Utrecht University

Master's Thesis

Title: "Public speaking anxiety, virtual reality exposure, and the role of expectancy violation and self-efficacy"

MSc. Levi van Ravesteyn

Student number: 553657

Thesis supervisor: Dr. Katharina Meyerbröker

Master program: Clinical Psychology 2021-2023

Wordcount: 4956

Date: March 17, 2023

Abstract

Aim: Virtual reality exposure therapy (VRET) is regarded as an effective treatment for public speaking anxiety (PSA). However, the effects are difficult to attribute to the Inhibitory Learning Model (ILM). In the ILM, the importance of expectancy violation during exposure is emphasized, meaning the occurrence of a mismatch between feared and actual outcomes. This study investigated potential working mechanisms of VRET by exploring the role of expectancy violation and self-efficacy. **Methods:** Participants reported their feared expectancies related to public speaking, and conducted 5 presentations in VR, split over two sessions. A total of 18 participants were randomly assigned to the experimental or control condition. Exclusively in the experimental condition participants were informed that the VR audience was unable to react to the presentations. **Results:** We found that PSA decreased based on one instrument but increased based on another instrument. Additionally, the magnitude of expectancy violation did not differ between treatment conditions. Interestingly, self-efficacy improvements were exclusively found in the experimental condition. **Conclusion:** The effects of VRET were not explicitly explained by the working mechanisms of expectancy violation and self-efficacy. However, there are indications that VRET in combination with a safe learning environment could be used to build self-efficacy.

Keywords: Virtual reality, exposure therapy, expectancy violation, self-efficacy

.

Introduction

Exposure therapy is an evidence-based treatment for various anxiety disorders such as social anxiety (Öst et al., 2015; Powers & Emmelkamp, 2008). Exposure involves clients confronting their fears by engaging in anxiety-provoking behavior (Scheveneels et al., 2021). Although exposure therapy is seen as effective, it can be challenging to expose someone with public speaking anxiety (PSA) to a large audience due to planning, cost, and resources (Scheveneels et al., 2019a). Virtual Reality Exposure Therapy (VRET) could be a viable solution as it can simulate a feared stimulus in a therapeutic setting without using a broad range of resources (Bun et al., 2017; Morina et al., 2014).

The Inhibitory Learning Model (ILM) is currently regarded as one of the main explanatory models of exposure therapy (Craske et al., 2022, 2014). In theory, successful exposure can lead to the formation of an inhibitory association, which competes with the prior excitatory fear association, responsible for fear responding (Craske et al., 2014). During exposure, the aim is to create the strongest mismatch between a client's expectancy for an unpleasant outcome and the actual outcome (Craske et al., 2014). The term expectancy violation (EV) is by Craske and colleagues (2014) used to describe this mismatch. Theoretically, the more profound EV is, the stronger the inhibitory learning processes are (Craske et al., 2014). Additionally, elements that diminish certain expectations prior to, or during exposure, potentially impede EV by lessening the mismatch between expectancies and actual outcome (Craske et al., 2014). Therefore, the use of cognitive restructuring or sharing safety information during exposure therapy about consequences that cannot happen may be problematic. As it could lower the mismatch between expectations and actual outcomes (Buchholz et al., 2022; Scheveneels et al., 2019b). For example, a client with PSA, who is instructed to give a presentation in VR, could be told that a digital audience is unable to react to the presentation. This type of safety information might be problematic, as it can alter feared expectations, impede expectancy mismatch processes and therefore negatively impact treatment results (Scheveneels et al., 2019a).

Based on the idea of expectancy violation (EV), certain forms of exposure in which EV is unlikely to occur should theoretically evoke less effective treatment results (Scheveneels et al., 2021). This could be the case for VRET since certain outcomes cannot occur in a VR environment (Meyerbröker, 2021). For example, an individual with aerophobia might fear the possibility of dying in a plane crash but is also likely aware that is outcome cannot occur during VRET (Scheveneels et al., 2021). As a result, this type of exposure

theoretically limits the potential for expectancy violation and therefore treatment efficacy (Craske et al., 2014). However, research on VRET for aerophobia has shown it to be as effective as cognitive therapy for aerophobia (Meyerbröker, 2014). Moreover, other studies on the efficacy of VRET for anxiety disorders such as social anxiety, agoraphobia, and panic disorder also show promising results (Meyerbröker & Morina, 2021). Interestingly, these findings are difficult to explain by the principle of EV, because participants' awareness that certain consequences cannot occur during VRET would theoretically lead to a lower mismatch between feared expectancies and actual outcomes (Craske et al., 2014). However, Scheveneels and colleagues (2021) explain why VRET for anxiety disorders can still be effective, even when limited possibilities for expectancy violation apply. Firstly, the authors state that the type of expectancy has an impact on the possibility for EV. Someone who fears getting a panic attack on a plane can in fact successfully test this expectation during VRET. The authors additionally state the possibility that VR could lead to a degree of immersion that makes participants forget about the virtual environment they are in, hence that certain outcomes cannot occur.

To our knowledge, one study has tried to assess whether the treatment effects of VRET for public speaking anxiety (PSA) can be attributed to the role of EV. In the study conducted by Scheveneels and colleagues (2019a), participants with PSA were asked to give presentations in VR. Participants also reported their feared expectancies related to public speaking and were asked whether these expectancies were testable in VRET. Scheveneels and colleagues (2019a) found treatment effects, however, the authors conclude that these cannot be attributed to the working mechanism of expectancy violation. Although, it is important to emphasize that EV was measured based on a proportion of testable expectancies related to own behavior and whether this predicted treatment outcome. This means that for the concept of EV, the actual occurrence of the feared outcome or change in harm expectations was not taken into consideration, which are key components related to the principle of expectancy violation (Pittig et al., 2022).

Interestingly, Scheveneels and colleagues (2019a) did not find a critical role for EV in VRET. Perhaps, other potential working mechanisms or factors play a role in effective VRET. Meyerbröker (2014), on the other hand, emphasizes the importance of self-efficacy, which refers to having faith in one's capacity to successfully execute certain behaviors (Bandura, 1977). Possibly, self-efficacy has an essential impact on the outcome of VRET as it is believed to decrease avoidance behavior and increase confidence in one's abilities (Meyerbröker, Morina, 2021). Moreover, shortcomings in self-efficacy promote the

utilization of dysfunctional coping strategies whenever individuals feel socially anxious (Kampmann et al., 2019). Two VRET studies have shown promising results regarding the impact of VRET on self-efficacy. Krijn et al (2007) conducted a study on patients with acrophobia and found that VRET led to a significant increase in self-efficacy. Additionally, Meyerbröker and Emmelkamp (2008) found that VRET was effective in improving self-efficacy among patients with specific phobias. The authors state the possibility that the efficacy of VRET could be explained by participants feeling more competent in dealing with challenging situations. Furthermore, Frisbee and colleagues (2020) state that VR can be used to create a realistic but safe practice environment that is ideal for building self-efficacy related to public speaking.

The present study aimed to broaden our knowledge on the working mechanisms of VRET for PSA, by exploring the role of expectancy violation and self-efficacy. This has been done by letting participants give 5 presentations in a VR environment and tracking their public speaking anxiety over time. We expected that participants taking part in VRET, would encounter significant improvements in self-reported PSA measures from baseline to follow-up (hypothesis 1). Secondly, we expected that, providing safety information exclusively in the experimental condition, would impede expectancy mismatch processes. Therefore, we predicted that participants belonging to the control condition were more likely to obtain better treatment outcomes (hypothesis 2). As an alternative, we also predicted that the change in harm expectancy from the 1st to the 4th exposure trial would be larger in the control condition (hypothesis 2.A). Finally, we expected that self-efficacy significantly increased from pre- to post-treatment regardless of treatment condition (hypothesis 3).

Methodology

Participants

This study targeted Dutch-speaking adults, aged 18-35 with elevated fear of public speaking. Participants were recruited through social media as Instagram, Facebook and study posters at Utrecht University. Participants were screened based on a two-item questionnaire that was used in earlier studies (Culver et al., 2012; Scheveneels et al., 2019a; Van Dis et al., 2021).

Exclusion criteria for participants included a history of motion sickness, problems with stereoscopic vision, epilepsy, having a DSM disorder (APA, 2013), receiving currently an anxiety treatment, recent changes in psychoactive medication, benzodiazepine use, or participation in another public speaking-related study. Furthermore, only participants with a total score of 17 or lower and a maximum score of 1 on suicidal ideation (item 9) on the Beck Depression Inventory-II were included (Beck, Steer & Brown 1996). The study was approved by the Utrecht University ethical commission of the faculty of social sciences (ID: FERB 22-0249).

Measures

Screening Visual Analogue Scales (VAS scales)

Participants were screened based on a two-item questionnaire that was used in previous studies (Culver et al., 2012; Scheveneels et al., 2019a; Van Dis et al., 2021). They were firstly asked how anxious they think they would feel when giving a formal speech in front of a live audience. Secondly, how likely it is that they would avoid taking a class that requires giving an oral presentation. Each question was rated on a 9-point scale, 0 indicating completely not, and 8 indicating extremely. Based on the work by Scheveneels et al. (2019a), participants scoring 6 or higher on anxiety and 5 or higher on avoidance were invited for further participation.

Beck Depression Inventory-II (BDI -II NL)

Depressive symptoms were measured with the Beck Depression Inventory-II (BDI-II; Beck, Steer & Brown, 1996). The questionnaire contains 21 themes with 3 answer statements about emotional well-being. Participants have to select a statement out of each row that resembles their feelings of the last two weeks. For example, theme 10 regarding crying has the answer options, 0 = ” I do not cry more than in the past”, 1 = “I cry more than in the past”, 2 = “I cry

about every little thing", 4 = "I would like to cry, but I can't". The BDI has shown good internal consistency, with a Cronbach's alpha of (.87) (Beck, Steer & Carbin, 1988).

Personal Report of Public speaking Anxiety (PRPSA)

Public speaking anxiety was measured with the Personal Report of Public Speaking Anxiety questionnaire (PRPSA; McCroskey, 1970). The questionnaire contains 34 statements about public speaking. Each statement is rated from strongly disagree (= 1) to strongly agree (= 5). For instance, "I get anxious when I think about a speech coming up." There is evidence for the internal consistency and validity of the PRPSA (Mörtberg et al., 2018). The instrument was administered during pre, post, and follow-up of VRET. In the current sample, the questionnaire demonstrated internal consistency, with a Cronbach's alpha of (.77).

Speech Anxiety Thoughts Inventory (SATI)

Thoughts associated with public speaking were measured by the Speech Anxiety Thoughts Inventory (SATI; Cho et al., 2004). The questionnaire contains 23 items, statements that need to be evaluated based on how much someone believes them. Answers range from not at all (= 1) to completely (= 5). For example, "my presentation will be incoherent". There is evidence for the internal consistency and validity of the SATI (Cho et al., 2004). The instrument was administered during pre, post, and follow-up of VRET. In the current sample, the questionnaire demonstrated internal consistency, with a Cronbach's alpha of (.85).

Evaluation of Expectancy Violation (EEV)

Before VRET, participants were asked to formulate concretely what their fear consisted of. Together with a researcher, this fear was formulated using an if-then statement, namely "if my mind goes blank then others will think I am weak". In addition, 3 questions regarding this expectancy were asked. These were: how likely it is that the specific expectancy will occur (harm expectancy); how bad it is if this expectancy will occur; what the current level of fear is regarding public speaking. The answers were ranked on a scale from 0-100. Harm expectancy ratings had to reach a minimum score of 60 to suffice, if not the feared expectancy was reformulated until a minimum harm expectancy score of 60 was obtained.

Exposure logs

Before every presentation, all participants were asked to answer two questions regarding their expectancy. Firstly, harm expectancy was assessed by the probability that participants believed that their expectancy would occur. Secondly, we asked how bad it would be if their expectancy would become reality. These questions were rated on a scale from 0-100, where 0 indicated not at all and 100 indicated extremely terrible. In addition, the duration of each presentation was registered in seconds.

Harm expectancy change was assessed by calculating the difference between the harm expectancy rating between the 1st and the 4th trial. The same method has been applied in a prior expectancy violation study (Buchholz et al., 2022).

Self-Efficacy (SE)

The 5-item self-efficacy questionnaire used by Krijn and colleagues (2007) was used and adapted to the current study. The self-efficacy questionnaire was designed to be used before VRET and directly after finishing the first session. The questionnaire measures the perceived confidence and ability to achieve specific goals related to managing fear and anxiety linked to public speaking. The questionnaire consists of five items, each of them connect to a specific goal: reducing fear, clear thinking, control over actions, control over anxious thoughts and images, and control over staying in a panic or fear-inducing situation. Answers can be provided on a scale from 0-100, with 0 indicating having no confidence at all, and 100 indicating feeling completely confident. For example, “how much confidence do you have in your ability to think clearly during a presentation?” In the current sample, the questionnaire demonstrated good internal consistency, with a Cronbach's alpha of (.78).

Manipulation

Before starting with the presentations, participants were randomized to receive either the experimental condition in which they were explicitly focused on not being able to disconfirm all expectations as the audience in VR is not real, or to the control condition with the same instructions except for the focus that a broader range of expectations can be violated. Participants received a document regarding the introduction of the intervention and were instructed to carefully read the rationale behind it.

Control condition

Participants in the control condition read; "*exposure therapy is a way to overcome anxiety by gradually confronting feared stimuli. Confronting feared situations is called exposure. For instance, when you fear public speaking the goal is to practice giving presentations in front of an audience instead of avoiding it. As a consequence, you will get more accustomed to the situation which gradually will lead to a decrease in public speaking anxiety".*

Experimental condition

Participants in the experimental condition read the same rationale in addition to the following information: "*It is important to remember that you will present in front of a digital audience those not being real people. In other words, the audience will consist of a 360-degree prerecorded audience, they are not able to react in real-time to your presentation. Therefore, it is a situation in which certain outcomes that you may fear in real life are not possible to occur".* Before every presentation, participants were also verbally reminded that they would present in front of a digital prerecorded audience, emphasizing that the audience could not react to their speech in real time.

Procedure

After completing eligibility screening, participants willing to take part in the experiment were invited to participate. During session 1, participants provided signed informed consent and completed the BDI-II, PRPSA, SATI, SE, and EEV. Participants were told that they would give 4 presentations in VR about various subjects, consisting of a job interview, a wedding toast, a presentation on climate change, and a presentation about the regulation of social media. They had to rank these subjects based on the degree in which these topics make them feel anxious (Scale: 0 = not anxious at all; 10 = extremely anxious). Afterwards, randomization took place.

During the experiment, participants gave four oral presentations in a row, each lasting 5 minutes with a 3-minute preparation window and they were allowed to repeat parts of the presentation whenever they felt "out of words". The presentation order was based on the ranked level of difficulty, with participants starting with the most difficult and ending with the least difficult topic. Participants were not allowed to use notes or other helpful material during the presentations. Each presentation topic was connected to a different matching virtual reality surrounding and audience. Before and after finishing each presentation, exposure logs were registered. After completion of the four presentations, participants were

instructed to complete the PRPSA, SATI, SE, and EEV again.

Following the first session, session two was conducted as an online video call one week later. Participants were instructed to complete the follow-up measure containing the PRPSA and SATI. Afterwards, participants were provided with five presentation subjects which they graded based on the degree of difficulty. These topics were: 1) education, 2) immigration and refugee crisis, 3) working remotely, 4) LGTBQ+, and 5) wokeness. Participants were instructed to present about the most difficult topic for a minimum of 1 minute but encouraged to continue for a maximum of two minutes. All participants got 3 minutes in order to prepare themselves. Furthermore, participants were told the presentation was recorded and later evaluated based on quality by a panel. In reality, no evaluation took place, this information was shared to raise anxiety levels before the presentation. While doing the presentation, the researcher shared his screen containing a prerecorded audience of three women. Upon finishing the presentation, the duration in seconds, subjective units of distress (SUDs), and potential safety behavior were registered.

Data Analysis

Power analysis

A statistical power analysis was used to estimate the minimum required number of participants to run a Repeated Measures Anova, containing a within-between interaction design, including four timepoints and two treatment groups. Using G*power it was estimated that for a small effect (Cohen's d) of 0.20, with an alpha = 0.05 and power = 0.80, the required sample size would be N= 36, meaning 18 participants in both treatment conditions were acquired to reach sufficient power.

Assumptions

Statistical analysis was carried out with the usage of SPSS (28) for Windows. Firstly, the assumptions of Anova related to normality and homogeneity were assessed and interpreted (Field, 2013). The Shapiro-Wilk, F max, and Levene's test statistic indicated that the assumptions were not violated.

Baseline equivalence of groups

Baseline equivalence of groups was assessed through an independent sample t-test. To understand whether the treatment and not a different factor led to differential outcomes in the treatment groups (Anderson & Maxwell, 2018). The treatment groups were at baseline compared based on age, educational level, and measurements on SATI, PRPSA, and SE.

Main effects

To investigate the main effects of the VRET we performed a 3(Time: pre-measure, post-measure, follow-up) x 2 (Group: A, B) Mixed Model Anova with SATI and PRPSA scores as dependent variables. Moreover, the effect of VRET on self-efficacy was assessed by a 2 (Time: pre-measure, post-measure,) x 2 (Group: A, B) Mixed Model Anova with self-efficacy scores as dependent variables

Harm expectation change

A 4(Time: trial 1-4) Repeated Measures Anova was executed to investigate changes in harm expectancies from the first to the fourth exposure trial. Furthermore, an independent sample t-test was executed to investigate mean differences in harm expectancy change between the two conditions.

Results

A total sample of 18 participants, 1 male, and 17 females; mean age = 22 years, SD = 3.8 participated in the current study. The participants were randomly allocated to experimental condition A (n = 9) and control condition B (n = 9). On average participants reported a highest obtained educational level of (M = 4.94, SD = 1.16), matching a level of higher vocational education.

Baseline equivalence of groups

An independent sample t-test was performed to compare the two treatment groups based on demographics and pre-measure scores on the SATI and PRPSA and SE. No significant differences were found between SATI-scores in the control condition (M = 73.00, SD = 6.80) and experimental condition (M = 74.66, SD = 12.17), t(12.55) = -.359, p = .726. In addition, no significant differences were found between PRPSA scores in the control condition (M = 137.67, SD = 6.20) and experimental condition (M = 136.22, SD = 13.87), t(16) = .29, p = .78. However, a significant difference was found between SE scores in the control condition (M = 278.89, SD = 42.71) and experimental condition (M = 187.33, SD = 65.79), t(16) = 3.50, p = .003. This result is an indication that participants in the control started the treatment with a higher level of self-efficacy. No significant differences were found between age in the control condition (M =23, SD =4.51) and experimental condition (M = 22, SD = 3.16), t(16) = .79, p = .44. Lastly, no significant differences were found in educational level between the control condition (M = 5.00, SD = 1.23) and experimental condition (M = 4.89, SD = 1.17), t(16) = 1.97, p = .85.

H1: Participation in VRET leads to significant improvements in self-reported PSA measures from baseline to follow-up.

Firstly, based on PRPSA-scores a significant main effect of time was obtained F(2,28) = 6.22, p = .006, partial $\eta 2$ = .31 with PRPSA-scores at pre-measure (M = 136.44, SD = 10.92), post-measure (M = 129.44, SD = 10.69) and follow-up (M =131.56, SD = 11.33).

Secondly, based on SATI-scores a significant main effect of time was obtained F(2,28) = 7.92, p = .002, partial $\eta 2$ = .36 with SATI-scores at pre-measure (M = 73.50, SD = 0.97), post-measure (M = 82.31, SD = 9.65) and follow-up (M =76.43, SD = 11.47). The results indicate a decrease in PSA based on PRPSA scores, but an increase in PSA based on SATI scores from pre-measure to follow-up. Results for each condition are presented in Table 1.

H2: Participants in the control condition are more likely to obtain better treatment outcomes.

Based on PRPSA-scores, no significant effect of time * condition was found $F(2,28) = 1.49$, $p = .24$. $\eta 2 = .10$. In accordance, based on SATI-scores, no significant effect of time * condition was found $F(2,28) = 2.05$, $p = 1.47$, $\eta 2 = .13$. Results are presented in Table 1.

H2.A: Harm expectancy change from the 1st to the 4th exposure trial would be larger in the control condition.

Firstly, the results suggest a significant effect of time, indicating a change in harm expectancy from the first to the 4^{th} trial, $F(3,42) = 4.12$, $p = .01$, partial $\eta 2 = .23$. The mean and standard error mean of harm expectancies are presented in Figure 1.

However, an independent sample t-test did not show a significant mean difference in harm expectancy change between the control and experimental condition, $t(14) = -.79$, $p = .44$.

H3: VRET leads to a significant increase in self-efficacy regardless of treatment condition.

No significant main effect of time on self-efficacy was obtained $F(1,14) = 4.67$, $p = .051$, $\eta 2 = .25$.

However, a significant effect of time * condition was found $F(1,14) = 7.61$, $p = .02$, $\eta 2 = .35$ with SE-scores at pre-measure in the control condition ($M = 278.89$, $SD = 42.72$) and experimental ($M = 170.14$, $SD = 62.99$) and SE-scores at post-measure in the control condition ($M = 271.89$, $SD = 58.23$) and experimental condition ($M = 225.14$, $SD = 99.25$). Indicating that participants in the experimental condition did encounter self-efficacy improvements that were absent in the control condition.

Table 1

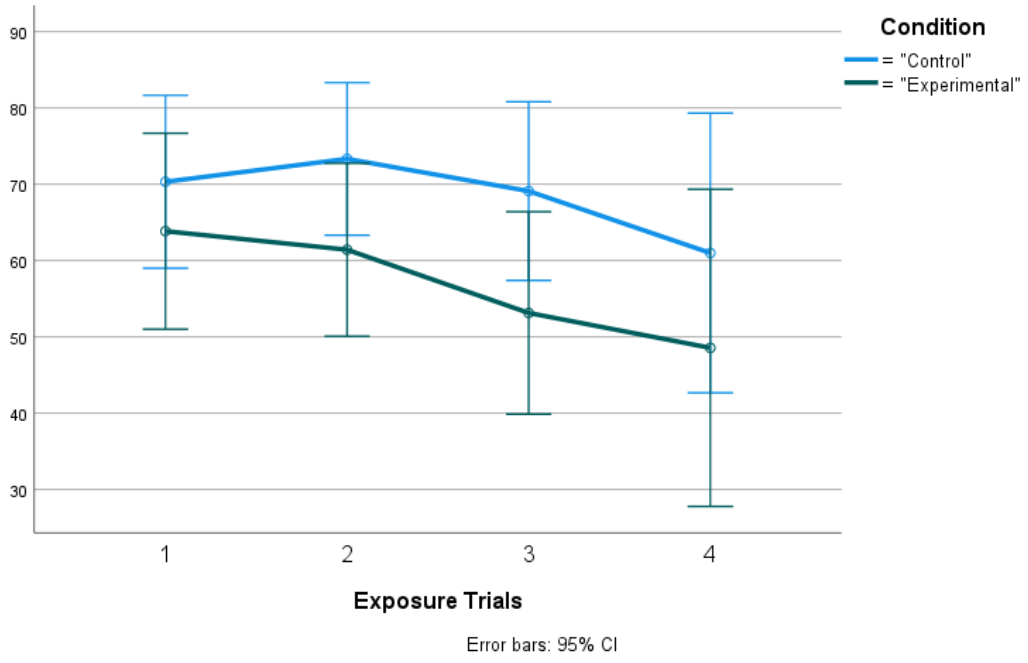
Descriptive data for outcome variables ordered by time and condition.

Tests	Group	Pre-test Mean (SD)	Post-test MEAN (SD)	Follow-up MEAN (SD)	ANOVA Time F	ANOVA Interaction F
PRPSA	Exp	134.86 (15.54)	125.29 (9.27)	131.43 (10.98)	6.22 *	1.49 (ns)
	Con	137.67 (6.20)	132.67 (11.10)	131.67 (12.25)		
SATI	Exp	74.14 (13.63)	78.86 (8.82)	77.29 (12.98)	7.92*	2.05 (ns)
	Con	73.00 (6.80)	85.00 (9.87)	75.78 (10.91)		
SE	Exp	170.14 (62.99)	225.14 (99.25)		4.57 (ns)	7.61 *
	Con	278.89 (42.72)	271.89 (58.23)			

Note. PRPSA = Personal Report of Public speaking Anxiety; SATI = Speech Anxiety Thought Inventory; SE = Self-Efficacy.

Figure 1.

Change in harm expectancies during 4 exposure trials of 1st treatment session.



Discussion

The current study investigated the working mechanisms of VRET on individuals with public speaking anxiety. Furthermore, we examined the potential role of expectancy violation and self-efficacy. We found that (1) public speaking anxiety declined based on PRPSA scores but increased based on SATI scores; (2) neither PRPSA nor SATI scores differed over time based on treatment condition; (3) harm expectancies decreased over the 4 trials for both conditions; (4) expectancy change did not differ between the two treatment conditions; and (5) self-efficacy only increased in the experimental condition.

Our first hypothesis that all participants would exhibit improvements in both public speaking anxiety outcome variables was not supported, because public speaking anxiety declined based on PRPSA scores, but increased based on SATI scores. This divergence could be attributed to the fact that the two scales have distinct focuses. The SATI predominantly focuses on maladaptive cognitions whereas the PRPSA also explores behavioral and physiological aspects of public speaking anxiety, with a greater emphasis on PSA in educational settings (Mörtberg et al., 2018). VRET employed in this study may have primarily targeted the behavioral and physiological aspects of public speaking anxiety. Alternatively, this outcome may suggest the possibility of iatrogenic effects, thus highlighting the importance for future studies to include multiple public speaking instruments that encompass physiological, behavioral, and cognitive aspects of PSA.

The second hypothesis, predicting superior treatment outcomes for participants in the control condition was not supported. No interaction effects between time and treatment condition were found for either the PRPSA or the SATI anxiety outcome variables. The initial notion was that the safety information provided only in the experimental condition would result in a reduction in the initial harm expectancy, thereby reducing the magnitude of expectancy violation. However, participants in both treatment conditions experienced a comparable degree of harm expectancy change, indicating no difference regarding expectancy violation processes as a result of experimental manipulation. Although, two aspects might have impacted the results. Firstly, various harm expectancies reported by participants, such as “If I do not know what to say, I will feel like a failure” or “If I do not know what to say, I will ruin the results of the study,” were unrelated to the digital audience. Therefore, these expectancies were perhaps not affected by the experimental manipulation. Future studies, could in their study designs make the distinction between expectancies related

to internal states and those related to the opinions/reactions of others more explicitly. Secondly, a weakness in the way how expectancy violation was measured in this study should be acknowledged. We were unable to utilize a more nuanced EV conceptualization as opposed to how it was measured in other studies since we did not inquire about the degree to which participants believed their feared expectations occurred after every trial (De Kleine et al., 2017; Pittig et al., 2022).

Our third hypothesis, that self-efficacy would increase regardless of treatment condition, was unsupported. Only participants in the experimental condition improved on the self-efficacy outcome measure. Possibly, these participants felt more reassured in their public speaking abilities as a result of the exposure rationale that was shared with them. Previous research has highlighted that VR can be used as an effective, low-risk, and comfortable practice environment for building self-efficacy related to public speaking (Frisby et al., 2020). Arguably, participants in the control group felt less comfortable practicing their presentations due to a fear of the VR audience's reaction, which could have resulted in lower scores on the post-measure self-efficacy instrument.

The findings should be interpreted while taking some limitations into account. Firstly, the small sample size led to insufficient power for identifying statistically significant differences. Secondly, our sample constituted predominantly of female students with a high educational level, which indicates a lack of diversity in gender and socioeconomic background. Finally, participants' ability to test their expectancies based on the VR audience, could have been impacted by the presence of the researcher in the room. Participants might have been predominantly occupied by how they were perceived by the researcher, instead of the VR audience. This may have been a confounding factor, impacting treatment results. Future studies could substantiate this argument by comparing the effects of having a researcher present versus having no researcher in the room during VR exposure trials.

Despite the limitations, this research has shown several strengths. Firstly, the current study is the first to explores the role of both EV and self-efficacy in VRET for PSA. Therefore, it provides critical directions on ways how EV can be measured and how PSA can be targeted by VRET. Moreover, EV was measured both by experimental manipulation and personal harm expectancy ratings. By having two distinct ways to assess EV we were able to test the principle of EV, as opposed by ILM, more rigorously.

In conclusion, no compelling evidence was found indicating that VRET was effective in decreasing public speaking anxiety. Consequently, we were unable to find a critical role for expectancy violation and self-efficacy. However, there are indications that VRET in

combination with a safe learning environment could be used to build self-efficacy (Frisby et al., 2020). Perhaps, VR could used as a tool in educational settings for individuals who struggle with giving presentations. Further exploration related to the working mechanisms of VRET is needed, as these findings are vital for the development of efficacious and affordable VRET.

Literature

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anderson, M. A., & Maxwell, N. (2018). Baseline Equivalence: What it is and Why it is Needed. *Mathematica 42*(4), 73-93.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory–II. *Psychological assessment.* https://doi.org/10.1037/t00742-000

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical psychology review, 8*(1), 77-100. https://doi.org/10.1016/0272-7358(88)90050-5

Buchholz, J. L., Blakey, S. M., Hellberg, S. N., Massing-Schaffer, M., Reuman, L., Ojalehto, H., ... & Abramowitz, J. S. (2022). Expectancy violation during exposure therapy: A pilot randomized controlled trial. *Journal of Behavioral and Cognitive Therapy, 32*(1), 13-24. https://doi.org/10.1016/j.jbct.2021.12.004

Bun, P., Gorski, F., Grajewski, D., Wichniarek, R., & Zawadzki, P. (2017). Low–cost devices used in virtual reality exposure therapy. *Procedia Computer Science, 104*, 445-451. https://doi.org/10.1016/j.procs.2017.01.158

Cho, Y., Smits, J. A., & Telch, M. J. (2004). The Speech Anxiety Thoughts Inventory: scale development and preliminary psychometric data. *Behavior Research and Therapy, 42*(1), 13-25. https://doi.org/10.1016/S0005-7967(03)00067-6

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour research and therapy, 58*, 10-23. https://doi.org/10.1016/j.brat.2014.04.006

Craske, M. G., Treanor, M., Zbozinek, T. D., & Vervliet, B. (2022). Optimizing exposure therapy with an inhibitory retrieval approach and the OptEx Nexus. *Behaviour Research and Therapy, 152*, 104069. https://doi.org/10.1016/j.brat.2022.104069

Culver, N. C., Stoyanova, M., & Craske, M. G. (2012). Emotional variability and sustained arousal during exposure. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(2), 787-793. https://doi.org/10.1016/j.jbtep.2011.10.009

De Kleine, R. A., Hendriks, L., Becker, E. S., Broekman, T. G., & van Minnen, A. (2017). Harm expectancy violation during exposure therapy for posttraumatic stress disorder. *Journal of Anxiety Disorders*, *49*, 48-52. https://doi.org/10.1016/j.janxdis.2017.03.008

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.

Frisby, B. N., Kaufmann, R., Vallade, J. I., Frey, T. K., & Martin, J. C. (2020). Using virtual reality for speech rehearsals: An innovative instructor approach to enhance student public speaking efficacy. *Basic Communication Course Annual*, *32*(1), 6.

Kampmann, I. L., Emmelkamp, P. M., & Morina, N. (2019). Cognitive predictors of treatment outcome for exposure therapy: do changes in self-efficacy, self-focused attention, and estimated social costs predict symptom improvement in social anxiety disorder?. *BMC psychiatry*, *19*(1), 1-8. https://doi.org/10.1186/s12888-019-2054-2

Krijn, M., Emmelkamp, P. M., Ólafsson, R. P., Bouwman, M., Van Gerwen, L. J., Spinhoven, P., ... & Van der Mast, C. A. (2007). Fear of flying treatment methods: virtual reality exposure vs. cognitive behavioral therapy. *Aviation, space, and environmental medicine*, *78*(2), 121-128.

McCroskey, J. C. (1970). Measures of communication-bound anxiety. *Speech Monographs, 37*, 269–277. https://doi.org/10.1080/03637757009375677

Meyerbröker, K. (2014). Virtual Reality Exposure Therapy. In P. M. G. Emmelkamp & T. Ehring (Eds.), *The Wiley handbook of anxiety disorders* (pp. 1310–1324*)*. Wiley & Blackwell. https://doi.org/10.1002/9781118775349.ch63

Meyerbröker, K. (2021). Virtual reality in clinical practice. *Clinical Psychology & Psychotherapy*, *28*(3), 463. https://doi.org/10.1002/cpp.2616

Meyerbröker, K., & Emmelkamp, P. M. (2008). Therapeutic processes in virtual reality exposure therapy: The role of cognitions and the therapeutic alliance. *Journal of Cybertherapy & Rehabilitation*, *1*(3), 247-257.

Meyerbröker, K., & Morina, N. (2021). The use of virtual reality in assessment and treatment of anxiety and related disorders. *Clinical Psychology and Psychotherapy, 28*(3), 466–476. https://doi.org/10.1002/cpp.2623

Morina, N., Brinkman, W. P., Hartanto, D., & Emmelkamp, P. M. (2014). Sense of presence and anxiety during virtual social interactions between a human and virtual humans. *PeerJ*, *2*, e337. https://doi.org/10.7717/peerj.337

Mörtberg, E., Jansson-Fröjmark, M., Pettersson, A., & Hennlid-Oredsson, T. (2018). Psychometric properties of the personal report of public speaking anxiety (PRPSA) in a sample of university students in Sweden. *International journal of cognitive therapy, 11*(4), 421-433. https://doi.org/10.1007/s41811-018-0022-0

Öst, L. G., Havnen, A., Hansen, B., & Kvale, G. (2015). Cognitive behavioral treatments of obsessive–compulsive disorder. A systematic review and meta-analysis of studies published 1993–2014. *Clinical psychology review, 40,* 156-169. https://doi.org/10.1016/j.cpr.2015.06.003

Pittig, A., Heinig, I., Goerigk, S., Richter, J., Hollandt, M., Lueken, U., Pauli, P., Deckert, J., Kircher, T., Straube, B., Neudeck, P., Koelkebeck, K., Dannlowski, U., Arolt, V., Fydrich, T., Fehm, L., Ströhle, A., Totzeck, C., Margraf, J., Wittchen, H.U. (2023). Change of Threat Expectancy as Mechanism of Exposure-Based Psychotherapy for Anxiety Disorders: Evidence From 8,484 Exposure Exercises of 605 Patients. *Clinical Psychological Science, 11*(2), 199–217. https://doi.org/10.1177/21677026221101379

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

Scheveneels, S., Boddez, Y., & Hermans, D. (2021). Uitdagingen voor het inhibitorisch-leermodel van exposuretherapie. *GEDRAGSTHERAPIE (UTRECHT), 54*(1), 3-25.

Scheveneels, S., Boddez, Y., De Ceulaer, T., & Hermans, D. (2019b). Ruining the surprise: The effect of safety information before extinction on return of fear. *Journal of behavior therapy and experimental psychiatry*, *63*, 73-78. https://doi.org/10.1016/j.jbtep.2018.11.001

Scheveneels, S., Boddez, Y., Van Daele, T., & Hermans, D. (2019a). Virtually unexpected: No role for expectancy violation in virtual reality exposure for public speaking anxiety. *Frontiers in psychology*,*10*, 2849. https://doi.org/10.3389/fpsyg.2019.02849

van Dis, E. A., Landkroon, E., Hagenaars, M. A., van der Does, F. H., & Engelhard, I. M. (2021). Old fears die hard: Return of public speaking fear in a virtual reality procedure. *Behavior Therapy*, *52*(5), 1188-1197. https://doi.org/10.1016/j.beth.2021.01.005