Kamilla Sanner
Ola Dyrdal

Virtual Reality Exposure Therapy (VRET) For Fear of Spiders: a Randomized study

Graduate thesis in Clinical Psychology
Supervisor: Stian Solem
December 2021

Graduate thesis

NTNU
Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Psychology

NTNU
Norwegian University of
Science and Technology

Kamilla Sanner
Ola Dyrdal

Virtual Reality Exposure Therapy (VRET) For Fear of Spiders: a Randomized study

Graduate thesis in Clinical Psychology
Supervisor: Stian Solem
December 2021

Norwegian University of Science and Technology
Faculty of Social and Educational Sciences
Department of Psychology

NTNU
Norwegian University of Science and Technology

Preface

Work on the thesis started with a shared interest in using VR for exposure therapy. We were contacted by Fornix, a student company at NTNU developing VRET programs for common phobias, wanting us to participate in developing and testing their arachnophobia VRET program. They were at an early stage of developing the program, and we started reading relevant literature to give input on how the scenarios and spiders should look like from a psychological stance. Together, we advanced five VRET scenarios designed for fear of spiders. The thesis design, hypotheses, literature review, testing, analyses and interpretation of the results were done by the authors in collaboration with our mentor. After composing the thesis design and hypotheses, we formulated an application to the Norwegian Centre of Research Data (NSD). The final approval was given in the middle of April. While waiting for approval, we completed a literature search and acquired an overview of the VRET research field.

Testing was demanding, as we needed a lot of space (3x3 meters) and the possibility of borrowing and transporting a tarantula, which required extensive planning before every testing session. Due to Covid-19, we also had to consider infection prevention and needed to disinfect the equipment between all participants to ensure safe testing. To obtain a considerable sample, the data collection demanded both a lot of time and effort. Composition of the graduate thesis, as well as analyses and interpretation of the results started in August. The total process has given us a greater insight in the role of VR in exposure therapy, as well as how to conduct clinical research. The amount of work invested in the project could not have been realized without the equal contribution of both parties.

We are especially grateful to our mentor Stian Solem for his invaluable input, effort and availability throughout the process. This thesis could not have been completed without his guidance and support. In addition, we would like to thank Fornix for lending us the VR equipment and including us in the development of the VRET program. We are especially grateful to Are Åberg for lending us his tarantula, Kai Arne, making it possible to complete our research design as desired. Furthermore, we would like to thank all participants willing to contribute to the research, by using their time and challenging their fears. Finally, we would like to thank our family, partners, and fellow students for contributing with support over the last year.

Abstract

Background: The use of Virtual Reality (VR) in psychological research has increased since the nineties, especially when studying anxiety disorders. For arachnophobia, virtual reality exposure therapy (VRET) might encourage more people to seek therapy. VRET seems to be about as efficient as in-vivo exposure for inducing fear. Still, methodological issues, a limited number of studies, and a rapid changing technology complicates conclusions about effectiveness of VRET. Newer VR technology has promising implications due to accessibility, lower prices, and more user-friendly devices.

Objective: This study tested whether a VRET intervention would be able to induce and reduce discomfort in participants, if it was experienced as immersive, and if participants in the VR condition were able to approach a real-life tarantula to a greater extent than participants in the control group.

Method: Forty-seven undiagnosed participants were assigned to two conditions: a VR condition and a No VR condition. Before testing, participants' severity of arachnophobia was assessed using the Spider Phobia Questionnaire-15. The VR condition completed five VR scenarios, where discomfort ratings were assessed, and the Presence Questionnaire thereafter. Both conditions then participated in a Behavior Avoidance Test (BAT), giving discomfort ratings during the test. The BAT involved approaching a real-life tarantula according to a specific sequence of steps.

Results: The VRET intervention induced and reduced discomfort ratings of the VR group. There was a significant difference between phobic and nonphobic participants in discomfort ratings throughout the scenarios. Overall, the participants experienced the VRET intervention as immersive. Both conditions obtained high scores on the BAT (8.7 vs. 8.8) and there was no indication that participants in the VR condition were able to approach a real-life spider to a greater extent than the No VR group. However, the VR condition reported lower levels of discomfort during the BAT.

Conclusion: Despite the non-significant findings between conditions, practical difficulties with using a real spider emphasized the potential use of VRET. As comparison to other relevant studies was challenging due to different assessment methods and reporting of results, new VRET studies would benefit from more standardization to ensure the validity of documented treatment effects.

Keywords: virtual reality (VR), exposure therapy, subjective discomfort, fear of spiders, arachnophobia, Behavioral avoidance test, immersion, presence.

Sammendrag

Bakgrunn: Bruken av virtuell virkelighet (VR) i psykologisk forskning har økt siden nittitallet, spesielt i studier av angstlidelser. Ved araknofobi kan virtuell virkelighet-eksponeringsterapi (VRET) muligens bidra til at flere personer oppsøker terapi. VRET virker å være omtrent like effektivt som in-vivo-eksponering når det gjelder fryktfremkalling. Likevel kompliserer metodologiske problemer, en begrenset mengde studier og en raskt endrende teknologi konklusjoner angående effektiviteten av VRET. Nyere VR-teknologi viser lovende implikasjoner grunnet tilgjengelighet, lavere priser og mer brukervennlige enheter.
Formål: Denne studien testet hvorvidt en VRET-intervensjon var i stand til å indusere og redusere ubehag hos deltakere, om den ville oppleves som involverende og om deltakerne i VR-betingelsen var i stand til å nærme seg en ekte tarantell i større grad enn deltakerne i kontrollgruppen.
Metode: Førtisju udiagnostiserte deltakere ble delt inn i to betingelser: en VR-betingelse og en Ikke VR-betingelse. Før testingen ble grad av araknofobi målt med The Spider Phobia Questionnaire-15. VR-betingelsen gjennomførte fem VR-scenarioer, der målinger av ubehag ble gjort, og deretter Presence Questionnaire. Begge betingelser deltok så i en atferdsunngåelsestest (BAT), hvor de ga målinger av ubehag gjennom testen. BAT testen innebar å nærme seg en ekte tarantell i en spesifisert rekkefølge av steg.
Resultater: VRET-intervensjonen induserte og reduserte ubehag i VR-gruppen. Det var en signifikant forskjell mellom fobiske og ikke-fobiske deltakere i ubehagsmålingene gjennom scenariene. Alt i alt, opplevde deltakerne VRET-intervensjonen som involverende. Begge betingelsene oppnådde høye skårer på BAT (8.7 vs. 8.8) og det var ingen indikasjon på at deltakerne i VR-betingelsen var i stand til å nærme seg den ekte edderkoppen i større grad enn Ikke VR-gruppen. Imidlertid rapporterte deltakerne i VR-betingelsen lavere nivå av ubehag underveis i BAT.
Konklusjon: Til tross for de ikke-signifikante funnene mellom betingelsene, så har praktiske utfordringer med å bruke en ekte edderkopp understreket den potensielle nytten med VRET. Da sammenligning med andre relevante studier var utfordrende grunnet forskjellige metoder og rapportering av resultater, så vil nye VRET studier vært tjent med mer standardisering for å sikre validiteten til dokumenterte behandlingseffekter.

Nøkkelord: virtuell virkelighet, eksponeringsterapi, subjektivt ubehag, frykt for edderkopper, araknofobi, atferdsunngåelsestest, immersjon, tilstedeværelse.

Introduction

The use of virtual reality technology (VR) has increased steadily in psychological research since the late 1990s as increasingly advanced digital tools have been developed. VR can be defined as technology that lets the user interact with a simulated environment in real time through several sensory modalities (Adamovich et al., 2009). The therapeutic potential of VR was soon applied in specially developed software for conditions like post-traumatic stress syndrome and anxiety disorders like specific phobias (Opriş et al., 2012). The gold standard for treating phobia has long been exposure therapy, and the similarity of VR to real-world experience was picked up as a promising new way to expose people with phobias to their feared stimuli, creating a new intervention known as virtual reality exposure therapy (VRET; Côté & Bouchard, 2005; Powers & Emmelkamp, 2008). Multiple studies have since documented VRET as a good treatment option for anxiety disorders (Côté & Bouchard, 2005; Garcia-Palacios et al., 2001; Lindner et al., 2020; Opriş et al., 2012). VRET has shown effectiveness across various assessment domains, including domain-specific subjective distress, general subjective distress, cognitive, behavioral, and psychophysiological measures (Powers & Emmelkamp, 2008).

Specific phobias

A specific phobia refers to an intense fear of a specific thing, situation, or activity (Bourne, 2011, p. 15). Arachnophobia is a specific phobia defined as a persistent fear of spiders, accompanied by an immediate anxiety response and avoidance when exposed to them (Garcia-Palacios et al., 2001). For many, a considerable part of the condition consists of anxiety about one´s own response when facing a feared stimulus – typically losing control or having a panic attack (Bouchard et al., 2006). People with strong phobic reactions tend to avoid situations where they might encounter their feared stimuli, which for arachnophobes often include basements, attics, or even entire countries (Lemelin & Yen, 2015). Specific phobias like arachnophobia can therefore come at a high cost for the person afflicted, by severely reducing their quality of life. In an extensive epidemiological study on specific phobias, Stinson et al. (2007) found prominent levels of comorbidity with other anxiety disorders, as well as a strong association with alcohol and substance dependence, personality disorders and bipolar II disorder. While little is known of the socio-economic costs of arachnophobia, there has been considerable focus on the economic impact of anxiety disorders, of which phobias constitute the majority (Craske & Waters, 2005). In 1990, it was estimated that anxiety disorders alone cost the US economy 46.6 billion dollars annually, with ¾ of these expenses caused by lost or reduced productivity (DuPont et al., 1996).

Virtual Reality Exposure Therapy (VRET)

Using VR in treatment of arachnophobia is especially relevant for several reasons. It is often challenging to find a suitable spider for in vivo exposure, due to shifting seasons in many countries and the general rarity of having spiders as pets. Finding a spider can therefore be a time- and resource-intensive process for the therapist. It could also be an advantage for the patients themselves, as a study by Garcia-Palacios et al. (2001) found that 81% of the participants high in fear of spiders strongly preferred VR exposure therapy compared to in vivo exposure therapy. The same study found that while 34% of the participants would "absolutely not" participate in in vivo exposure, only 8% answered the same for VR exposure. It can therefore be argued that by providing an opportunity to confront one´s fear through VR exposure, the treatment acceptability increases for a patient group characterized by avoidance (Rinck & Becker, 2007). However, the preference for VRET could also be seen as avoidance for the actual feared stimulus (Powers & Emmelkamp, 2008). Nevertheless, with only 15-20% of people with specific phobias ever seeking treatment, the availability of VR-based treatments might encourage more people to do so (Bouchard et al., 2006).

The possibility to control the environment used to expose for an activating stimulus is essential for exposure therapy. A VR-based approach is assumed to increase this control by customizing the exposure situation to the individual patient's needs and thereby achieving the best treatment effect possible (Adamovich et al., 2009; Bouchard et al., 2006; Lindner et al., 2020). Still, VR-based exposure therapy is not used in most ordinary treatment clinics. Lindner et al. (2020) states that reasons might be that the earlier generation VR equipment has been too heavy, too expensive, and in most cases required a high degree of technical skill to operate. In addition, VR can be problematic concerning symptoms of cybersickness (a condition resembling seasickness) (Rebenitsch & Owen, 2016). However, VR technology has evolved significantly since its beginning (Lindner et al., 2020), now offering wireless head-mounted displays (HMD).

While state of the art equipment for immersive VR still can cost up to 30.000 dollars (Minns et al., 2019), cheaper equipment is now also commercially available for the price of about 300 dollars. This has made VR technology much more available to the public and the equipment also gives the possibility of free movement within the virtual environment. This might prevent cybersickness. At the same time, Chen et al. (2021) argued that free movement might increase the likelihood of cybersickness due to larger visual turning angles. Moreover, recently developed hand tracking systems enable the use of hands and fingers to control the programs, which frees participants from the traditional stationary computer or hand-held controllers. Another advantage posed by VR is the possibility of exposure to feared stimuli from the safe confines of one’s home. This could have an impact on public health by making evidence-based treatment easily accessible to an unprecedented number of people (Lindner et al., 2020).

The role of presence and immersion in VRET

The level of realism in VR and its effect on presence for the user can have some interesting and sometimes problematic consequences, as is evident in the uncanny valley effect (Lindner et al., 2017). The uncanny valley refers to a phenomenon in digital settings, where virtual objects can produce discomfort and revulsion in the user if they are seen as almost identical to a real human, but not perfectly so (Lindner et al., 2017). Therefore, programs designed for VR should be wary of the graphic detail level and work better with semi-realistic human representations if they cannot achieve photorealistic ones. While this effect has been extensively reported in virtual human representations, there is less research on how this affects perception of virtual animals, and no studies specific to spiders. One study by Schwind et al. (2018) found support for the effect extending to virtual cats, and identified facial expressions, body pose and violation of naturalness as the major factors affecting human reactions.

To measure the similarity of VR to real-life experiences it is common to investigate two related concepts, known as presence and immersion. While often considered the same and used synonymously, they do theoretically refer to distinct aspects of VR. Slater and Wilbur (1997, p. 4) defined presence in a virtual environment as a state of consciousness, meaning a subjective experience of the user. Cummings and Bailenson (2016, p. 274) build on Slater and Wilbur’s distinction of the two and describe immersion on the other hand as “a quality of the system´s technology, an objective measure of the extent to which the system presents a vivid virtual environment”. Further they go on to sum up the definition of the two as: “immersion—defined as a technological quality of media—and presence—defined as the psychological experience of ‘‘being there”.” (Cummings and Bailenson, 2016, p. 273). As such, immersion can be regarded as a subcomponent of the larger presence construct, and is found to have a medium-sized effect on presence (Cummings & Bailenson, 2016). While regarded as two different measures, one as the objective technical qualities of the VR program and the other as a subjective feeling, they are undeniably intertwined and often used interchangeably in research on the topic. We therefore also use the word “immersive” in our study while referring to measured levels of presence, since we consider it an appropriate description of the subjective feeling of presence.

Research on VRET for arachnophobia

Several studies have looked specifically at VRET as a treatment intervention for arachnophobia since the technology first became available. One of the first clinical studies documented VRET as an effective intervention, finding significant improvement in 83% of the VRET group compared to 0% of the waiting list control group (Garcia-Palacios et al., 2002). This corresponds well with later meta-analysis finding that VRET outperform inactive control groups and are highly efficient in inducing and reducing fear (Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008). Hoffman et al. (2003) added tactile cues, making the virtual spider touchable and found that the participants in this group both went closer to the real spider afterwards and did so with less anxiety than the VRET group without tactile cues. This hints at presence as important for inducing fear, which is considered essential to the extinction process in exposure therapy. Further investigating the presence-fear connection, Peperkorn et al. (2016) found that the representation of a virtual hand perceived as the participants own, modulated fear reactions in a clinical sample.

Studies comparing VRET to the traditional in-vivo method for arachnophobia have yielded mixed results. Miloff et al. (2019) found VRET non-inferior in a comparison study, but only at 3- and 12-months follow-up, indicating that the VRET group continued to improve more over time. This seems in line with another Swedish study, which showed that after one session, participants with in vivo exposure showed better results at approaching a spider than participants in virtual reality exposure (Lindner et al., 2020). The difference decreased after several sessions, which suggests that repeated sessions might be more important for treatment effect in VRET than in vivo exposure. The claim of non-inferiority has been substantiated, but also challenged by newer studies and meta-analyses on the growing body of research in the recent decade. Especially meta-analytical studies have given mixed and inconsistent findings showing both inferiority, non-inferiority, and even slight superiority (Opriş et al., 2012; Wechsler et al., 2019). The consensus, however, seems to be that VRET is about as efficient as in-vivo for inducing fear, and that the uncertainty of inferiority stems from methodological issues with the limited number of studies in question (Opriş et al., 2012; Parsons & Rizzo, 2008).

Much of the research on VRET uses the same measures such as Spider Phobia Questionnaire (SPQ), Behavioral Avoidance Test (BAT), Subjective Units of Distress (SUDS) and Presence Questionnaire (PQ). Yet, uncertainty on what variables need to be examined and a lack of uniform reporting on these often causes trouble when comparing VRET studies (Parsons & Rizzo, 2008). Non-standardized instruments such as BAT and SUDS might be used with different scales in different studies. There are also two different questionnaires both commonly used for measuring fear of spiders, the Fear of Spiders Questionnaire (FSQ) and the more commonly used SPQ. SPQ also has an abbreviated version with fewer items (Olatunji et al., 2009). The same issue applies to questionnaires used for measuring presence in VR, with Igroup Presence Questionnaires (IPQ) and Presence Questionnaire (PQ) being of the most frequently used. However, many studies utilize customized versions of these, and Ling et al. (2014) problematize the lack of research on the validity of all presence scales (Cummings & Bailenson, 2016; Morina et al., 2015). Some studies use single-session design, while others have as many as eight treatment sessions (Michaliszyn et al., 2010), potentially leading to different treatment outcomes.

Another essential difference between most of the VRET studies is the use of different hardware and software, mirroring the rapid pace of technological change (Lindner et al., 2020). Studies also differ in choice of research design. Randomized control trial (RCT) design has long been considered the gold standard for treatment intervention studies (Ginsburg & Smith, 2016, p. 1). While many VRET studies utilize RCT design, it is often of an insufficient standard (Parsons & Rizzo, 2008) and Morina et al. (2015) recommend stronger methodology in future studies to be able to draw clearer conclusions about treatment efficiency. Randomizing into different treatment conditions also divides already small samples into smaller groups, with some experimental conditions reporting as few as 12 participants (Hoffman et al., 2003). This does not reach the minimal standard calculated by Parsons and Rizzo (2008), that for a VRET study to have adequate power to detect an effect, the treatment condition should consist of at least 30 participants.

Objective of the study

This study used an undiagnosed sample and assumed various degrees of phobic reactions among participants. It therefore did not aim to show the efficacy of the program as a treatment intervention on people with arachnophobia. It did, however, focus on presence and subjective discomfort and seek to establish whether the intervention could induce this in the participants regardless of their comfort level with spiders. This study stands out with its use of completely new and unstudied software. It also utilizes new hardware in the form of commercially available equipment, increasing the potential reach of the intervention to a patient group that is inherently avoidant of treatment.

This study tested the following hypotheses:

1. The VRET intervention will be able to both induce and reduce discomfort in the participants.
2. The VRET intervention will be experienced as immersive.
3. Participants in the VR condition will be able to approach a real-life tarantula to a greater extent than participants in the control group.

Method

Participants and procedure

The study protocol was approved by the Norwegian Centre for Research Data (NSD, reference number: 717189). The respondents did not receive any incentives for partaking in the study. Data collection took place from 22nd of April to 23rd of September 2021 at the Norwegian University of Science and Technology in Trondheim. To recruit participants for the study, we asked students and acquaintances using snowball sampling, and we posted on a student Facebook group searching exclusively for participants feeling discomfort towards spiders. The inclusion criteria specified that participants had to be at least 18 years old and consent to participate. There were no exclusion criteria. The final sample consisted of 47 consenting adults (22 men, 24 women and 1 “other”). The exact age of the participants was not obtained, due to NSD wanting the participants to remain anonymous in the data set. Instead, we requested their age category – with all participants fitting in the 18-24 and the 25-29 categories.

The participants were randomly assigned to one of two conditions (VR/No VR), using www.randomizer.org. We initially used a randomization list for 40 participants and thereafter another list for 10 participants. No variables were controlled for when randomizing. In the VR condition, the participants tried five VR scenarios. Distress ratings were recorded at peak level and before switching scenarios. After completing the scenarios, they were given a behavioral avoidance test (approaching a spider). Participants in the second condition (No VR) completed the avoidance test but had no VR training before starting. Participants in the

VR condition used approximately 1 hour (including filling out forms) compared to 30 minutes for the No VR condition.

Measurements

The Behavioral Avoidance Test (BAT) (Muris et al., 1998; Öst et al., 1991) was used for evaluating avoidance of spiders. To be able to measure the effect of the VR exposure, it was considered suitable to use BAT. The test consisted of the participant entering a room with a spider on a table approximately three meters away. The spider, a female Chilean rose haired tarantula with the diameter of about 13 cm, was inside a lidded glass cage. The Chilean rose haired tarantula has a docile temperament and is not particularly venomous (Van Vlierberghe, n.d.). Before starting BAT, the test leader informed the participants about the steps, that they could stop whenever they wanted, and that SUDS ratings (see next paragraph) would be requested throughout the test. The BAT instructions were written down and maintained across all tests. The participant was asked to approach the spider until they did not want to get any closer. The test leader did not encourage the participant to go any further in the BAT, and therefore had a passive role in the testing.

BAT was rated on a scale from 1 to 10. Step 1 involved the participant being at least ¼ of the length from the spider, step 2 involved being at least ½ of the length from the spider and step 3 involved being at least ¾ of the length from the spider. Step 4 encompassed the participant touching the glass cage. Step 5 entailed lifting the lid off the cage. Step 6 involved touching any place inside the cage with at least one finger. Step 7 contained touching the spider. Here, the test leader advised the participant to touch the spider´s back, because the spider then moved away from you. After step 7, the participants were asked if they wanted the test leader to place the spider on the table outside of the cage. If they agreed to this, step 8 was to touch the spider again once it was outside the cage. Step 9 and 10 both involved holding the spider, either for at least 10 seconds (9), or for at least 20 seconds (10). The participants were given a score according to the step they completed before wanting to stop the BAT.

The Subjective Units of Distress Scale (SUDS) (Wolpe, 1969) was used to measure intensity of stress or nervousness when meeting a stress triggering stimulus. It is a self-assessment instrument with a scale from 0 to 100, where 0 indicates *no distress/totally relaxed* and 100 indicates *highest anxiety/distress you have ever felt*. The instrument is subjective, and its purpose was to give the test leader some insight to the participant´s experience during the exposure situation. SUDS was utilized both throughout the VR scenarios and BAT. When starting the exposure, the test leader asked the participant to scale their discomfort. The test leader continued to ask the participant for SUDS frequently during both exposures.

The Spider Phobia Questionnaire 15 (SPQ-15) (Olatunji et al., 2009) is a questionnaire with 15 questions that is used to measure fear of spiders. The questionnaire consists of 15 statements that are all categorical, either true or false. Olatunji et al. (2009) reported that the SPQ-15 is found to be sensitive to effects of exposure therapy and it is therefore often used combined with BAT in research designs. The internal consistency of SPQ-15 ranges between Cronbach´s α = .87-.95 in the early studies of Olatunji et al. (2009). However, no later studies using the questionnaire report the internal consistency. The SPQ-15 was translated to Norwegian by the authors.

There seems to be no published cut-off values for the SPQ-15 (Cowdrey & Walz, 2015), only for the original 31 item SPQ (Zsido, 2017; Åhs et al., 2011). Zsido (2017) reported 22-23 as a typical cut-off point for participants with diagnosed phobia on the original SPQ. Calculations based on SPQ-15 scores from phobic and nonphobic samples (Olatunji et al., 2009) yielded a value of 5.2 for clinical cut-off score (Jacobson & Truax, 1992). Therefore, participants with a total score of 6 or above were categorized as “phobic”, and the remaining participants as “nonphobic”. Twenty-one participants obtained a phobic score (11 [48%] in the VR condition and 10 [42%] in the No VR condition).

The Presence Questionnaire (PQ) (Witmer & Singer, 1998) is a questionnaire used to evaluate the degree of presence when using virtual technology. In this context, presence is defined as the subjective experience of being in a place – when one is physically located in another place (Witmer & Singer, 1998). An abbreviated version of the PQ (UQO Cyberpsychology Lab, 2004) was utilized in the study. The scoring procedure in this version was also applied, referring to the following factors: realism, possibility to act, quality of interface, possibility to examine and self-evaluation of performance. Some questions were excluded because the VRET program analyzed in the study did not include any sound effects. These include question 20, 21 and 22. Question 23 and 24 were also excluded, since the virtual environment did not entail a sense of touch. The remaining 19 questions were scored on a 7-point Likert scale with different endpoints on the response choices. Many of the scales ranged from (1) *not at all* to (7) *completely*, and another example was scales ranging from (1) *extremely artificial* to (7) *completely natural*. The internal consistency of this abbreviated version of the PQ (PQ-19) varies between Cronbach´s α = .81-.84 (Michaliszyn et al., 2010; UQO Cyberpsychology

Lab, 2004). We were not able to find other comparable Cronbach´s α, due to lack of reporting it and other PQ-versions utilized. Sagnier et al. (2020) reported Composite Reliability (CR) when measuring the internal consistency of the five factors, which is comparable to Cronbach´s α, with CR ranging from .67-.87. PQ-19 was, as SPQ, translated to Norwegian by the test leaders.

Interventions

The VRET software used in the study was developed by Fornix in Trondheim. The program operated with the VR hardware Oculus Quest 2, a wireless head-mounted display (HMD) (Holzwarth et al., 2021). It utilizes Oculus insight tracking, which computes an accurate and real-time position for the headset and controllers every millisecond (Hesch et al., 2019). This translates to the user´s precise movements into VR, which increases the experience of presence. With its internal processing power, the Oculus Quest 2 does not require the setup of auxiliary equipment to be operative (Holzwarth et al., 2021). The program utilizes a controller-free hand tracking feature, which enables the users to interact with and control the program using their hands. This means that the hands are visible within the simulation and move in a manner highly similar to real life. Certain hand signals are then used to navigate the program. The HMD had a low input lag, hypothetically leading to less cybersickness (Kim et al., 2020). Oculus Quest 2 offers high mobility, fast setup, and high quality.

The VR scenarios required an unobstructed movement area of 3x3 meters. Before starting the exposure, the participants were told what they will encounter in different scenarios. They were given instructions on how to attach the Oculus Quest 2, the HMD used in the study, and how to interact with the program. Furthermore, they were told that they could take a break (taking the HMD off) during the scenarios if they were feeling any motion sickness or too much discomfort during the exposure. The test leader was following the participant through the scenarios by streaming from the Oculus Quest 2 to their phone.

The first scenario was like a traditional BAT. The participant is in a room that is approximately 3 x 3 meters, with a tarantula in a glass cage a little less than three meters away (Figure 1). The tarantula is about 15 centimeters in diameter, with the purpose of it being easy to spot right from the beginning. The participants were asked to move towards the table in their own pace and were free to continue to the next scenario whenever they felt finished with the current scenario.

When entering the second scenario, the participant would stand in front of a table with a spider, smaller than the first spider, standing on the table (Figure 2). The participant was encouraged to touch the spider, and this caused the spider to move in the opposite direction of the touch. The spider walked towards the end of the table, where it would stop. However, the participant could touch the spider several times while moving, which led the spider to change direction according to where the participant touched it.

Before entering the third scenario, the participants were asked to sit down on a chair. When entering this scenario, the participant would have a simplified body in VR. This body was sitting down with the arms laid down on the sides of the thighs. The participant was told to position their own body to fit the virtual body. On the right knee, they would see a spider with the same looks as the previous spider (Figure 3). The spider walked slowly toward the virtual hip, and the participant was not able to interact with it. When arriving at the hip, the spider stopped. The participants were encouraged to try this scenario several times, until their SUDS rating declined. Then, when the participants wanted to proceed to the next scenario, they were asked to stand up before proceeding.

The fourth scenario had the same location as the second one, but the spider was smaller. The participant was encouraged to put their hand next to the spider, which led the spider to crawl onto the participant´s virtual hand. By placing the other hand next to the hand with the spider, the spider would start crawling onto the other hand (Figure 4). The participant was told which hand sign to do if the discomfort became too intense, which restarted the scenario – with the spider securely on the table.

The fifth scenario was in the same room as the other scenarios, but with several more items in the room (Figure 5). The table, as the first scenario, is about 3 meters away. Behind the table was a fireplace, with bookcases on both sides. There were also two stacks of cardboard boxes in front of the table, on both sides. In addition, parts of the walls, ceiling and the items mentioned were covered in spider webs. When entering the scenario, the participants would not be able to immediately spot any spiders. They were encouraged to explore the room to look for different spiders. Some spiders stood still in their spider webs, but most of them crawled around on the floor, ceiling, or the furniture. The participants were told that the spiders would not move in a surprising or frightening manner, and that the spiders would not react to them coming closer. When the participants were satisfied with the exploration, they were told to take off the HMD.

Figure 1

Inner Perspective of Scenario 1 (BAT)

Figure 2

Inner and Outer Perspectives of Scenario 2 (Touch Spider)

Figure 3

Inner and Outer Perspectives of Scenario 3 (Spider on thigh)

Figure 4

Inner and Outer Perspectives of Scenario 4 (holding Spider)

Figure 5

Inner Perspective of Scenario 5 (Exploring the room)

Data analyses

The collected data was analyzed with Statistical Package for the Social Sciences (SPSS), version 27. Independent samples T-test was conducted to determine if a difference existed between the means of the two groups (VR/No VR) on continuous dependent variables and whether this difference was significant. The variables tested were total fear of spiders and total score on BAT. Additionally, chi squares (χ^2) was calculated for the sex-, age- and phobic vs. nonphobic-variables to test whether the distribution of cases in these variables follows a known or hypothesized distribution. Cronbach´s alpha (α) was used to measure internal consistency in the questionnaires SPQ and PQ, as well as for the five factors in PQ. When interpreting Cronbach´s alpha (a value between 0 and 1), $\alpha < .5$ is seen as unacceptable, $.6 > \alpha \geq .5$ is considered poor, $.7 > \alpha \geq .6$ is seen as questionable, $.8 > \alpha \geq .7$ is considered acceptable, $.9\ \alpha \geq .8$ is seen as good and $\alpha \geq .9$ is considered excellent (George & Mallery, 2003, p. 231; Gliem & Gliem, 2003).

Pearson´s Correlations was utilized to assess the relationship between completion of exposure and immersion (all questions in PQ summed up), as well as the five factors in SPQ. Shapiro-Wilk´s test was used to test whether these variables were normally distributed. For correlations incorporating not normally distributed variables, the 95% bias corrected accelerated confidence intervals are reported. When interpreting r (the Correlations coefficient, a value between 0 and 1), a small correlation is defined as $.1 < r < .3$, a medium correlation is defined as $.3 < r < .5$ and a large correlation is defined as $r > .5$ (Cohen, 2013, p. 77-80).

One-way repeated measures ANOVA (analysis of variance) was conducted to determine if there was a difference, and whether this was significant, between the phobic and nonphobic group in the VR scenarios, as well as between the four groups (phobic VR, nonphobic VR, phobic No VR and nonphobic No VR) on BAT, when it came to SUDS ratings. The Greenhouse-Geisser estimate of the departure from sphericity (ε) was calculated as part of the one-way repeated measures ANOVA to correct the degrees of freedom of the *F*-distribution (Abdi, 2010). A Greenhouse-Geisser estimate of 1 means sphericity is met, whereas a Greenhouse-Geisser estimate less than 1 indicates sphericity not being met – and the degrees of freedom will be overestimated, while the F-value will be inflated.

Effect sizes was calculated using Cohen´s d (*d*) and partial eta squared (η^2). Cohen´s *d* was implemented when conducting the independent samples t-test, that is the effect of the relationship between the VR and No VR groups regarding total fear of spiders and completion of exposure. Cohen´s *d* is defined as the mean difference between two means divided by the pooled standard deviation (Cohen, 2013, p. 67). When interpreting Cohen´s *d*, Cohen (2013, p. 184-185) recommend .20 as a small effect size, .50 as a medium effect size and .80 as a large effect size. Partial eta squared (η^2) was implemented when conducting repeated measures ANOVA, and is defined as the proportion of variance that a variable explains when excluding other variables in the analysis (Richardson, 2011). It was therefore utilized to consider the effect of the relationship between the VR and No VR groups concerning SUDS ratings in completion of exposure, the relationship between the phobic and nonphobic groups in the VR condition relating to SUDS values, as well as for the relationship between the phobic and nonphobic groups in completion of exposure. When interpreting Partial eta squared (Cohen, 2013, p. 283-287), the rules of thumb are that .01 indicates a small effect, .06 signifies a medium effect and .14 indicates a large effect.

Two participants in the VR condition had missing values in their Presence Questionnaire. The missing values were replaced with the mean score for the remaining, answered questions. For one of the respondents, the BAT-score was missing due to the test being interrupted. The interruption involved the spider being hungry and therefore too aggressive to continue testing. It was not possible to resume the BAT another time, and it was consequently decided to set the score to the interruption point – a score of 6 on the BAT. This was considered valid because of a high SUDS score (90). To enable comparisons between participants in SUDS on BAT, we gave the SUDS value 100 for all remaining steps when participants did not complete all steps.

Results

Comparisons of the two conditions are presented in Table 1. The VR condition consisted of 23 participants and the No VR condition consisted of 24 participants. The distribution of males and females was equal for the two conditions. As measured with SPQ-15 (both ranging from 0-13 of a possible 0-15), participants in the VR condition reported discomfort for spiders with a mean score of 5.7 while the mean score of the participants in the No VR condition was 5.2 (no significant difference). SPQ-15 had high internal consistency ($\alpha = 0.88$) in this sample. Phobic participants (SPQ-score of 6 or more) made up 47.8 % of the VR group and 45.8 % of the No VR group. The two conditions did not contain a systematic age difference.

The total scores on BAT (the possible range was 1-10) for the two conditions were similar, with a mean score of 8.74 (*SD* = 1.42) for the VR condition and 8.79 (*SD* = 1.69) for the No VR condition. The difference was not significant, $t(45) = 0.12$, $p = 0.909$, $d = 0.03$. The high scores on the BAT indicates that most participants were able to approach the spider.

Table 1

Comparisons of the VR condition (n =23) and the No VR condition (n = 24)

	VR	*No VR*	t/χ^2	*p*
1. Male sex*	47.8 % (23)	45.8 % (24)	0.08	0.777
2. SPQ	5.74 (4.06)	5.17 (3.64)	-0.51	0.613
3. Phobic	47.8 % (23)	41.7 % (24)	0.18	0.671
4. Age			1.02	0.312
18-24	62.5% (23)	47.8% (24)		
25-29	37.5% (23)	52.2% (24)		
5. Total score on BAT	8.74 (1.42)	8.79 (1.69)	0.12	0.909

*Note. *"Other" sex was excluded in the table. SPQ = Total Fear of Spiders.*

Did the participants feel present in the virtual scenarios?

The Presence Questionnaire, consisting of 19 variables, had a Cronbach's α of .85 in this study. Scores (range 1-7) on the five factors of the PQ are described in Table 2. The participants were positive about all aspects of the VR experience as all scores were above the mid-point (4). The total presence score (ranging from 81-107 of a possible 19-133) had a mean of 92.50 (SD = 7.56). Self-evaluation of performance had the highest mean score of 5.77 (SD = 0.73), but also had an unacceptable internal consistency (α = .45). Realism was the factor with the lowest mean score (M = 4.81, SD = 0.72). The mean scores for the five factors were equal for the phobic and nonphobic group, as seen in Table 2.

Table 2

Mean item scores on the Presence Questionnaire

	Phobic	*Nonphobic*	α
Realism	4.93 (0.90)	4.71 (0.52)	.78
Possibility to act	5.53 (0.78)	5.56 (0.59)	.52
Quality of interface	5.29 (1.24)	5.27 (1.13)	.77
Possibility to examine	5.71 (1.03)	5.64 (0.59)	.61
Self-evaluation of performance	5.85 (0.65)	5.71 (0.81)	.45

Note. All PQ scales range from 1-7.

Pearson´s Correlations was run to assess the relationship between the total score on BAT and the total presence score, as well as the five factors in PQ. Not all variables were normally distributed, as assessed by Shapiro-Wilk´s test ($p < 0.05$). There was a statistically significant relationship between BAT and Realism, $r(23) = -0.47$, $p = 0.023$. The medium negative correlation between the two indicated that the participants experiencing the VR scenarios as more realistic completed less steps of BAT. There were no significant correlations between BAT and the other presence factors, nor with the total presence score.

A few participants removed the HMD during testing, two due to cybersickness and two due to experiencing overwhelming anxiety when confronting the scenarios. The cybersickness probably occurred because of problems with height glitches. These glitches were quickly fixed by the company, and no participants reported cybersickness afterwards. One participant with monocular vision commented surprisingly low levels of cybersickness compared to other VR software previously tested.

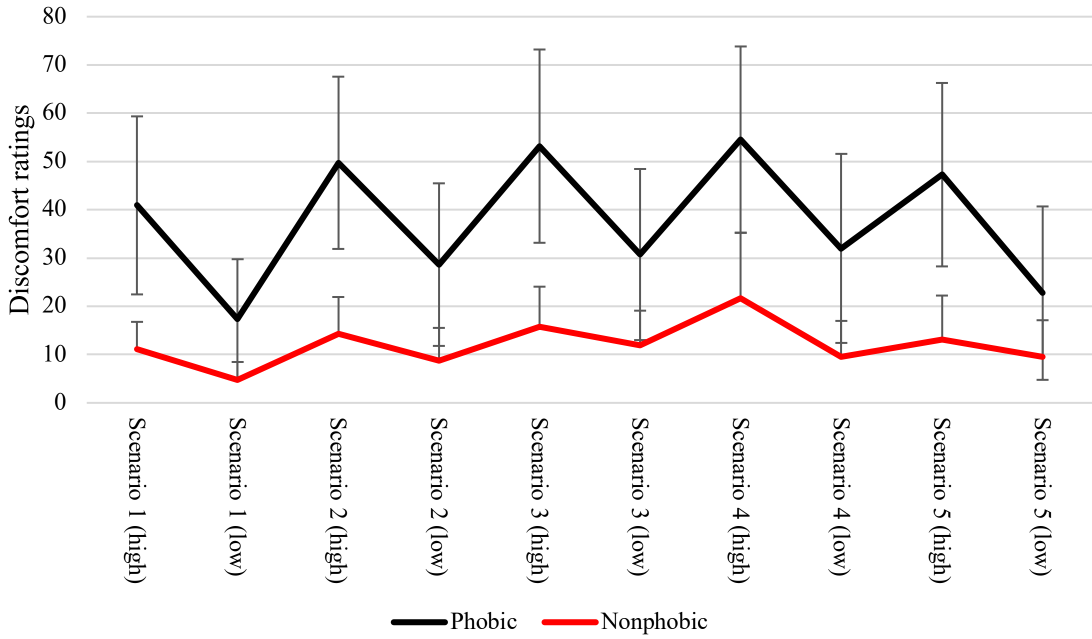
Changes in discomfort following the VRET intervention

The phobic (n = 11) and nonphobic (n = 12) participants in the VR condition was compared regarding SUDS ratings throughout the VR scenarios (Figure 6). The Greenhouse-Geisser estimate of the departure from sphericity was $\varepsilon = 0.43$. The difference was significant, $F(3.87,81.22) = 4.17$, $p = 0.004$, $\eta^2 = 0.17$. This indicates a large effect. As seen in Figure 6, the phobic participants score higher than the nonphobic participants – with nonoverlapping error bars with 95% confidence intervals for the highest scores on all scenarios except number 4. At the same time, the phobic participants did get overlapping scores with the nonphobic participants when it comes to the lowest scores in the scenarios, indicating that they also managed to get a reduction in discomfort in the scenarios. Scenario 4 is the only scenario with overlapping confidence intervals for the two groups, suggesting that this scenario provoked discomfort even in nonphobic participants.

When summing up the highest SUDS score from each scenario, the mean of the phobic group was 49.13 (*SD* = 31.37), and for the lowest score 26.29 (*SD* = 28.03). In contrast, the mean of the highest score for the nonphobic group on all scenarios was 15.22 (*SD* = 15.29), and for the lowest score 8.92 (*SD* = 11.30). The highest value in the nonphobic group throughout the scenarios was 80, documented in scenario 4. The highest value in the phobic group throughout the scenarios was 100, also documented in scenario 4.

Figure 6

SUDS ratings from all VR scenarios for phobic and nonphobic participants



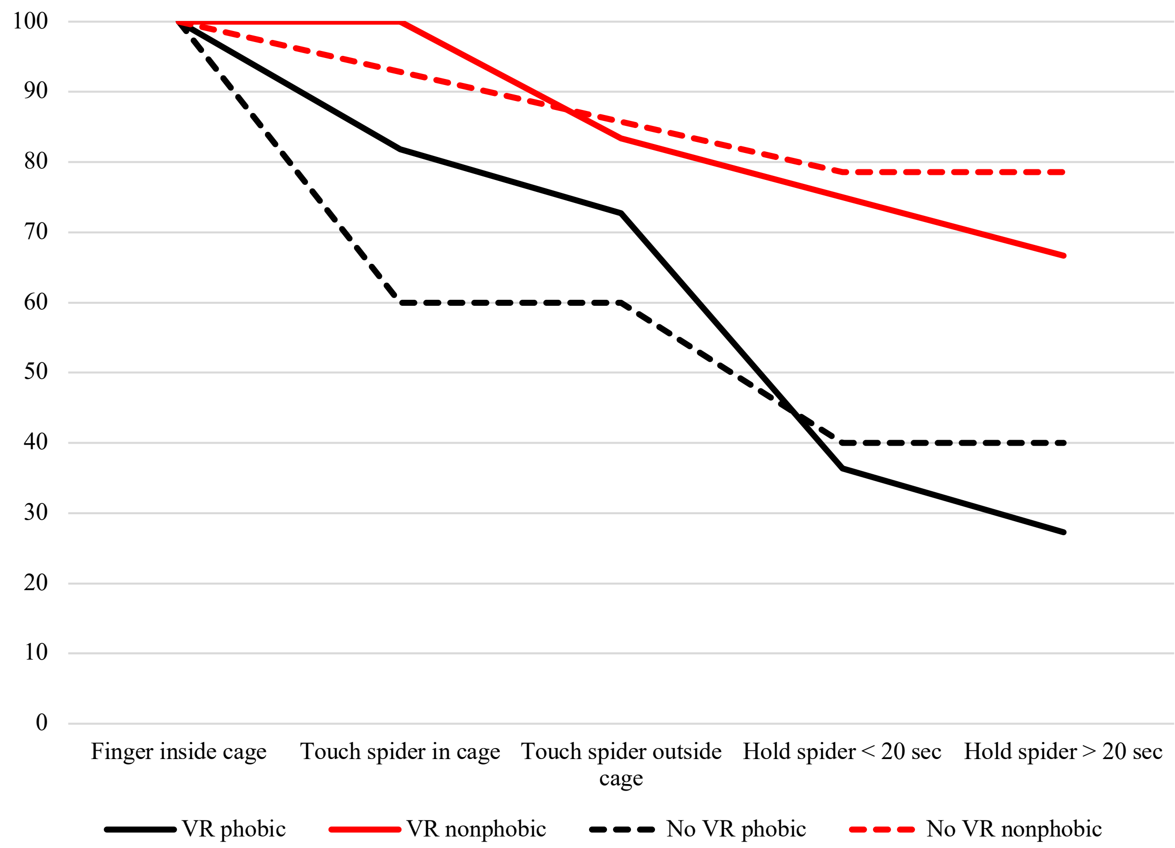
Note. *Scenario 1 = BAT, scenario 2 = touching spider, scenario 3 = spider on thigh, scenario 4 = holding spider, scenario 5 = exploring the room. With 95 % confidence intervals.*

Were participants in the VR condition able to approach a real-life tarantula to a greater extent than participants in the control group?

As previously mentioned in Table 1, both conditions obtained very high BAT scores (8.7 vs 8.8 on a 1-10 scale) and the difference was not significant. Exposure in VR was not associated with completing more steps on the BAT. As seen in Figure 7, the survival curves for phobic and nonphobic participants were similar irrespective of condition. There was a high percentage completing the entire BAT (held the spider more than 20 seconds) in both groups, constituting 47.8 % of the VR group and 62.5 % of the No VR group. The remaining participants completed 6-9 steps of BAT. A summary of the BAT steps completed for both conditions is displayed in Figure 7. In the VR condition, 20.8 % stopped after step 6 (one finger inside cage), whereas 8.7 % of the No VR condition stopped here. At step 7 (touched spider in cage), 4.2 % of the VR group and 13 % of the No VR group stopped the BAT. Next, 12.5 % of the VR group and 21.7 % of the No VR group stopped after step 8 (touched the spider outside the cage). Finally, 8.7 % of the No VR group, and none in the VR group, stopped after step 9 (held spider less than 20 seconds).

Figure 7

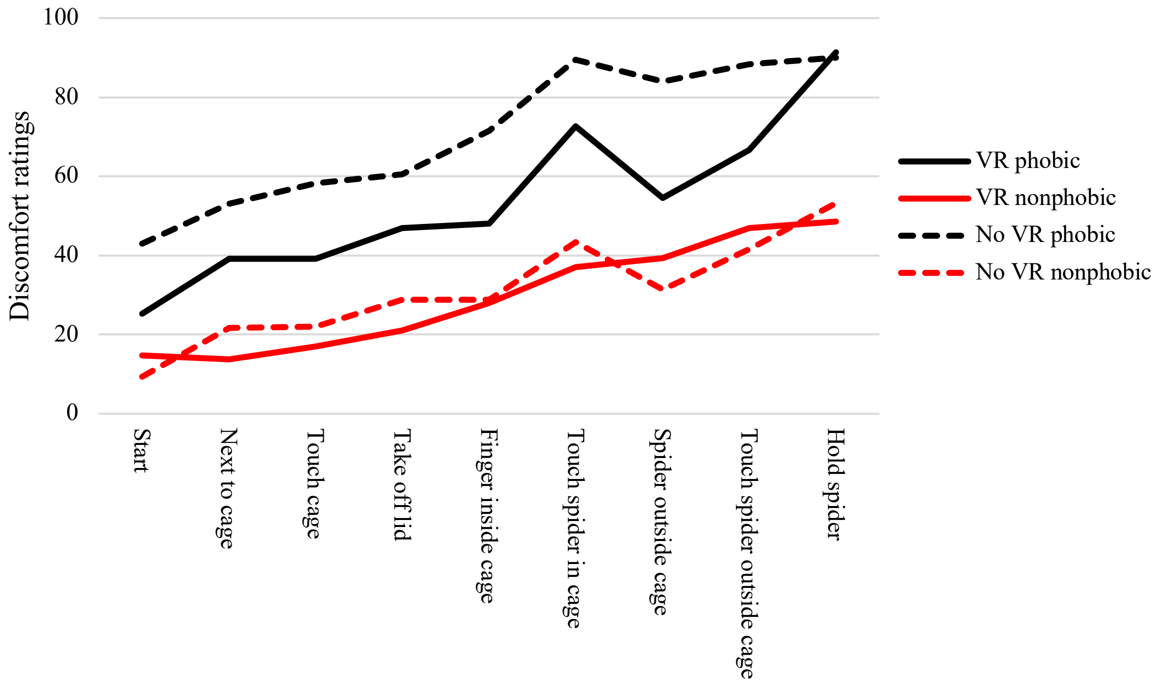
Percentage of steps conducted in BAT



As seen in Figure 8, the phobic groups reported lower SUDS ratings than the nonphobic groups throughout BAT. There was not a significant overall difference in SUDS ratings throughout the BAT for the four groups (phobic VR, nonphobic VR, phobic No VR and nonphobic No VR), $F(8.24,118.09) = 1.10$, $p = 0.370$, $\eta^2 = 0.07$. However, phobic participants in the VR group reported consistently lower SUDS ratings than phobic participants in the No VR group (except for the last step). There were significant differences in discomfort ratings between the two phobic groups when the spider was brought out of its cage (54.5 vs. 84.0, $p = .029$) and when touching the spider outside the cage (66.6 vs. 88.3, $p = .041$).

Figure 8

Visualization of SUDS Ratings Throughout BAT, VR vs. No VR and phobic vs. nonphobic



Discussion

The results showed that, in line with our first hypothesis, the intervention program was successful in both inducing and eventually reducing the participants' discomfort. A comparison between the nonphobic and phobic participants on the VR condition showed a significant difference in reported discomfort on SUDS during the VRET intervention. The VR participants also reported a level of presence that indicated that they experienced the program as immersive. A strong negative correlation was found between the perceived realism of the VRET intervention and steps finished in the BAT, suggesting that the participants categorized as phobic experienced higher degrees of immersion. There was no significant difference in the performance between the VRET and the control group on the BAT, implying that the VRET intervention did not help the participants approach the real-life spider more than they otherwise would. However, they did report lower levels of discomfort throughout the BAT.

Did the VRET intervention induce and reduce discomfort in the participants?

Like in-vivo exposure therapy, VRET depends on inducing the fear it is supposed to extinguish gradually. It is therefore crucial that VRET programs have properties that make them induce fear to a similar degree as the original stimuli to achieve treatment effect (Opriş et al., 2012). Other VRET studies on specific phobias have typically used SUDS as an indicator of when to progress to next exposure level without reporting the data (Bouchard et al., 2006; Krijn et al., 2004), making direct comparison with similar studies challenging. Based on the reported SUDS scores from the participants, as well as verbal feedback and observed test behavior, the VRET intervention was able to induce fear in the participants. Some reported such discomfort during the testing that they looked away from the virtual spiders or in a couple of cases removed the VR headset temporarily. A significant difference was found in this study between the phobic and nonphobic groups in the VRET intervention. This supports our first hypothesis as a stronger fear reaction for those categorized as phobic indicates that the VRET intervention invokes fear correlational to the participants' pre-measured fear of spiders in the SPQ. This is in accordance with earlier research showing good treatment effects in clinical samples (Côté & Bouchard, 2005; Garcia-Palacios et al., 2002; Lindner et al., 2020).

Upon investigating the phobic group within the intervention group, scenario three induced the second highest level of distress with a group average of 53 on the SUDS scale. This contrasted with an average score of 20 for the nonphobic group, showing that this scenario seemed to affect those high in fear of spiders more specifically than the others. As invasion of personal space is thought to be a considerable part of the fear in arachnophobia, it is likely that stationary positioning increases fear response in the participants (Lindner et al., 2017). This fits well with other research that has shown that the presence of virtual limbs modulates fear in people with arachnophobia (Peperkorn et al., 2016). The combination of a perceived virtual body and the complete lack of mobility may therefore have worked to create a particularly strong fear response in the phobic group.

Scenario four was the one that triggered the most distress in both groups. Since this scenario contained an interactive spider moving between the participants hands, it was one of the most technically challenging and hence the one most prone to technical "bugs". Earlier research in VRET interventions have highlighted cybersickness as a potential source of discomfort (Lindner et al., 2017). Since the test leaders used the words "fear" and "distress" synonymously while instructing the participants to rate their SUDS scores, it is possible that the increase in the nonphobic group is due partially to discomfort with these visual bugs. The relatively lower score on the subcategory "realism" from the PQ could also be interpreted to support such an explanation, as according to de Borst and de Gelder (2015) unnatural movement is considered to strongly affect perceived realism in virtual environments.

Central to exposure therapy is the theory that prolonged exposure to feared stimuli eventually causes the fear response to weaken and disappear in a process of extinction (Craske et al., 2008). In line with this theory, our hypothesis also stated that the SUDS scores would be reduced after repeated exposure to the virtual spider. The result from the SUDS showed that both groups had reduced scores in all scenarios after the first measurement. This indicates that the intervention program works as intended by also weakening the fear response it initiates in the participants. These results are in line with previous research showing that VRET as an intervention is effective in reducing fear, which is essential for transfer of learning to real-life situations (Lindner et al., 2020; Miloff et al., 2019). Given these findings, the program has the properties essential for exposure therapy to work.

Was the VRET intervention experienced as immersive?

Immersion has become an increasingly studied topic in VRET research, as it is thought to mediate whether the VR-environment engages “real emotions” or not (Peperkorn et al., 2015). As such, the second hypothesis of the study was that the VR program would be experienced as immersive. The results from the PQ show that all subcategories of the questionnaire except Realism had a mean score above 5, out of 7. Being above the mid-point, this supported our hypothesis as it indicated that the participants felt immersed and present in the virtual environment. This also matched the verbal feedback given by many participants during and after the testing. A related study on VR for spider phobia reported a mean total score of 84.21 (*SD* = 14.35) on the PQ, lower (*d* = .72) than the score of 92.50 (*SD* = 7.56) in this study. Our results were more comparable (*d* = .17) to a VRET study on social phobia (Parrish et al., 2016) that reported a mean PQ score of 89.61 (*SD* = 22.76). However, the large standard deviation in that study suggests more variation among participants’ experience of presence than in this study.

The results also showed a significant negative correlation between the subcategory “realism” and the participants total BAT score. This implies that the people who completed less steps on the BAT perceived the VRET intervention as more realistic but it does not tell us the direction of this relationship. Bouchard et al. (2008) found in a VRET study on people with snake phobia, that anxiety increased the feeling of presence in the VR environment. Interpreting the correlation based on this knowledge, it seems that the participants who finished the least number of steps in the BAT were the ones who also experienced the most anxiety in VRET, thus feeling more present. This becomes even more likely when looking at the differences between the phobic and nonphobic groups on number of steps in the BAT completed. As about 70% of the nonphobic group completed all the steps and only about 30% of the phobic group did the same, it seems likely that arachnophobia is the third variable mediating both the number of steps completed in the BAT as well as the perceived realism of the VRET intervention.

Several respondents expressed to the test leaders that they found questions in the PQ confusing or unclear. Cronbach's alpha values, however, indicated that the factors were reliable and almost identical to the one in the original validation for the PQ-19 (UQO Cyberpsychology Lab, 2004). Of the five factors investigated, only "Self-evaluation of performance" had a Cronbach's alpha value found to be at an unacceptable level of internal consistency ($\alpha = .44$). As there was no validated translation of the questionnaire to Norwegian, this was done by the authors, raising the question if this affected the validity of the presence results.

While realism was slightly lower rated than the other categories in the PQ it still is up to question whether high realism is necessary to induce fear and achieve a good treatment outcome (Miloff et al., 2019). The goal of high immersion in the environment is creating a reaction similar to that of real-life meetings with spiders (Lindner et al., 2017). High realism in virtual human characters, while not entirely photorealistic, could on the other hand end up in the "uncanny valley" (Lindner et al., 2017). This phenomenon causes unease and revulsion, and while it has been extensively studied in virtual human characters, as well as documented in virtual pets (Schwind et al., 2018), we do not yet know if this effect also extends to virtual animals such as spiders. Other research suggests that the essential predictor for perceived realism in VR is the degree of naturalism in the movement of the virtual characters, and not the level of detail in the graphic (de Borst & de Gelder, 2015).

Does VRET ease real life exposure to a spider?

Results showed that contrary to the expectation of our third hypothesis, the participants in the VR condition did not score higher on average in the BAT. The difference between the groups was found to be insignificant, indicating that the VRET intervention did not affect performance in the BAT. It is worth considering that the VRET intervention had different virtual spiders in its scenarios, with only the one in the first level resembling the tarantula used in the later BAT. Michaliszyn et al. (2010) refers to probable cultural differences in the expression of arachnophobia and recommends using local spiders in both VRET and BAT. It is possible that the VRET experience with the virtual spider did not generalize sufficiently to the real-world meeting with a spider of different appearance in this study. Another VRET study by Shiban et al. (2015) performed a BAT with two different spiders, one household spider, and one tarantula, with significant differences found only in the household spider condition. The same study also found multiple spiders and single context in VRET to give lower SUDS scores at post-test, indicating better treatment effect. Considering these points, it is possible the VRET design and the BAT with a tarantula hindered us in measuring any significant differences.

The BAT is still only a behavioral measure (Shiban et al., 2015), thought to be mediated by fear, but not a measurement of the experience of fear itself. And while the VRET group scored lower on the BAT than the control group, they did have a lower average SUDS score throughout the BAT. Though these differences were not significant, this could still indicate that the intervention did transfer into learning, causing the VRET group to feel less as they encountered the real spider. Interestingly, the SUDS scores of the two groups shift towards converging once the participants are required to touch the spider, the same point where more VRET participants stopped their BAT. This could be explained by the lack of tactile perception in the VRET scenarios involving touching the spider. Both Garcia-Palacios et al. (2002) and Hoffman et al. (2003) found tactile cues to strongly increase treatment effect in their studies. Meaning the absence of tactile cues in our study, as well as different visual appearance, could have been too different from the real-life BAT, rendering the results on these scenarios unaffected by learning from VRET.

Almost half of the VR group and over 60% of the No VR group finished the BAT test and the mean score for the two conditions on the BAT was 8.7 and 8.8. This could indicate a ceiling effect (Cramer & Howitt, 2004, p. 21), meaning scores on BAT were approaching the maximum they can be, and were therefore not being normally distributed in the two conditions. In other words, reducing the sensitivity of the BAT, that was used to determine if the average of the two groups were significantly different in spider approaching behavior. At worst, the ceiling effect may lead to the mistaken conclusion that the independent variable has no effect. One way to avoid this could have been a time limit on each step of the BAT, for instance of one minute like in a study by Olatunji et al. (2009).

In traditional exposure therapy, having the patient observe the therapist perform the exposure behaviors themselves without displaying a phobic reaction is a powerful, if not essential part of the therapy (Lindner et al., 2017; Öst, 1989). In this study, the test leaders lifted the tarantula from the glass cage to the table while conducting the BAT (between step 7 and 8) while the participants were watching. Several participants noted that it felt safer to touch the spider again and potentially hold it after seeing the test leader carrying it. While there was an expected steady increase in SUDS scores for both groups throughout the BAT, there is an exception to this trend with a dip in scores between step 7 and 8. This dip was 7.5 points for the VR group and 9.25 for the No VR group, a reduction that was larger than the average increase between most steps in the BAT for both groups. As this dip corresponded with the test leaders lifting the spider, we assume that some level of social model learning occurred because of this step and affected the participants' behavior. It is therefore worth considering when interpreting the results of the study, that social model learning might have affected the results of the BAT, potentially making them less valid.

While the VRET intervention was not shown to improve BAT scores, the VRET group had lower, if not significantly, SUDS scores. This indicates that they did not feel as much anxiety when confronted with the spider. It could therefore still be argued that it VRET might pose an advantage, whether stand-alone, or as a supplementary preparational intervention before in vivo exposure. Adding tactile cues and designing the VRET for further steps similar to the BAT could have given better learning and continued the lowered SUDS ratings throughout the entire BAT, transferring into a better score.

Strengths and limitations

Compared to most studies on use of VRET for arachnophobia this study had a relatively large sample. While this increases the chance that findings can be generalized to the broader population as well as increasing their reliability, it should be noted that not all the participants were considered phobic. It also did not meet the minimum standard of 30 recommended by Parsons and Rizzo (2008) for VRET studies, and the significant difference found was within the VR group of 23 participants, of which only 11 were categorized as phobic. The sample was also characterized by an equal gender distribution. Many earlier studies have tended to use mostly or exclusively female samples, as there likely are gender differences and most self-reported arachnophobes are female (Miloff et al., 2019; Peperkorn et al., 2016). As this study had an undiagnosed sample with no inclusion-criteria, the equal gender balance should be regarded as a strength. The inclusion of the BAT in the research procedure makes it easier to analyze the effect of VRET on real anxiety provoking situations and generalization to reality (Powers & Emmelkamp, 2008).

One of the participants had to stop their BAT testing prematurely as the real-life spider used in the testing was acting unpredictable. The test leader therefore decided to interrupt the testing and cancel the remaining appointments of the day. While the spider was able to continue testing afterwards, the erratic behavior persisted, which could have affected the validity of the BAT scores for participants being tested after this point.

Recruitment was done through a convenience and snowball sampling method. This can be problematic as it creates a selection bias increasing the risk of ending up with a sample that is neither representative nor strictly random in a statistical sense (Parker et al., 2019). Another limitation involves the subclinical sample instead of people with diagnosed spider phobia. Among participants with subclinical spider phobia, it is expected a lower treatment effect, due to the lower baseline at the starting point compared to clinical spider phobia (Lindner et al., 2020; Minns et al., 2019). In addition, this study did not include physiological measures like heart rate and skin conductance, nor measures such as perceived self-efficacy and dysfunctional beliefs. Côté and Bouchard (2009) found that changes in perceived self-efficacy and dysfunctional beliefs were the best predictors of change in general outcome and cardiac response, and should therefore be included to measure therapeutic improvements (Michaliszyn et al., 2010).

The study could have benefited from using the FSQ instead of SPQ for measuring fear of spiders. While SPQ is more commonly used in VRET research, the FSQ has been found to be superior in measuring fear in the nonphobic range (Muris & Merckelbach, 1996). And while this was not a treatment study, it can still be seen as a limitation that self-reported fear of spiders was not re-assessed at a later stage giving us no direct indication of treatment effect.

Implications

The opportunity to manipulate both the environment and the fear-inducing stimuli in VRET should be further studied, as a highly customized scenario could lead to better treatment outcomes (Trappey et al., 2021). One aspect of this customization could be the mobility of the subject in the VR-interventions, and further research should investigate static vs. mobile VRET scenarios in their ability to induce fear in a clinical sample (Lindner et al., 2017). Such studies should also investigate virtual representations of body parts, as well as the possibility of adding tactile cues (Hoffman et al., 2003; Peperkorn et al., 2016), but also needs to be aware of potential effect of cybersickness when making more technically advanced scenarios. Another possibility is creating graded intervention programs with fear hierarchies similar to the different scenarios applied in this study. This could be achieved by having different spiders of varying size, color, movement pattern etc., helping phobic participants move gradually from easier first meetings with virtual spiders than in vivo exposure could provide (Lindner et al., 2017).

The reduction in subjective fear reported when the test leader carried the spider implies that social model learning occurred in this study. While research on the topic within VRET is virtually non-existent, new VRET studies should study the potential of integrating social model learning in future VRET applications (Lindner et al., 2017). Such implementation would necessarily involve the use of a virtual person to demonstrate and should therefore be cautious of the possibility of the uncanny valley effect.

Conclusion

This study found the VRET intervention used was able to induce and reduce discomfort in the participants as intended. The participants reported levels of presence that indicate an immersive experience in the VRET program. The results did not show that VRET made it easier for the participants to approach a real-life spider. Yet the program still succeeded in raising and reducing discomfort, a tendency that was even stronger in the phobic group and is a prerequisite for exposure therapy to work. This can be seen as supporting, if not directly contributing, to a steadily expanding body of research documenting VRET for specific phobia as an effective intervention that could widen treatment accessibility as well as drastically reduce costs. In this study we also experienced many of the practical and logistical difficulties of using a live spider for testing, underlining the strong potential for use of such VR-based interventions in its place. The advent of reliable and user-friendly VR technology could lead to the development of many new VRET programs intended for treating anxiety disorders like arachnophobia. While this could ease the everyday life of both patients and therapists, it also puts stronger requirements on standardizing empirical research in this field to be able to make relevant comparisons. Lack of reported data was a recurring problem while trying to compare our results with other similar studies. VRET is a rapidly evolving technological method and therefore a strong methodology is especially important, as it needs constantly updated research to be documented as a viable treatment option.

References

Abdi, H. (2010). The greenhouse-geisser correction. *Encyclopedia of research design*, *1*(1), 544-548.

Adamovich, S. V., Fluet, G. G., Tunik, E., & Merians, A. S. (2009). Sensorimotor training in virtual reality: a review. *NeuroRehabilitation*, *25*(1), 29-44. https://doi.org/10.3233/NRE-2009-0497

Bouchard, S., Côté, S., St-Jacques, J., Robillard, G., & Renaud, P. (2006). Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games. *Technology and health care*, *14*(1), 19-27. https://doi.org/10.3233/THC-2006-14103

Bouchard, S., St-Jacques, J., Robillard, G., & Renaud, P. (2008). Anxiety increases the feeling of presence in virtual reality. *Presence: Teleoperators and Virtual Environments*, *17*(4), 376-391. https://doi.org/10.1162/pres.17.4.376

Bourne, E. J. (2011). *The anxiety and phobia workbook* (J. Blomquist, Ed.). New Harbinger Publications. https://books.google.no/books?hl=no&lr=&id=d3eYG3WKBs8C&oi=fnd&pg=PR9&dq=phobia&ots=GdfXPrMo1F&sig=Kmag762MGkWI9l0MF2GjPRvG1N8&redir_esc=y#v=onepage&q=phobia&f=false

Chen, C.-Y., Chuang, C.-H., Tsai, T.-L., Chen, H.-W., & Wu, P.-J. (2021). Reducing cybersickness by implementing texture blur in the virtual reality content. *Virtual Reality*, 1-12. https://doi.org/10.1007/s10055-021-00587-2

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (Rev. ed.). Academic Press. https://books.google.no/books?hl=no&lr=&id=rEe0BQAAQBAJ&oi=fnd&pg=PP1&dq=Cohen,+J.+(2013).+Statistical+power+analysis+for+the+behavioral+sciences.&ots=swWUItSTo5&sig=qdO2DvGNz4gxYo_tGabvZcm4-IY&redir_esc=y#v=onepage&q=Cohen%2C%20J.%20(2013).%20Statistical%20power%20analysis%20for%20the%20behavioral%20sciences.&f=false

Côté, S., & Bouchard, S. (2005). Documenting the efficacy of virtual reality exposure with psychophysiological and information processing measures. *Applied psychophysiology and biofeedback*, *30*(3), 217-232. https://doi.org/10.1007/s10484-005-6379-x

Côté, S., & Bouchard, S. (2009). Cognitive mechanisms underlying virtual reality exposure. *CyberPsychology & Behavior*, *12*(2), 121-129. https://doi.org/10.1089/cpb.2008.0008

Cowdrey, F. A., & Walz, L. (2015). Exposure therapy for fear of spiders in an adult with learning disabilities: a case report. *British Journal of Learning Disabilities*, *43*(1), 75-82. https://doi.org/10.1111/bld.12106

Cramer, D., & Howitt, D. L. (2004). *The Sage dictionary of statistics: a practical resource for students in the social sciences*. Sage. https://books.google.no/books?hl=no&lr=&id=pa3_49Mpso4C&oi=fnd&pg=PP1&dq=Cramer,+Duncan%3B+Howitt,+Dennis+Laurence+(2005).+The+SAGE+Dictionary+of+Statistics:+A+Practical+Resource+for+Students+in+the+Social+Sciences+(Third+ed.).+SAGE.+p.+21+(entry+%22ceiling+effect%22).+ISBN+978-0-7619-4138-5.+Lay+summary+(1+August+2010).&ots=YstjgOm0Cs&sig=AZ50wEQxKlaFqrIrSy7pHVTPMFU&redir_esc=y#v=onepage&q&f=false

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour research and Therapy*, *46*(1), 5-27. https://doi.org/10.1016/j.brat.2007.10.003

Craske, M. G., & Waters, A. M. (2005). Panic disorder, phobias, and generalized anxiety disorder. *Annu. Rev. Clin. Psychol.*, *1*, 197-225. https://doi.org/10.1146/annurev.clinpsy.1.102803.143857

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272-309. https://doi.org/10.1080/15213269.2015.1015740

de Borst, A. W., & de Gelder, B. (2015). Is it the real deal? Perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in psychology*, *6*, 576. https://doi.org/10.3389/fpsyg.2015.00576

DuPont, R. L., Rice, D. P., Miller, L. S., Shiraki, S. S., Rowland, C. R., & Harwood, H. J. (1996). Economic costs of anxiety disorders. *Anxiety*, *2*(4), 167-172. https://doi.org/10.1002/(SICI)1522-7154(1996)2:4%3C167::AID-ANXI2%3E3.0.CO;2-L

Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness III, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour research and Therapy*, *40*(9), 983-993. https://doi.org/10.1016/S0005-7967(01)00068-7

Garcia-Palacios, A., Hoffman, H. G., Kwong See, S., Tsai, A., & Botella, C. (2001). Redefining therapeutic success with VR exposure therapy. *Cyberpsychology and Behavior*, *4*, 431-348. https://doi.org/10.1089/109493101300210231

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4 ed.). Allyn & Bacon.

Ginsburg, A., & Smith, M. S. (2016). Do randomized controlled trials meet the "gold standard". *American Enterprise Institute*.

Gliem, J. A., & Gliem, R. R. (2003, October 8-10). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales* [Conference Paper]. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, Colombus. https://scholarworks.iupui.edu/bitstream/handle/1805/344/Gliem%20%26%20Gliem.pdf?sequence=1&isAllowed=y

Hesch, J., Kozminski, A., & Linde, O. (2019, 22. August). Powered by AI: Oculus Insight. https://ai.facebook.com/blog/powered-by-ai-oculus-insight/

Hoffman, H. G., Garcia-Palacios, A., Carlin, A., Furness Iii, T. A., & Botella-Arbona, C. (2003). Interfaces that heal: coupling real and virtual objects to treat spider phobia. *international Journal of Human-Computer interaction*, *16*(2), 283-300. https://doi.org/10.1207/S15327590IJHC1602_08

Holzwarth, V., Gisler, J., Hirt, C., & Kunz, A. (2021, March 20-22). *Comparing the Accuracy and Precision of SteamVR Tracking 2.0 and Oculus Quest 2 in a Room Scale Setup.* [Conference Paper]. ICVARS, Melbourne. https://dl.acm.org/doi/abs/10.1145/3463914.3463921

Jacobson, N. S., & Truax, P. (1992). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. https://doi.org/10.1037/0022-006X.59.1.12

Kim, J., Luu, W., & Palmisano, S. (2020). Multisensory integration and the experience of scene instability, presence and cybersickness in virtual environments. *Computers in Human Behavior*, *113*, 106484. https://doi.org/10.1016/j.chb.2020.106484

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical psychology review*, *24*(3), 259-281. https://doi.org/10.1016/j.cpr.2004.04.001

Lemelin, R. H., & Yen, A. (2015). Human-spider entanglements: Understanding and managing the good, the bad, and the venomous. *Anthrozoös*, *28*(2), 215-228. https://doi.org/10.1080/08927936.2015.11435398

Lindner, P., Miloff, A., Bergman, C., Andersson, G., Hamilton, W., & Carlbring, P. (2020). Gamified, automated virtual reality exposure therapy for fear of spiders: A single-subject trial under simulated real-world conditions. *Frontiers in psychiatry*, *11*. https://doi.org/10.3389/fpsyt.2020.00116

Lindner, P., Miloff, A., Hamilton, W., Reuterskiöld, L., Andersson, G., Powers, M. B., & Carlbring, P. (2017). Creating state of the art, next-generation Virtual Reality exposure therapies for anxiety disorders using consumer hardware platforms: design considerations and future directions. *Cognitive behaviour therapy*, *46*(5), 404-420. https://doi.org/10.1080/16506073.2017.1280843

Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., & 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*(5), e96144. https://doi.org/10.1371/journal.pone.0096144

Michaliszyn, D., Marchand, A., Bouchard, S., Martel, M.-O., & Poirier-Bisson, J. (2010). A randomized, controlled clinical trial of in virtuo and in vivo exposure for spider phobia. *Cyberpsychology, Behavior, and Social Networking*, *13*(6), 689-695. https://doi.org/10.1089/cyber.2009.0277

Miloff, A., Lindner, P., Dafgård, P., Deak, S., Garke, M., Hamilton, W., Heinsoo, J., Kristoffersson, G., Rafi, J., & Sindemark, K. (2019). Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. *Behaviour research and Therapy*, *118*, 130-140. https://doi.org/10.1016/j.brat.2019.04.004

Minns, S., Levihn-Coon, A., Carl, E., Smits, J. A., Miller, W., Howard, D., Papini, S., Quiroz, S., Lee-Furman, E., & Telch, M. (2019). Immersive 3D exposure-based treatment for spider fear: A randomized controlled trial. *Journal of anxiety disorders*, *61*, 37-44. https://doi.org/10.1016/j.janxdis.2018.12.003

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour research and Therapy*, *74*, 18-24. https://doi.org/10.1016/j.brat.2015.08.010

Muris, P., & Merckelbach, H. (1996). A comparison of two spider fear questionnaires. *Journal of behavior therapy and experimental psychiatry*, *27*(3), 241-244. https://doi.org/10.1016/S0005-7916(96)00022-5

Muris, P., Merckelbach, H., Holdrinet, H., & Sijsenaar, M. (1998). *Behavioral Avoidance Test (BAT)*. https://doi.org/10.1037/t11195-000

Olatunji, B. O., Woods, C. M., de Jong, P. J., Teachman, B. A., Sawchuk, C. N., & David, B. (2009). Development and initial validation of an abbreviated Spider Phobia Questionnaire using item response theory. *Behavior Therapy*, *40*(2), 114-130. https://doi.org/10.1016/j.beth.2008.04.002

Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and anxiety*, *29*(2), 85-93. https://doi.org/10.1002/da.20910

Parker, C., Scott, S., & Geddes, A. (2019). Snowball Sampling. *SAGE research methods foundations*. https://doi.org/10.4135/9781526421036831710

Parrish, D. E., Oxhandler, H. K., Duron, J. F., Swank, P., & Bordnick, P. (2016). Feasibility of virtual reality environments for adolescent social anxiety disorder. *Research on Social Work Practice*, *26*(7), 825-835. https://doi.org/10.1177/1049731514568897

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of behavior therapy and experimental psychiatry*, *39*(3), 250-261. https://doi.org/10.1016/j.jbtep.2007.07.007

Peperkorn, H. M., Diemer, J., & Mühlberger, A. (2015). Temporal dynamics in the relation between presence and fear in virtual reality. *Computers in Human Behavior*, *48*, 542-547. https://doi.org/10.1016/j.chb.2015.02.028

Peperkorn, H. M., Diemer, J. E., Alpers, G. W., & Mühlberger, A. (2016). Representation of patients' hand modulates fear reactions of patients with spider phobia in virtual reality. *Frontiers in psychology*, *7*, 1-11, Article 268. https://doi.org/10.3389/fpsyg.2016.00268

Powers, M. B., & Emmelkamp, P. M. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of anxiety disorders*, *22*(3), 561-569. https://doi.org/10.1016/j.janxdis.2007.04.006

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality*, *20*(2), 101-125. https://doi.org/10.1007/s10055-016-0285-9

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational research review*, *6*(2), 135-147. https://doi.org/10.1016/j.edurev.2010.12.001

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of behavior therapy and experimental psychiatry*, *38*(2), 105-120. https://doi.org/10.1016/j.jbtep.2006.10.001

Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I., & Valléry, G. (2020). User acceptance of virtual reality: an extended technology acceptance model. *International Journal of Human–Computer Interaction*, *36*(11), 993-1007. https://doi.org/10.1080/10447318.2019.1708612

Schwind, V., Leicht, K., Jäger, S., Wolf, K., & Henze, N. (2018). Is there an uncanny valley of virtual animals? A quantitative and qualitative investigation. *International Journal of Human-Computer Studies*, *111*, 49-61. https://doi.org/10.1145/3236673

Shiban, Y., Schelhorn, I., Pauli, P., & Mühlberger, A. (2015). Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: a randomized clinical trial in virtual reality. *Behaviour research and Therapy*, *71*, 45-53. https://doi.org/10.1016/j.brat.2015.05.014

Stinson, F. S., Dawson, D. A., Chou, S. P., Smith, S., Goldstein, R. B., Ruan, W. J., & Grant, B. F. (2007). The epidemiology of DSM-IV specific phobia in the USA: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological medicine*, *37*(7), 1047-1059. https://doi.org/10.1017/S0033291707000086

Trappey, A., Trappey, C. V., Chang, C.-M., Tsai, M.-C., Kuo, R. R., & Lin, A. P. (2021). Virtual reality exposure therapy for driving phobia disorder (2): System refinement and verification. *Applied Sciences*, *11*(1), 347. https://doi.org/10.3390/app11010347

UQO Cyberpsychology Lab. (2004). Presence Questionnaire (Rev. ed.). Retrieved 22. January 2021, from https://marketinginvolvement.files.wordpress.com/2013/12/pq-presence-questionnaire.pdf

Van Vlierberghe, D. (n.d.). *Grammastola Rosea*. Retrieved 10. August 2021 from https://www.theraphosidae.be/en/grammostola-rosea/

Wechsler, T. F., Kümpers, F., & Mühlberger, A. (2019). Inferiority or even superiority of virtual reality exposure therapy in phobias?—A systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard in vivo exposure in agoraphobia, specific phobia, and social phobia. *Frontiers in psychology*, *10*, 1758. https://doi.org/10.3389/fpsyg.2019.01758

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, *7*(3), 225-240. https://doi.org/10.1162/105474698565686

Wolpe, J. (1969). *The Practice of Behavior Therapy*. Pergamon Press.

Zsido, A. N. (2017). The spider and the snake–A psychometric study of two phobias and insights from the Hungarian validation. *Psychiatry research*, *257*, 61-66. https://doi.org/10.1016/j.psychres.2017.07.024

Öst, L.-G. (1989). One-session treatment for specific phobias. *Behaviour research and Therapy*, *27*(1), 1-7. https://doi.org/10.1016/0005-7967(89)90113-7

Öst, L.-G., Salkovskis, P. M., & Hellström, K. (1991). One-session therapist-directed exposure vs. self-exposure in the treatment of spider phobia. *Behavior Therapy*, *22*(3), 407-422. https://doi.org/10.1016/S0005-7894(05)80374-0

Åhs, F., Palmquist, Å. M., Pissiota, A., Appel, L., Frans, Ö., Liberzon, I., Furmark, T., & Fredrikson, M. (2011). Arousal modulation of memory and amygdala-parahippocampal connectivity: A PET-psychophysiology study in specific phobia. *Psychophysiology*, *48*(11), 1463-1469. https://doi.org/10.1111/j.1469-8986.2011.01231.x

NTNU

Norwegian University of
Science and Technology