolute values of the CSQ-VR were very low.

3.3.1 Limitations of Experiment 1

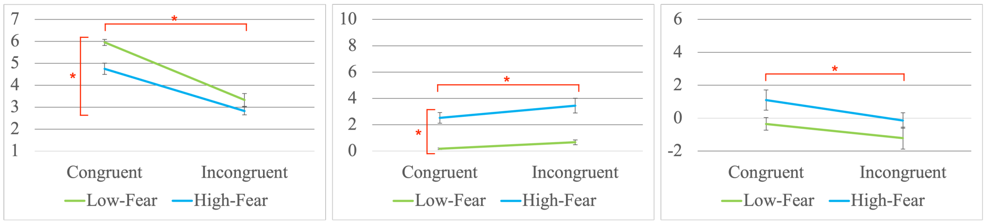
There are a few problems with Experiment 1. First, the NASA-TLX showed a significant difference in task load between the congruence conditions. The effect of this confound might have on other factors, like fear and perceived plausibility, is unclear. We used the gravity of the objects with which the participants interacted. While it makes sense to use a conspicuous manipulation for an initial study, it would be useful to manipulate the object behavior in a way that does not interfere with the completion of the task.

Secondly, the PPQ showed that the high-fear condition had unwanted a priori incongruencies, lowering the overall perceived plausibility in this condition. This might also be due to the nature of the environment. Standing on a skyscraper and walking over a plank is not a normal situation for participants to be in and might, therefore, be seen as implausible.

In this study, we also did not examine the role that presence could have played. However, the feeling of presence could differ between the conditions, and it should, therefore, be considered in a follow-up experiment.

4 Experiment 2

To address the confounds of Experiment 1, we designed a follow-up experiment. As discussed, there was a significant effect in the NASA-TLX in Experiment 1. This may partly be because sorting floating bottles is harder than sorting bottles with normal behavior. Thus, we adapted the congruence manipulation for the objects to change size when interacted with to avoid this issue. The en-



(a) Perceived Plausibility (b) Fear (SUDS) (c) VR Sickness (CSQ-VR)

Figure 6: Means and standard error of the PPQ, SUDS, and CSQ-VR for each condition for Experiment 1. A * marks a significant effect.

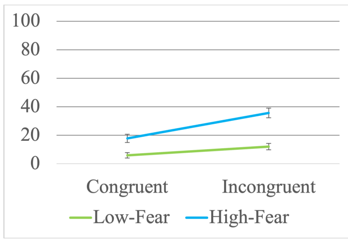


Figure 7: Mean and standard error for the NASA-TLX in Experiment 1

		NASA-TLX			
		Mean Difference	SE	t	p_{holm}
high-fear cong.	low-fear cong.	2.73	.491	4.83	<.001
	high-fear incong.	-3.60	.451	-7.97	<.001
	low-fear incong.	1.14	.524	2.18	.033
low-fear cong.	high-fear incong.	-5.97	.524	-11.39	<.001
	low-fear incong.	-1.23	.451	-2.73	.018
high-fear incong.	low-fear incong.	4.74	.491	9.65	<.001

Table 2: Post hoc comparisons of the interaction effects for the NASA-TLX in Experiment 1.

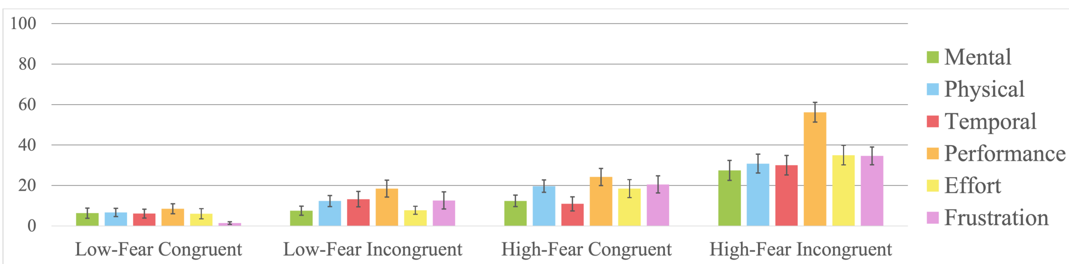


Figure 8: Means and standard deviations of each NASA-TLX subscale for Experiment 1.

vironment in Experiment 1 seemed to introduce unwanted a priori incongruencies. Thus, we also changed the environment to avoid the plank and have a more realistic setting. Lastly, we did not examine the role that presence could have played in the first experiment. However, the feeling of presence could differ between the conditions, so it should be included. Therefore, we used the IPQ to measure presence.

As we changed the manipulation of the object H1 has to be adapted as follows:

- **H1** Incongruent object behavior, i.e., size changes, will reduce perceived plausibility.

With the added measurement of presence, we introduce a third hypothesis. Previous research showed that a higher feeling of presence leads to higher fear [7, 25]. Therefore, our third hypothesis is as follows:

- **H3** Higher presence will lead to a higher feeling of fear.

After Experiment 1, we hypothesized that the uncertainty introduced by the congruence manipulation might have influenced the participants' perceived plausibility. We introduced an uncertainty tolerance questionnaire to see if this is the case. This was designed to show whether participants' tolerance of uncertainty moderated their perception of the environment and the manipulations.

The study received ethical approval from the Institutional Review Board (IRB) (Ethics Committee of the Institute Human-Computer-Media, University of Würzburg).

4.1 Methods

4.1.1 Study Design

We also used a randomized 2 x 2 within-subject design for the second experiment. The first factor was *height*. Participants were either down on the ground in front of the radio tower (low-fear) or on top of the radio tower, with an open elevator to bring them up (high-fear). The second factor was the *congruence*. Participants had to repair a fuse box of the radio tower. The tools either behaved normally, i.e., they kept their size when they were picked up (congruent), or their size behavior was manipulated, i.e., they randomly got bigger or smaller when they were picked up (incongruent). The size was randomly set between 70 and 130% of the original size. This congruence manipulation should lead to a reduced perceived plausibility. The manipulations can be seen in figure 9.

4.1.2 Application

We used a high-end computer with an Nvidia Geforce RTX 3080 GPU with 64 GB of RAM and an Intel i9-11900K CPU. The application was developed in the Unity Engine (v2022.3.23f1) using the Open XR Plugin (v1.10.0). The HTC Vive headset was used in combination with the Valve Index controllers for development and user study.

For our second experiment, we wanted a scenario where participants could see the height more easily without needing a plank. We opted for a radio tower in a forest setting. The tower had an open elevator on the side, bringing participants up to the top to actively

(a) Smaller. (b) Normal Size. (c) Larger.

Figure 9: Incongruent tool behavior.

watch the height increase. Additionally, some of the floor plates on the top were made of glass. At the top and bottom were an open fuse box and a table that held all the tools needed. Participants remained in their place while repairing the fuse box. The size of the tools did not influence the task difficulty. The participants only had to touch the parts (i.e., screws, nails) they were supposed to repair with the tool. The area to be repaired was marked in yellow. They had to carry out 6 repairs by touching the parts 3 times each. The environment can be seen in figure 10.

4.1.3 Measures

As in Experiment 1, we used the PPQ, the vHISS, the SUDS, and the NASA-TLX.

Additionally, we used the Intolerance of Uncertainty (IUS) questionnaire by Freeston et al. [10] in the appropriate translation by Gerlach et al. [11]. It measures participants' emotional, cognitive, and behavioral reactions to ambiguous situations, the implications of uncertainty, and attempts to control the future. It has 27 items on a 5-point Likert scale from 1 (not at all representative) to 5 (completely representative).

We measure presence with the Igroup Presence Questionnaire (IPQ) by Schubert et al. [26], with three subscales: *spatial presence* (five questions), *involvement* (four questions), and *experienced realism* (four questions), as well as one item that does not belong to a subscale. It consists of fourteen items on a scale from 0 to 6, with varying endpoints between the questions.

The Virtual Reality Sickness Questionaire (VRSQ) by Kim et al. [15] was used, which measures sickness caused by virtual reality with the two dimensions *oculomotor* and *disorientation*. It consists of five items for *disorientation* and four items for *oculomotor* on a scale from *not at all* (0) to *strong* (3) to describe the symptoms of VR sickness.

4.1.4 Procedure

The study procedure is the same as in Experiment 1 and can be seen in figure 5. The main differences are the pre- and post-questionnaires. In the second experiment, we used the VRSQ instead of the CSQ-VR. Additionally, we surveyed uncertainty tolerance and presence. The experiment took approximately one hour. The VR part took about 5 minutes for each condition.

4.1.5 Participants

A power analysis determined that the experiment required a minimum of 24 participants (effect size 0.25, estimated power 0.8). Thirty-one participants took part in the experiment. The pool was divided into twenty females, ten males, and one trans-feminine participant. They were between 19 and 63 years old and had a mean age of $M = 31.10 (SD = 14.34)$ years. Ten participants were employees, sixteen were students, three were unemployed, and two had other unspecified occupations. Seven participants had less than one hour of VR experience, ten had between one to five hours of experience, five had between five and ten hours of experience, five had more than ten hours of experience, and four had more than twenty hours of experience. As in Experiment 1, the results of the vHISS questionnaire were analyzed before the VR part started. No participant showed a conspicuous disposition for acrophobia.

4.2 Results Experiment 2

We calculated a repeated measures ANOVA with a significance level of $p < .05$ for all variables. Means and standard deviations of the PPQ, SUDS, and IPQ for each condition for Experiment 2 can be seen in table 3.

4.2.1 Control Variables

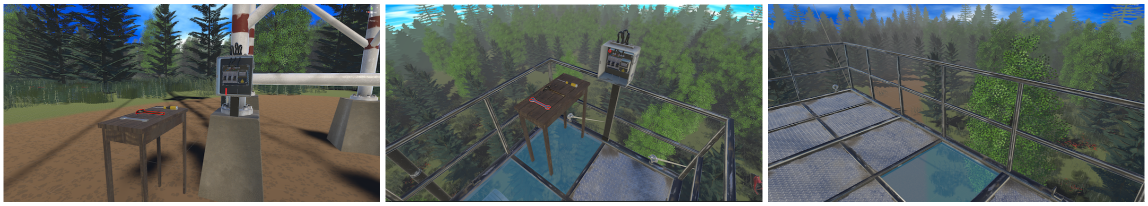
VRSQ We found no significant main effects in the VRSQ for the height ($F(1,30) = 3.19, p = .084, \eta_p^2 = .096$) and the congruence ($F(1,30) = 0.67, p = .420, \eta_p^2 = .022$). There was also no significant interaction effect between the two conditions ($F(1,30) = 1.56, p = .222, \eta_p^2 = .049$).

NASA-TLX We found no significant main effects in the NASA-TLX for the congruence ($F(1,30) = 0.32, p = .575, \eta_p^2 = .011$) and the height ($F(1,30) = 2.58, p = .119, \eta_p^2 = .079$). There was also no significant interaction effect between the two conditions ($F(1,30) = 0.38, p = .541, \eta_p^2 = .013$).

vHISS and IUS Again, the results of the vHISS questionnaire showed that no participant reached the threshold to consider "clicinal acrophobia". We conducted a moderation analysis to determine whether the vHISS significantly predicts fear of heights measured by the SUDS. The overall model was significant for the height conditions, $F(3,120) = 30.59, p < .001$, predicting 13.3% of the variance. Moderation analysis showed that visual height intolerance moderated the effect between height and fear significantly, $\Delta R^2 = 13.3\%, F(1,120) = 28.17, p < .001, 95\%\ CI[-0.866, -0.396]$. For the congruence conditions, the analysis did not show that visual height intolerance moderated the effect between congruence and fear significantly, $\Delta R^2 = 0.6\%, F(1,120) = 3.57, p < .366, 95\%\ CI[-0.162, 0.437]$.

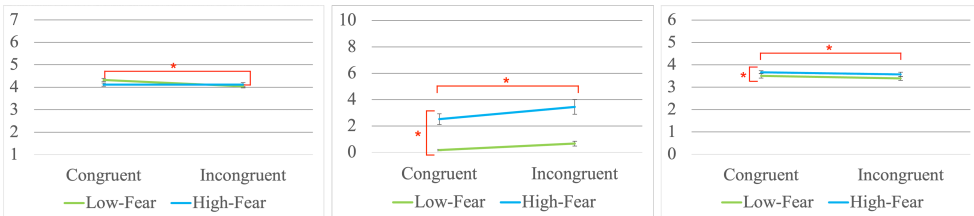
We conducted a moderation analysis to determine whether the IUS significantly predicts fear of heights measured by the SUDS. For the height conditions, the analysis did not show that uncertainty intolerance moderated the effect between height and fear significantly, $\Delta R^2 = 1\%, F(1,120) = 1.71, p < .194, 95\%\ CI[-0.077, 0.016]$. For the congruence conditions, the analysis did not show that uncertainty intolerance moderated the effect between congruence and fear significantly, $\Delta R^2 = 1\%, F(1,120) = 2.80, p < .700, 95\%\ CI[-0.064, 0.043]$.

Lastly, we also conducted a moderation analysis to determine whether the IUS significantly predicts perceived plausibility measured by the PPQ. For the height conditions, the analysis did not show that uncertainty intolerance moderated the effect between height and fear significantly, $\Delta R^2 = 1\%, F(1,120) = 0.379, p < .459, 95\%\ CI[-0.008, 0.017]$. For the congruence conditions, the analysis did not show that uncertainty intolerance moderated the effect between congruence and fear significantly, $\Delta R^2 = 1\%, F(1,120) = 1.54, p < .419, 95\%\ CI[-0.007, 0.017]$.



(a) Bottom of the radio tower. (b) Top of the radio tower. (c) Elevator bringing the participants to the top.

Figure 10: The environment used in Experiment 2.



(a) Perceived Plausibility (b) Fear (SUDS) (c) Presence (IPQ)

Figure 11: Means and standard error of the PPQ, SUDS, and IPQ for each condition for Experiment 2. A * marks a significant effect.

4.2.2 Subjective Units of Distress Scale

In the SUDS, there was a significant main effect for the height ($F(1,30) = 23.90, p < .001, \eta_p^2 = .443$). The stress was significantly higher in the high-fear conditions. We found no significant main effect for the congruence ($F(1,30) = 0.29, p = .596, \eta_p^2 = .009$) and no significant interaction effect between the two conditions ($F(1,30) = 0.34, p = .567, \eta_p^2 = .011$).

4.2.3 Perceived Plausibility Questionnaire

In the PPQ, there was a significant main effect for the congruence ($F(1,30) = 6.24, p = .018, \eta_p^2 = .172$). The perceived plausibility was significantly lower in the incongruent conditions. We found no significant main effect for the height ($F(1,30) = .42, p = .522, \eta_p^2 = .014$). We also found a significant interaction effect between the two conditions ($F(1,30) = 5.97, p = .021, \eta_p^2 = .166$).

Post-hoc tests using the Holm correction showed that the significant effect is between the low-fear congruent and the low-fear incongruent condition ($p_{holm} = .016$), with a higher perceived plausibility in the low-fear congruent condition. There was no other significant interaction effect.

4.2.4 IPQ

We found significant main effects for both the congruence ($F(1,30) = 5.13, p = .031, \eta_p^2 = .146$) and the height ($F(1,30) = 12.18, p = .002, \eta_p^2 = .289$). We found no significant interaction effect between the two.

4.3 Discussion Experiment 2

The NASA-TLX did not show a significant difference between the conditions. We can, therefore, say that we eliminated the problem with the different difficulties between the conditions.

The visual height intolerance scale (vHISS) and the intolerance of uncertainty scale (IUS) both showed a significant effect on the SUDS. Therefore, both these factors are predictors of the fear participants felt during the high-fear condition. This is an expected effect as participants with a predisposition for fear of heights (vHISS) or a low tolerance for unknown events (IUS) are prone to experience more fear of heights in an unknown, virtual height environment.

We did not find an effect of the intolerance of uncertainty scale (IUS) on the perceived plausibility. We expected that a lower tolerance to uncertainty would make subjects more susceptible to the congruence manipulation as it made the objects less predictable. This does not appear to be the case.

On the positive side, it should be noted that the height conditions did not significantly differ in perceived plausibility. The high-fear condition of Experiment 2 thus does not appear to have any a priori incongruencies compared to Experiment 1. In Experiment 2, we can partly accept H1: "Incongruent object behavior, i.e., size changes, will reduce perceived plausibility." While the perceived plausibility was lower in the incongruent condition, this was only true for the low-fear condition. In the high-fear condition, we could not find a significant difference in the perceived plausibility. One explanation could be that participants' heightened fear overshadowed their perception of the incongruence. Compared to Experiment 1, the incongruencies did not affect the participants' ability to complete the task. It was, therefore, easier for participants to overlook the manipulated behavior.

As in Experiment 1, we can see that our height manipulation worked as the SUDS showed significant differences between the conditions. Compared to the low-fear condition, participants had higher stress ratings in the high-fear conditions. However, because we did not find significant differences in the perceived plausibility for the high-fear condition, we have to reject H2, "Participants' fear will be lower with incongruent object behavior." again. However, for different reasons as in Experiment 1. In Experiment 1, we found a significant effect for the SUDS between the congruent and the incongruent conditions. This was not the case in Experiment 2. Therefore, the plausibility manipulation did not influence the participants' fear. It seems to be more the case that the increased fear captivates the participants so much that the incongruencies play a minor role. This is why they have less of an effect on perceived plausibility.

We can accept our third hypothesis: "Higher presence will lead to a higher feeling of fear." We found a significant effect between the low-fear and the high-fear condition. Participants had higher presence scores in the high fear condition. Additionally, they had higher scores in the congruent conditions. This contradicts previous research from Brübach et al. [5] where a break in plausibility did not affect presence. The incongruent object behavior caused a lower

Height Plausibility	Low-Fear Congruent		Low-Fear Incongruent		High-Fear Congruent		High-Fear Incongruent	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NASA-TLX	12.97	14.84	14.30	15.79	16.21	16.09	16.05	16.31
VRSQ	0.13	0.18	0.15	0.19	0.17	0.22	0.17	0.19
SUDS	0.84	1.07	0.68	0.83	2.58	2.31	2.58	2.16
PPQ	4.32	0.42	4.04	0.43	4.12	0.38	4.12	0.49
IPQ	3.51	0.55	3.39	0.48	3.67	0.41	3.57	0.53

Table 3: Means and standard deviations of all questionnaires for each condition in Experiment 2.

feeling of presence in the participants.

5 Discussion

In our two experiments, we were able to see that the high-fear conditions did trigger higher stress and, thus, higher fear. Therefore, we can say that the fear manipulation worked for both experiments.

In the first experiment, we could see that the perceived plausibility was lower in both the incongruent and high-fear conditions. They did not, however, influence each other. However, the first experiment had confounds, which we discussed earlier. Therefore, we carried out the second experiment to eliminate the unwanted a priori incongruence of the high-fear condition present in Experiment 1. We, therefore, chose a more believable way to show virtual height. Additionally, we changed the congruence manipulation. In this experiment it was intended to be conspicuous but not to influence the task's difficulty between conditions.

Controlling for these two factors has led to different results than in Experiment 1. Perceived plausibility and fear seem to influence each other, as we found an interaction effect between them. We can see that the congruence manipulation worked as we can see a significant difference in the low-fear conditions. Results showed that the stress ratings in the high-fear condition were independent of the congruence manipulation. In the high-fear conditions, the fear seemed to overshadow the incongruence. The participants' mental capacities seem to be fully utilized by the fear that they no longer perceive the incongruencies caused by the object behavior. This shows us that slight incongruities should not influence the success of VRET.

Considering Lin's results, our findings are surprising. However, there are clear differences between their work and ours. While they directly compared Psi to PI elements, we focused on plausibility alone. Their Psi elements would be on the cognitive level of the CaP model and the PI elements on the perceptual level. However, previous work has shown that the lower perceptual layer should have a stronger effect [5]. They also dealt with a different kind of fear. While we used a specific anxiety, namely fear of heights, they used a broader fear triggered by a horror game. However, we can confirm the results of Brübach et al. [5, 6]. The manipulation of fear is on the lower perceptual layer of the CaP model. According to previous results, the perceptual layer has a stronger influence on the perceived plausibility than the cognitive congruence manipulation. This was also confirmed in our study, further strengthening the assumptions of the CaP model. This insight should be kept in mind when designing VRET. Incongruencies on the cognitive level, which do not prevent participants from fulfilling their task, appear to have a minor influence on the fear response. However, it can be assumed that incongruences from the perceptual level could have an influence. A greater focus should, therefore, be placed on avoiding incongruencies.

5.1 Limitations and Future Work

As discussed, there is a significant effect in the NASA-TLX in Experiment 1. This may partly be because sorting floating bottles is harder than sorting bottles with normal behavior. We were able to eliminate this effect in Experiment 2. We also changed the environment to get rid of any incongruencies it caused. Overall, there were quite a few changes between the two experiments. In the future, testing different incongruencies within the same environment would be very interesting. This could give better insights into which effect the strength of a congruence manipulation has.

The perceived plausibility questionnaire has not yet been fully validated. Additionally, we changed the wording of some of the questions compared to the original questionnaire. While it seems sensitive to congruence manipulations [5, 6, 36], it would be beneficial to validate this questionnaire further.

In the first study, we did not examine the role that presence could have played. In the second experiment, we found that the feeling of presence differs between the conditions.

Additionally, a larger sample size would be beneficial for the results to be more valid. In the future, looking at other congruence manipulations would also be useful. We used the behavior of objects with which participants interacted. While it makes sense to use a conspicuous manipulation for an initial study, it would be useful to manipulate more subtle aspects, such as the environment. Investigating other layers of the CaP model in the future would be especially interesting.

Another limiting factor is the composition of the sample. Two-thirds of the participants were female. A diverse sample would be desirable for future work.

6 Conclusion

With the rise of the use of VRET, it is necessary to understand the factors that influence these applications better. Previous research focused on the connection and causality between presence and fear, using immersion to manipulate PI and overall presence. However, recent discussions show that the perceived plausibility should not be neglected. We present two studies that manipulated semantic congruence to influence plausibility and height to influence fear. We kept the technological aspects consistent between the conditions. Our results show that the strength of an incongruency determines whether participants notice them in stress- or fearful situations. If the incongruence is too subtle or does not hinder the participant in fulfilling the task, it might be overshadowed by the fear. It is important to consider such results when designing future therapy applications. The perceived plausibility and behavior of the VR application should be adapted to the patient's expectations in order to avoid unwanted anxiety effects. Further research should be conducted in this area, especially with the strength of the congruence manipulations in mind. This is the only way to maximize patient safety and the success of the therapy.

Acknowledgments

This research has been funded by the Bavarian State Ministry For Digital Affairs in the project XR Hub (project number A5-3822-2-16).

REFERENCES

[1] C. Botella, J. Fernández-Álvarez, V. Guillén, A. García-Palacios, and R. Baños. Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports*, 19(7):42, 2017. doi: 10.1007/s11920-017-0788-4 1, 2, 3

[2] S. Bouchard, S. Côté, J. St-Jacques, G. Robillard, and P. Renaud. Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3d games. *Technology and health care*, 14(1):19–27, 2006. 1

[3] L. Brübach, M. Röhm, F. Westermeier, M. E. Latoschik, and C. Wienrich. Manipulating immersion: The impact of perceptual incongruence on perceived plausibility in vr. In *23rd IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE Computer Society, 2024. Accepted for publication. 2, 3

[4] L. Brübach, M. Röhm, F. Westermeier, C. Wienrich, and M. E. Latoschik. The influence of a low-resolution peripheral display extension on the perceived plausibility and presence. In *30th ACM Symposium on Virtual Reality Software and Technology (VRST)*. ACM, 2024. Accepted for publication. doi: 10.1145/3641825.3687713 2, 3

[5] L. Brübach, F. Westermeier, C. Wienrich, and M. E. Latoschik. Breaking Plausibility Without Breaking Presence - Evidence For The Multi-Layer Nature Of Plausibility. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2267–2276, May 2022. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2022.3150496 2, 3, 5, 8, 9

[6] L. Brübach, F. Westermeier, C. Wienrich, and M. E. Latoschik. A systematic evaluation of incongruencies and their influence on plausibility in virtual reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 894–901. IEEE, Sydney, Australia, 2023. doi: 10.1109/ISMAR59233.2023.00105 2, 3, 9

[7] P. L. Chard Ian, Van Zalk Nejra. Virtual reality exposure therapy for reducing social anxiety associated with stuttering: the role of outcome expectancy, therapeutic alliance, presence and social presence. *Frontiers in Virtual Reality*, 4, 2023. doi: 10.3389/frvir.2023.1159549 1, 2, 6

[8] C. M. Coelho, A. M. Waters, T. J. Hine, and G. Wallis. The use of virtual reality in acrophobia research and treatment. *Journal of Anxiety Disorders*, 23(5):563–574, 2009. doi: 10.1016/j.janxdis.2009.01.014 2

[9] A. Francová, M. Jablonská, and I. Fajnerová. Design and Evaluation of Virtual Reality Environments for Claustrophobia. *PRESENCE: Virtual and Augmented Reality*, 32:23–34, 12 2023. doi: 10.1162/pres_a_00385 1

[10] M. H. Freeston, J. Rhéaume, H. Letarte, M. J. Dugas, and R. Ladouceur. Why do people worry? *Personality and Individual Differences*, 17(6):791–802, 1994. doi: 10.1016/0191-8869(94)90048-5 7

[11] A. L. Gerlach, T. Andor, and J. Patzelt. Die bedeutung von unsicherheitsintoleranz für die generalisierte angststörung modellüberlegungen und entwicklung einer deutschen version der unsicherheitsintoleranz-skala. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 37(3):190–199, 2008. 7

[12] D. Gromer, M. Reinke, I. Christner, and P. Pauli. Causal interactive links between presence and fear in virtual reality height exposure. *Frontiers in Psychology*, 10:141, 2019. doi: 10.3389/fpsyg.2019.00141 1, 2

[13] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, Human Factors and Ergonomics Society., -, 2006. doi: 10.1177/154193120605000909 4

[14] D. Huppert, E. Grill, and T. Brandt. A new questionnaire for estimating the severity of visual height intolerance and acrophobia by a metric interval scale. *Frontiers in Neurology*, 8, 2017. doi: 10.3389/fneur.2017.00211 3

[15] H. K. Kim, J. Park, Y. Choi, and M. Choe. Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment. *Applied ergonomics*, 69:66–73, 2018. 7

[16] K. Kim, C.-H. Kim, K. R. Cha, J. Park, K. Han, Y. K. Kim, J.-J. Kim, I. Y. Kim, and S. I. Kim. Anxiety provocation and measurement using virtual reality in patients with obsessive-compulsive disorder. *CyberPsychology & Behavior*, 11(6):637–641, 2008. doi: 10.1089/cpb.2008.0003 1, 2

[17] P. Kourtesis, J. Linnell, R. Amir, F. Argelaguet, and S. E. MacPherson. Cybersickness in virtual reality questionnaire (CSQ-VR): A validation and comparison against SSQ and VRSQ. *Virtual Worlds*, 2(1):16–35, 2023. doi: 10.3390/virtualworlds2010002 4

[18] M. Krijn, P. M. Emmelkamp, R. Biemond, C. De Wilde De Ligny, M. J. Schuemie, and C. A. Van Der Mast. Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behaviour Research and Therapy*, 42(2):229–239, 2004. doi: 10.1016/S0005-7967(03)00139-6 1, 2

[19] M. E. Latoschik and C. Wienrich. Congruence and Plausibility, Not Presence: Pivotal Conditions for XR Experiences and Effects, a Novel Approach. *Frontiers in Virtual Reality*, 3:694433, June 2022. doi: 10.3389/frvir.2022.694433 1, 2, 3

[20] J.-H. T. Lin. Fear in virtual reality (VR): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game. *Computers in Human Behavior*, 72:350–361, 2017. doi: 10.1016/j.chb.2017.02.057 1, 2, 3, 5

[21] Y. Ling, W.-P. Brinkman, H. T. Nefs, C. Qu, and I. Heynderickx. Effects of stereoscopic viewing on presence, anxiety, and cybersickness in a virtual reality environment for public speaking. *Presence: Teleoperators and Virtual Environments*, 21(3):254–267, 2012. doi: 10.1162/PRES_a_00111 1, 2, 3

[22] Y. Ling, H. T. Nefs, N. Morina, I. Heynderickx, and W.-P. Brinkman. A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PloS one*, 9(5):e96144, 2014. doi: 10.1371/journal.pone.0096144 1, 2, 3

[23] H. M. Peperkorn, J. Diemer, and A. Mühlberger. Temporal dynamics in the relation between presence and fear in virtual reality. *Computers in Human Behavior*, 48:542–547, 2015. doi: 10.1016/j.chb.2015.02.028 1, 2, 4

[24] J. S. Pereira, L. M. Faêda, and A. M. Coelho. Evolution of vret to assist in the treatment of phobias: a systematic review. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, pp. 386–390, 2020. doi: 10.1109/SVR51698.2020.00064 2

[25] M. Price and P. Anderson. The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders*, 21(5):742–751, 2007. doi: 10.1016/j.janxdis.2006.11.002 1, 2, 6

[26] T. Schubert, F. Friedmann, and H. Regenbrecht. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, June 2001. doi: 10.1162/105474601300343603 7

[27] M. Schuemie, M. Bruynzeel, L. Drost, M. Brinckman, G. De Haan, P. Emmelkamp, and C. Van der Mast. Treatment of acrophobia in virtual reality: A pilot study. In *Conference Proceedings Euromedia*, pp. 271–275, 2000. 1, 2

[28] R. Skarbez, F. P. Brooks, Jr., and M. C. Whitton. A Survey of Presence and Related Concepts. *ACM Computing Surveys*, 50(6):1–39, Nov. 2018. doi: 10.1145/3134301 2

[29] R. Skarbez, S. Neyret, F. P. Brooks, M. Slater, and M. C. Whitton. A Psychophysical Experiment Regarding Components of the Plausibility Illusion. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1369–1378, Apr. 2017. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2017.2657158 2

[30] R. T. Skarbez. *Plausibility illusion in virtual environments*. PhD thesis, The University of North Carolina at Chapel Hill, 2016. 2

[31] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557, Dec. 2009. doi: 10.1098/rstb.2009.0138 2

[32] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A Separate Reality: An Update on Place Illusion and Plausibility in Virtual Reality. *Frontiers in Virtual Reality*, 3:914392, June 2022. doi: 10.3389/frvir.2022.914392 1, 2

[33] H. Song, S. Park, H. Kim, S. Jo, J.-I. Lee, S.-J. Han, I. Choi, and J.-

Y. Lee. Is anxiety-inducing vr experienced differently depending on personality? the mediating role of presence. *IEEE Access*, 9:42161–42168, 2021. doi: 10.1109/ACCESS.2021.3064251 1, 2

[34] D. Vansteenwegen, C. Iberico, B. Vervliet, V. Marescau, and D. Hermans. Contextual fear induced by unpredictability in a human fear conditioning preparation is related to the chronic expectation of a threatening us. *Biological psychology*, 77(1):39–46, 2008. doi: 10.1016/j.biopsycho.2007.08.012 5

[35] D. Villani, C. Repetto, P. Cipresso, and G. Riva. May i experience more presence in doing the same thing in virtual reality than in reality? an answer from a simulated job interview. *Interacting with Computers*, 24(4):265–272, 2012. doi: 10.1016/j.intcom.2012.04.008 1

[36] F. Westermeier, L. Brübach, M. E. Latoschik, and C. Wienrich. Exploring Plausibility and Presence in Mixed Reality Experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2680–2689, 2023. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2023.3247046 2, 9

[37] F. Westermeier, L. Brübach, C. Wienrich, and M. E. Latoschik. Assessing depth perception in vr and video see-through ar: A comparison on distance judgment, performance, and preference. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2140 – 2150, 2024. doi: 10.1109/TVCG.2024.3372061 2

[38] J. Wolpe. Subjective units of distress scale. *Journal of EMDR Practice and Research*, 1969. 4