Available online at www.sciencedirect.com
ScienceDirect
Behavior Therapy 52 (2021) 1188–1197
Behavior Therapy
www.elsevier.com/locate/bt

Methodological paper

Old Fears Die Hard: Return of Public Speaking Fear in a Virtual Reality Procedure

Eva A.M. van Dis
Elze Landkroon
Muriel A. Hagenaars
Florentine H.S. van der Does
Iris M. Engelhard
Utrecht University

Exposure-based therapy is an effective treatment for social anxiety, but some patients relapse. We used a novel virtual reality procedure to examine spontaneous recovery (i.e., a return of fear over time) and fear renewal (i.e., the return of fear after a context switch) in individuals with fear of public speaking. On Day 1, 32 participants received exposure training before a virtual audience. On Day 8, participants completed a spontaneous recovery phase, followed by a fear renewal test, in which they gave a presentation in front of a new (context switch) or the same audience (no context switch). After exposure, participants exhibited a lower heart rate, subjective distress, negative valence, and arousal. One week later, participants showed spontaneous recovery of heart rate, and the context switch group showed renewal of subjective distress, negative valence, and arousal. Future studies can use this procedure to test interventions aimed at improving long-term exposure effects in individuals with public speaking fear.

Florentine H. S. van der Does is now at Leiden University Medical Center. This study is registered on the Open Science Framework (https://osf.io/udny4/). Declaration of interest: None. This study was supported by a Vici grant (453-15-005) from the Netherlands Organization for Scientific Research (NWO) awarded to Iris M. Engelhard. We thank Nique van Outvorst, Wouter van Haperen, and Evy van Weelden for recruiting and testing participants.

Address correspondence to Eva van Dis, Department of Clinical Psychology, Utrecht University, PO Box 80140, 3508 TC Utrecht, The Netherlands. e-mail: e.a.m.vandis@uu.nl.

Keywords: public speaking anxiety; virtual reality exposure; return of fear; fear renewal

About 1 in 10 individuals will meet diagnostic criteria for social anxiety disorder at some point in their lives (Kessler et al., 2012). Social anxiety disorder is characterized by exaggerated fear of social rejections and avoidance of social situations (American Psychiatric Association, 2013). Although many patients with social anxiety disorder benefit from exposure-based therapy (Loerinc et al., 2015; Van Dis et al., 2020), about 13% of recovered patients experience relapse (Fava et al., 2001). One commonly accepted explanation for relapse is that fear learning is context-dependent (e.g., Vervliet, Craske, et al., 2013). Specifically, research has shown that the original fear learning (e.g., the belief that other people will judge oneself negatively) is not erased during exposure therapy and can easily resurface in new contexts (i.e., fear renewal; Bouton, 2002). Therefore, research paradigms on fear renewal may be useful for acquiring knowledge to eventually increase long-term treatment success for social anxiety disorder.

Fear renewal has been extensively studied in fear conditioning paradigms (Vervliet, Baeyens, et al., 2013). These usually start with a fear-learning phase in which a danger cue is repeatedly paired with an aversive outcome (e.g., a mild electric shock). Then, in an extinction-training phase, this cue is no longer paired with the aversive outcome, which typically extinguishes fear responses toward the cue. Finally, the danger cue is presented after a time lapse (i.e., spontaneous recovery test) or in a different context (i.e., fear-renewal test), generally resulting in a return of fear. Even though fear conditioning studies have substantially contributed to our understanding of learned fear (Vervliet, Craske, et al., 2013), their ecological validity has been criticized. First, most fear conditioning paradigms use simple, generally nonmeaningful stimuli (e.g., geometrical shapes) rather than personally meaningful and complex multimodal stimuli that are involved in clinical disorders (Landkroon et al., 2019; Scheveneels et al., 2018). Second, they often rely on passive learning (Scheveneels et al., 2016), while in real-life, individuals actively approach or avoid feared stimuli or situations. Third, fear conditioning paradigms typically instill new fear memories, whereas existing old fear memories are harder to modulate (Eisenberg et al., 2003). Thus, more ecologically valid fear renewal paradigms with complex stimuli, active behavior, and preexisting fear memories are needed.

To our knowledge, only one study has successfully demonstrated fear renewal in individuals with social anxiety using a more ecologically valid procedure (Culver et al., 2011; Study 1). In that study, participants with a fear of public speaking (a type of social anxiety; American Psychiatric Association, 2013) first received exposure training in front of a live audience in Context A. One week later, they received brief exposure again in the same room (Context A) or a new room (Context B). Participants in Context B showed increased subjective fear and heart rate (HR) during the exposure task 1 week later compared to their initial exposure, while those in Context A did not show an increase in fear or HR from the first exposure to the second (Culver et al., 2011, Study 1). Follow-up studies with this procedure have shown that the use of retrieval cues during exposure slightly reduces fear renewal (Culver et al., 2011, Studies 2 and 3) if participants do not perceive them as safety cues (Shin & Newman, 2018). This illustrates the procedure's potential to reveal mechanisms that may reduce fear renewal. Yet, there is room for improvement—that is, a fear-renewal procedure conducted in virtual reality (VR) would allow for more standardization of the audience (Parsons, 2015) and allow more researchers to use the procedure. In addition, the procedure of Culver et al. did not control for spontaneous recovery, even though it may have overshadowed fear renewal in previous studies (Craske et al., 2019; Shin & Newman, 2018).

We aimed to address these issues by developing and validating a 2-day novel VR procedure (following Culver et al., 2011) and testing whether it can be used to examine a return of public speaking fear 1 week after exposure. In our study, participants with a fear of public speaking first received exposure training in VR. After 1 week, they were all tested for spontaneous recovery of fear and they received additional exposure training in the same virtual environment as the previous week. The experimental manipulation was that at the end of the additional exposure training, the virtual context switched for one of the two groups. We expected spontaneous recovery of fear for all participants, but expected fear renewal only for the group that received a context switch relative to the group with no context switch. We also explored the return of subjective negative valence and arousal to delineate the specific emotional responses of this setup.

Method

PARTICIPANTS

Native Dutch-speaking students were recruited via Utrecht University, Facebook, and Proefbunny.nl to fill out two questions assessing how anxious they thought they would feel when giving a formal speech in front of a live audience and how likely they would avoid taking a class that requires giving an oral presentation, each rated on an 9-point scale (0 = *none/never* and 8 = *extremely/always*; see Culver et al., 2011). If they scored ≥6 on both questions, they were further screened and excluded if they reported heart, respiratory, or neurological problems or 3-D motion sickness. Thirty-seven eligible participants were invited to the first lab session and completed an informed consent procedure. They then completed the Beck Depression Inventory–II (BDI-II; Beck et al., 1996). If they reported 1 or higher on the item measuring suicidal ideation or had a total score of 18 or higher, they were then excluded from the study (n = 3) to prevent a potential worsening of symptoms. This resulted in 34 participants. Data from one participant were excluded from the final analyses because of a technical issue with the VR equipment, and data from one participant were excluded because of noncompliance. The final sample size comprised 32 participants (10 males, 22 females; mean age = 22.41 years, SD = 3.29) who were allocated to Context AA (n = 16) or AB (n = 16) groups (in random order; stratified for gender). The ethics committee of the Faculty of Social and Behavioral Sciences at Utrecht University (FETC17-073) approved this study. We pre-registered the study (including a power analysis) on the Open Science Framework (https://osf.io/udny4/).

POWER ANALYSIS

A power analysis (using G*Power 3.1.9.2) for a mixed-factorial analysis of variance (ANOVA) with two groups and two measures ($f = 0.25$, $\alpha = .05$, power = 0.80) yielded a total sample size of 34 participants (i.e., 17 per group). Although we pre-registered a power analysis using a power level of 0.95 (yielding a sample size of 54 participants), a power of 0.80 is often considered preferable (Cohen, 1992). We used an optimal stopping procedure, which allowed us to stop our data collection whenever we found strong evidence in favor of the null or alternative hypothesis (i.e., Bayes factor >10). After testing 32 eligible participants (i.e., the sample size of Culver et al., 2011), we obtained a strong effect on fear renewal for subjective units of distress (SUDS; $BF_{10} > 10.0$) and therefore stopped our data collection. Although stopping rules are considered problematic for frequentist statistics, they are appropriate and commonly used in Bayesian statistics (e.g., Rouder, 2014).

MEASURES

Questionnaires and Interview

Structured Clinical Interview for DSM-5 Disorders (SCID-5-CV). Social anxiety disorder was assessed using Questions F32–F41 of the SCID-5-CV (First et al., 2016) by trained clinical psychology students. The sections were translated from English to Dutch and back-translated by independent researchers. Independent raters (EvD and EL) evaluated the presence of a diagnosis (interrater reliability $\kappa = 0.79$).

Personal Report of Public Speaking Anxiety (PRPSA). Fear of public speaking was assessed with the PRPSA (McCroskey, 1970), which also has been validated in a Dutch sample (Cronbach's $\alpha = .83$; van Veen et al., 2020). It has good convergent validity ($r = .41$ with a communication apprehension scale) and high internal consistency ($\alpha = .94$; McCroskey, 1970). This 34-item scale consists of negative and positive statements that are rated on a 5-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Positive items are reverse scored. Cronbach's α was .90 in this study.

Brief Fear of Negative Evaluation Scale–II (BFNE-II). Fear of negative evaluations was measured with a validated Dutch version of the BFNE-II (Carleton et al., 2007; Cieraad & de Jong, 2007). Carleton et al. reported good construct validity (convergent validity with social phobia scales: $rs = .60–.64$; discriminant validity with illness and injury scales: $rs = .29–.38$) and excellent internal consistency ($\alpha = .97$). This scale consists of 12 statements that are rated on a 5-point Likert scale, ranging from 0 (*not at all characteristic of me*) to 4 (*extremely characteristic of me*). Cronbach's α was .95 in this study.

Behaviors Checklist (BCL). Perceived speech performance of one's speech was assessed with the 18-item BCL (derived from Mansell & Clark, 1999; Stopa & Clark, 1993; Vasey et al., 2012). The items were translated from English to Dutch and back-translated by independent researchers. This 18-item scale consists of negative and positive speech characteristics that are rated on a 9-point Likert scale, ranging from 0 (*not at all*) to 8 (*extremely*). Positive items are reverse scored. Cronbach's α was .80–.85 in this study.

Neuroticism Scale of the Eysenck Personality Questionnaire (EPQ-N)[2]. Neuroticism was assessed with a validated Dutch version of the EPQ-N (Eysenck & Eysenck, 1991; Sanderman et al., 2012). Its convergent validity is demonstrated by a strong correlation with another neuroticism scale ($r = .78$), and discriminant validity is indicated by a negative correlation with an emotional stability subscale ($r = -.70$). It also has excellent internal consistency ($\alpha = .87$; Barelds & Luteijn, 2002). This scale has 22 questions that are rated on a dichotomous scale (0 = *no*, 1 = *yes*). Cronbach's α was .90 in this study.

Anxiety Sensitivity Index (ASI). Anxiety sensitivity was measured with a validated Dutch version of the ASI (Reiss et al., 1986; Vujanovic et al., 2007). The convergent and discriminant validity is high: the ASI total score was positively associated with scales measuring anxious arousal ($r = .42$) and negative affectivity ($r = .35$). It did not show significant correlations with anhedonic depression ($r = .07$) and positive affectivity ($r =$

[2] Measures of anxiety-relevant personality traits, such as neuroticism, anxiety sensitivity, emotional reasoning (Arntz et al., 1995; Engelhard et al., 2001), and personalized implicit associations (Vasey et al., 2012) were included in this study to explore whether these predicted return of fear. Yet, our power analysis and stopping rule were aimed at obtaining sucient statistical power for the renewal analyses. Because the renewal eects were already large after testing 32 participants, we decided to stop testing, and we do not report the return of fear predictors due to limited statistical power.

.02). In addition, the internal consistency is good (α = .83; Vujanovic et al., 2007). The ASI consists of 16 statements (e.g., "It scares me when my heart beats rapidly") that are rated on a 5-point Likert scale, ranging from 0 (*very little*) to 4 (*very much*). Cronbach's α was .79 in this study.

VR Experiences Scale. A self-constructed seven-item questionnaire assessed three physiological complaints (nausea, headache, and dizziness), realness, immersion, presence, and whether presenting in VR was equally challenging as presenting in real life. Items were rated on a 5-point Likert scale, ranging from 1 (*barely*) to 5 (*very much*).

Subjective Ratings

Subjective Units of Distress Scale (SUDS). Subjective distress was assessed with the SUDS, a 100-point scale with five anchors: 0 (*no distress*), 25 (*mild distress*), 50 (*moderate distress*), 75 (*severe distress*), and 100 (*very severe distress*).

VR Valence and Arousal. VR valence and arousal were assessed by the following two questions: "How positive or negative do you find this audience?" and "How aroused do you feel when seeing this audience?" on an 11-point scale ranging from 0 (*negative/not arousing*) to 10 (*positive/arousing*). VR valence ratings were reverse scored for ease of interpretation.

Speech Topic Difficulty. Speech topic difficulty was measured with the following question: "How difficult do you find it to give a speech on this topic?" that was rated on a 10-point scale, ranging from 1 (*very easy*) to 10 (*very difficult*).

Heart Rate

HR was measured with a Polar H10 chest strap electrocardiogram (ECG) monitor that was connected to the free iOS HRV+ app (ZUZ LLC) on an iPad. Polar wearable HR monitors are reliable (Georgiou et al., 2018) and have often been used in similar research (e.g., Culver et al., 2011). For each day, the average beats per minute (BPM) during the baseline measurement was subtracted from the average BPM during each speech (Culver et al., 2011). We used the final minute of the 5-min baseline period to ensure that participants' HR returned to their baseline level. Following Vasey et al. (2012), we additionally analyzed the average BPM during each speech, without correcting for the baseline measurements (see Supplemental Materials). Data were monitored at 130 Hz and analyzed with Kubios HRV Standard (version 3.2).

PROCEDURE

On Day 1, participants provided informed consent and completed the BDI-II, PRPSA, BFNE-II, EPQ-N, and the ASI, followed by the SCID-5. Next, they put on the Polar chest strap and were instructed to "Please remain seated quietly, without speaking to the experimenter" for a 5-min HR baseline measurement. They then put on the VR headset and practiced for 2 minutes with SUDS, VR valence, and arousal ratings in a neutral VR environment. Hereafter, participants faced the virtual audience for 10 sec (Context A) and completed the baseline VR valence and arousal ratings. This was followed by an exposure phase in which they gave four 5-min speeches. Before each speech, they received three unique topics that were somewhat controversial (e.g., euthanasia, death penalty, and immigration; see Table S1 in Online Supplemental Material). From these, they chose one topic and rated its difficulty. They had 1 minute of preparation time in which they were not allowed to make notes. They then gave a speech in Context A during which they indicated the SUDS rating out loud at the start of the speech and 1-min intervals. If they stopped the presentation within 5 minutes, the experimenter instructed them to continue presenting even if that meant they had to repeat themselves. After each speech, participants took off the VR headset for a 1-min rest. After finishing the last speech, they were asked to complete VR valence and arousal ratings again. Finally, they filled out the BCL regarding the last speech. They also completed an emotional reasoning task and personalized implicit association test,[1] and these data are not reported further.

On Day 8, participants put on the Polar chest strap followed by another 5-min HR baseline measurement. Hereafter, participants gave four 5-min speeches, following the same procedure as Day 1. All participants gave the first three speeches in Context A (i.e., spontaneous recovery test). The fourth speech was either in Context A or Context B (i.e., fear renewal test). Afterward, they completed the VR valence and arousal ratings, the BCL, and the VR questionnaire. When the experiment was finished, participants were asked final questions for exploratory research purposes and were debriefed.

VR ENVIRONMENTS

The speech environments were two 360-degree videos depicting an audience (freely derived from https://virtualspeech.com; see Figure S1 in the Online Supplemental Material). The virtual audience consisted of either 11 or about 75 individuals with mixed gender and ethnic background. Their facial expressions were neutral to positive, and they had different levels of attentiveness. The environments were fully balanced across participants. The neutral environment was a 360-degree picture of a room with a couch, a desk, and a computer (purchased from TurboSquid, see https://www.turbosquid.com). The VR environments were presented with an Oculus Rift headset (version CV1; Oculus, USA) and the Oculus Rift App (version 1.19.0.456194).

DATA ANALYSES

First, to examine whether randomization was successful, we performed one-way ANOVAs to assess the effects of group (Context AA, AB) on age, public speaking fear (PRPSA, BFNE-II), speech performance (BCL Day 1), EPQ-N scores, ASI scores, average speech topic difficulty, HR baseline on Day 1, pre-exposure VR valence and arousal ratings, and the VR questionnaire. A Bayesian Contingency Tables Test assessed group differences in social anxiety disorder diagnosis. Second, to assess whether exposure training was successful (manipulation check), we performed two 4 (Time: four speeches Day 1) × 2 (Group: AA, AB) mixed ANOVAs with HR and SUDS ratings as dependent variables. For each speech, the average HR and the highest SUDS rating were selected for statistical analyses (see Shin & Newman, 2018). We additionally performed two 2 (Time: pre-exposure, post-exposure Day 1) × 2 (Group: AA, AB) mixed ANOVAs with VR valence and arousal ratings as dependent variables. To test whether HR baselines differed across time, we used a repeated-measures ANOVA with time (Day 1, Day 8) as an independent variable. Finally, we tested whether spontaneous recovery occurred for both groups and whether fear renewal occurred for the AB group, with 2 (Time: final speech Day 1, first speech Day 8 [spontaneous recovery]; and third speech Day 8, final speech Day 8 [fear renewal]) × 2 (Group: AA, AB) mixed ANOVAs performed separately for HR and SUDS. We also tested these hypotheses with VR valence and arousal ratings as dependent variables: 2 (Time: post-exposure Day 1, pre-exposure Day 8 [spontaneous recovery]; and pre-exposure, post-exposure Day 8 [fear renewal]) × 2 (Group: AA, AB).

All analyses were performed in JASP version 0.12.2.0 within the Bayesian hypothesis testing framework using the default settings (JASP Team, 2020). Bayes factors quantify the likelihood of the data under one hypothesis relative to another. For example, BF_{10} = 3.0 would mean that the data are three times more likely under the alternative than the null hypothesis (and vice versa for BF_{10} < 0.33; Aczel et al., 2020). We interpreted a BF_{10} between 1.0 and 3.0 as anecdotal evidence, values between 3.0 and 10.0 as moderate evidence, and values greater than 10.0 as strong evidence in favor of the alternative hypothesis. A BF_{10} below 0.33 indicates evidence in favor of the null hypothesis (Jeffreys, 1961; Wagenmakers et al., 2018). These classifications should only be used as a general rule of thumb and not as an absolute rule (Wagenmakers et al., 2018).

HR data were missing at random for Day 1 (n = 5) and Day 8 (n = 4). Missing values were imputed in R version 3.6.1. We generated five imputed data sets with predictive mean matching (five iterations) using the mice package version 3.0 (van Buuren & Groothuis-Oudshoorn, 2011) in R. Analyses of these data sets did not differ from complete case analyses (see also Figure S2 in Online Supplemental Material).

Results

RANDOMIZATION AND MANIPULATION CHECKS

We found no evidence that groups differed in age (BF_{10} = 0.38), public speaking fear (BFs_{10} = 0.34–0.37), speech performance on Day 1 (BF_{10} = 0.35), EPQ-N scores (BF_{10} = 0.34), ASI scores (BF_{10} = 0.45), average speech topic difficulty (BF_{10} = 0.62), HR baseline on Day 1 (BF_{10} = 0.45), pre-exposure VR valence (BF_{10} = 0.38) and arousal (BF_{10} = 0.59), VR questionnaire items (BFs_{10} = 0.34–1.11), and social anxiety disorder diagnosis (BF_{10} = 0.40), suggesting successful randomization (see Table 1). On Day 1, HR and subjective distress decreased during the four speeches (Time: BFs_{10} > 8.81); see Figure 1. There was no evidence for main or interaction effects of group (BFs_{10} < 1.30). From pre- to post-exposure, participants rated the VR environment as less negative (BF_{10} = 3.26) and less arousing (BF_{10} = 4.12), with no evidence for group differences (BFs_{10} < 0.49); see Figure 2. Thus, for both groups, HR, subjective distress, negative valence, and arousal ratings declined after exposure.

SPONTANEOUS RECOVERY

HR increased from the final speech on Day 1 to the first speech on Day 8 (BF_{10} = 30,061.15), with a stronger increase for the AA group (Time × Group: BF_{10} = 2.56), but no evidence for a main effect of group (BF_{10} = 0.55); see Figure 1a. Baseline HR did not differ between Day 1 and Day 8 (BF_{10} = 0.52), with no evidence for a main or interaction effect for group (BFs_{10} < 0.64). Subjective distress did not change over time (Time: BF_{10} =

Table 1
Overview of Randomization Variables

Questionnaire	Mean (*SD*)	Group	
		Context AA	Context AB
Diagnosis; no. (%)	12 (37.50)	6 (37.50)	6 (37.50)
PRPSA	131.63 (16.03)	133.00 (16.98)	130.25 (15.44)
BFNE-II	36.56 (12.76)	37.06 (13.04)	36.06 (12.88)
BCL (Day 1)	83.25 (16.82)	82.31 (17.07)	84.19 (17.07)
EPQ-N	10.75 (5.88)	10.69 (6.34)	10.81 (5.58)
ASI	32.72 (8.22)	31.44 (4.93)	34.00 (10.58)
HR baseline Day 1	84.03 (16.22)	81.73 (18.95)	86.69 (12.58)
Speech topic difficulty	6.44 (1.47)	6.11 (1.58)	6.77 (1.32)
VR negative valence	5.38 (2.21)	5.19 (2.48)	5.56 (1.97)
VR arousal	6.19 (2.04)	6.63 (1.89)	5.75 (2.15)
VR questionnaire			
Physiological complaints	1.18 (0.29)	1.19 (0.32)	1.18 (0.26)
Realness	2.47 (0.86)	2.47 (0.74)	2.47 (0.99)
Immersion	2.80 (1.06)	2.47 (0.99)	3.13 (1.06)
Presence	2.80 (0.89)	2.67 (0.82)	2.93 (0.96)
VR equally challenging as in real-life	2.03 (0.89)	2.07 (0.80)	2.00 (1.00)

Note. SD = standard deviation; PRPSA = Personal Report of Public Speaking Anxiety; BFNE = Brief Fear of Negative Evaluation Scale–II; BCL = Behaviors Checklist; EPQ-N = Neuroticism scale of the Eysenck Personality Questionnaire; ASI = Anxiety Sensitivity Index; HR = heart rate; VR = virtual reality.

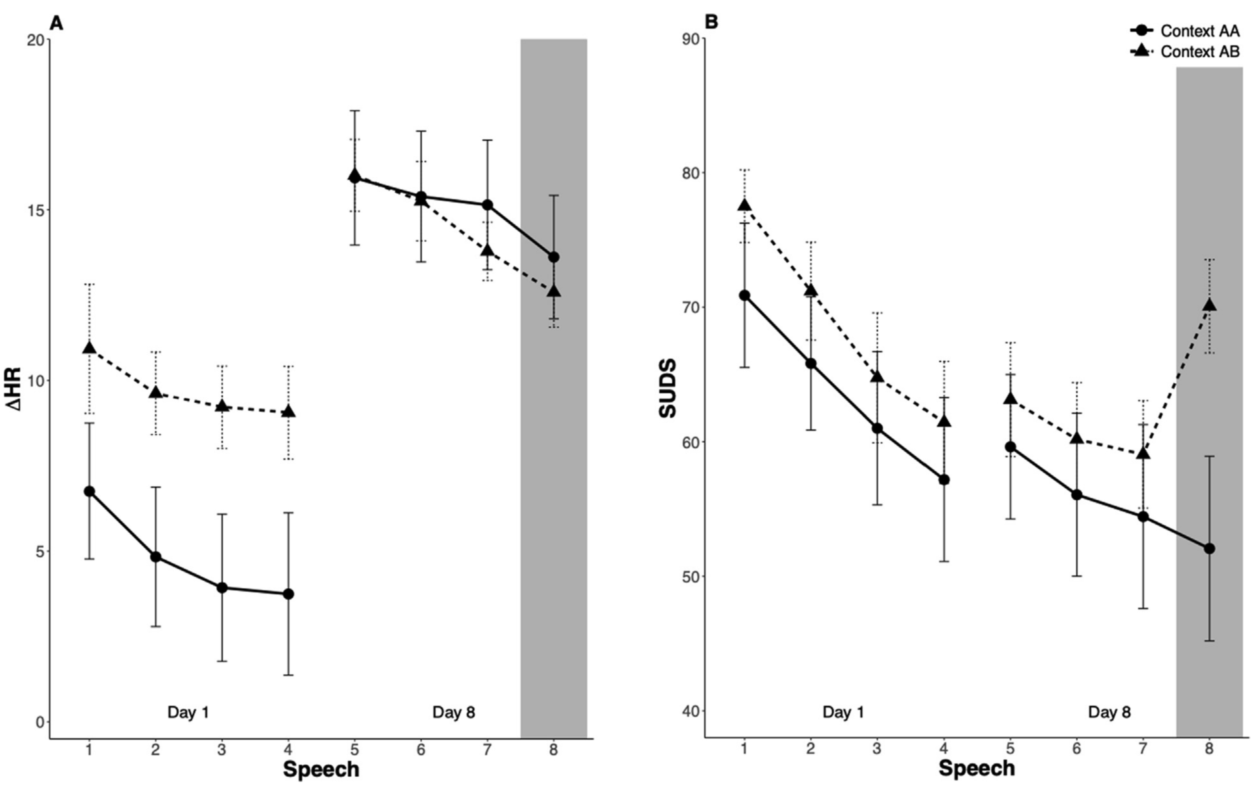

FIGURE 1 Average Δ HR (change scores from baseline) and maximum level of subjective distress (SUDS) during the speeches across groups. *Note.* Error bars reflect standard errors of the mean. They gray shaded areas indicate the context switch; HR = heart rate.

0.33) and there was no evidence for group differences (Time × Group: BF_{10} = 0.33; Group: BF_{10} = 0.57); see Figure 1b. In addition, VR negative valence and arousal ratings did not differ from post-exposure Day 1 to pre-exposure Day 8 (BFs_{10} < 0.26), with no evidence for main or interaction effects for group (BFs_{10} < 0.40); see Figure 2. Thus, both groups demonstrated spontaneous

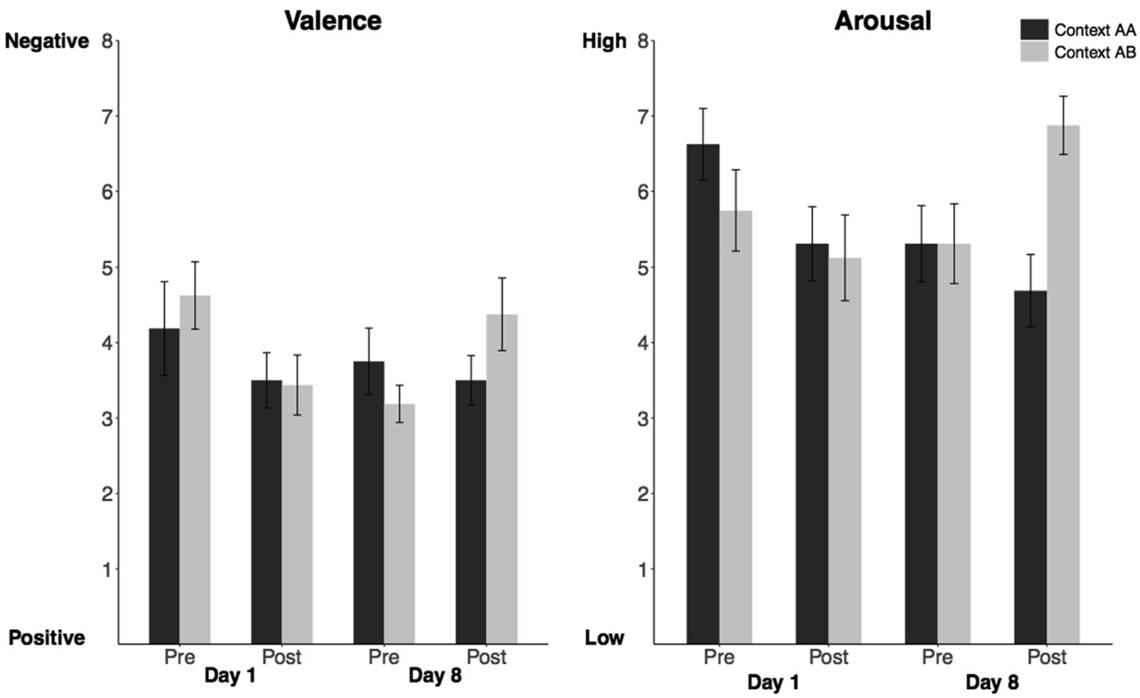

FIGURE 2 Virtual reality audience valence and arousal ratings across groups before and after exposure on Day 1 and Day 8. *Note.* Error bars reflect standard errors of the mean; The context switch occurred before the post-measurement on Day 8.

recovery of HR while giving a speech, but not of subjective distress, negative valence, and arousal ratings.

FEAR RENEWAL

Regarding HR during presenting, there was no evidence for a Time × Group effect ($BF_{10} = 0.40$); see Figure 1a. Yet, for subjective distress, the expected Time × Group effect indicated large differences in renewal between groups ($BF_{10} = 400.93$); see Figure 1b. Post hoc analyses, using paired samples *t*-tests, showed that the AB group reported an increase in SUDS ratings from the third to final speech on Day 8 ($BF_{10} = 82.78$), while the AA group did not ($BF_{10} = 1.22$). We also observed anecdotal evidence for a Time × Group effect for VR valence ($BF_{10} = 2.78$) and strong evidence for a Time× Group effect for VR arousal ratings ($BF_{10} = 14.49$); see Figure 2. Post hoc analyses, using paired samples *t*-tests, showed that the AB group reported an increase in negative valence and arousal ratings from pre- to post-exposure on Day 8 ($BFs_{10} > 3.68$), while there was no evidence for an increase in the AA group ($BFs_{10} < 0.86$). Robustness checks on these post hoc analyses indicated that effects were robust for SUDS, and to a lesser extent for VR valence and arousal (see Figure S3 in Online Supplemental Material). Thus, fear renewal was not observed for HR, but it was observed for subjective distress, negative valence, and arousal ratings.

Discussion

This study aimed to validate a newly developed VR paradigm to examine spontaneous recovery and fear renewal in individuals with fear of public speaking. The main findings can be summarized as follows. First, HR, subjective distress, negative valence, and arousal decreased during exposure. Second, 1 week after exposure training, spontaneous recovery occurred for HR during a presentation, which is in line with previous research (Vasey et al., 2012). Third, fear renewal was observed on all subjective measures, which is consistent with Culver et al. (2011, Study 1), except that they also found fear renewal on HR.

Our study expands previous research on the renewal of public speaking fear (Culver et al., 2011; Shin & Newman, 2018) by using VR to enhance experimental control and to facilitate applications in other research labs. One study also tested the renewal of public speaking fear in VR (Craske et al., 2019), but fear renewal did not occur in their setup. One likely explanation for this discrepancy in findings is that, in contrast to our study, they did not control for spontaneous recovery, which may have obscured their fear renewal effect.

Several findings of the current study should be highlighted. First, we found emotional concordance patterns during exposure (reduced HR, subjective distress, negative valence, and arousal ratings), but not during spontaneous recovery and fear renewal tests. This may reflect random variation across response indices and is consistent with findings of similar studies (e.g., Craske et al., 2019; Vasey et al., 2012) and with the general fear conditioning literature (e.g., Mertens et al., 2018) in which subjective and physiological responses also substantially varied. One plausible explanation for this variation is that lab studies may not always evoke sufficient fear for full emotional concordance to occur (Hollenstein & Lanteigne, 2014). Indeed, in our study, we observed high subjective distress ratings and emotional concordance patterns on the first day, whereas 1 week later, lower subjective distress was associated with weaker emotional concordance patterns. Another explanation may be that autonomic fear responding is highly variable across individuals, with high fear associated with HR increases as well as decreases (Hagenaars et al., 2014). It should be noted that spontaneous recovery was observed only for HR (and not subjective distress), while fear renewal occurred only on subjective distress (and not HR). Potentially, a ceiling effect prevented HR renewal effects because HR was already significantly higher during the presentations on the second test day. Future studies that use this procedure could examine whether spontaneous recovery of HR and renewal of subjective distress is a robust pattern or random variation across response indices.

The second finding that should be highlighted is that in our study, participants indicated that they found presenting in VR less challenging than in real life (see Table 1). Nevertheless, our subjective distress ratings were equal to (e.g., Tsao & Craske, 2000) or higher than (Culver et al., 2011; Shin & Newman, 2018) studies with real-life exposure. This is in line with findings among individuals with spider phobia, who exhibited equal fear levels in a VR and real-life setting (Shiban et al., 2015). Taken together, these findings underscore the potential clinical utility of VR in lowering the threshold to start with exposure, albeit being as fear-provoking as real-life exposure.

A third noteworthy finding was that the patterns of subjective valence and arousal closely mirrored those of subjective distress, although effects were stronger for subjective distress. Future research could examine the unique explanatory value of these measures—for example, by testing strategies aimed at reducing negative valence or arousal (e.g., van Dis et al., 2019).

Our VR procedure may pave the way for testing a variety of important research questions. One important question is whether treatment strategies that modulate emotional memories associated with performance anxiety (Kearns & Engelhard, 2015) could reduce fear renewal. In addition, future research may test whether this procedure could help to identify patients with social anxiety who are at risk for clinical relapse after exposure-based therapy (i.e., predictive validity). Another relevant research avenue could be to add a threat expectancy measure (see van Veen et al., 2020) to examine whether within-session fear reduction and renewal can be explained by expectancy violation.

Several limitations of the current study need to be addressed. First of all, statistical power was sufficient for testing the return of fear but limited for exploring individual differences. In addition, our procedure did not measure avoidance responses even though they may play a critical role in relapse (Craske et al., 2018). Future studies may use measures for behavioral and attentional avoidance, such as eye-tracking, to examine whether and when participants avoid facing the audience. Finally, we included only HR as a physiological outcome measure. Skin conductance and fear-potentiated startle measures can be informative additional indices, which may be included in future research (see Constantinou et al., 2021; van Veen et al., 2020). This may also enlighten concordance patterns across fear indices. Strengths include the controlled experimental setup by using VR and the addition of a spontaneous recovery phase using a 2-day procedure.

To conclude, this VR procedure successfully induced spontaneous recovery of HR and renewal of subjective distress, negative valence, and arousal ratings in individuals with a fear of public speaking. Future studies may use this ecologically valid and well-controlled procedure to test strategies aimed at attenuating the return of fear after exposure in individuals suffering from social anxiety.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.beth.2021.01.005.

References

Aczel, B., Hoekstra, R., Gelman, A., Wagenmakers, E. J., Klugkist, I. G., Rouder, J. N., Vandekerckhove, J., Lee, M. D., Morey, R. D., Vanpaemel, W., Dienes, Z., & van Ravenzwaaij, D. (2020). Discussion points for Bayesian

inference. *Nature Human Behaviour, 4*, 561–563. https://doi.org/10.1038/s41562-019-0807-z.
American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.
Arntz, A., Rauner, M., & van den Hout, M. A. (1995). "If I feel anxious, there must be danger": Ex-consequentia reasoning in inferring danger in anxiety disorders. *Behaviour Research and Therapy, 33*(8), 917–925. https://doi.org/10.1016/0005-7967(95)00032-s.
Barelds, D. P. H., & Luteijn, F. (2002). Measuring personality: A comparison of three personality questionnaires in the Netherlands. *Personality and Individual Differences, 33*(4), 499–510. https://doi.org/10.1016/S0191-8869(01)00169-6.
Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory–II* (2nd ed.). Psychological Corporation.
Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*(10), 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9.
Carleton, R. N., Collimore, K. C., & Asmundson, G. J. G. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders, 21*(1), 131–141. https://doi.org/10.1016/j.janxdis.2006.03.010.
Cieraad, R., & de Jong, P. J. (2007). BFNE-II-NL Geautoriseerde Nederlandse vertaling. Rijksuniversiteit Groningen.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155.
Constantinou, E., Purves, K. L., McGregor, T., Lester, K. J., Barry, T. J., Treanor, M., Craske, M. G., & Eley, T. C. (2021). Measuring fear: Association among different measures of fear learning. *Journal of Behavior Therapy and Experimental Psychiatry, 70*(101618), 1–9. https://doi.org/10.1016/j.jbtep.2020.101618.
Craske, M. G., Fanselow, M., Treanor, M., & Bystritksy, A. (2019). Cholinergic modulation of exposure disrupts hippocampal processes and augments extinction: Proof-of-concept study with social anxiety disorder. *Biological Psychiatry, 86*(9), 703–711. https://doi.org/10.1016/j.biopsych.2019.04.012.
Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(20170025), 1–10. https://doi.org/10.1098/rstb.2017.0025.
Culver, N. C., Stoyanova, M., & Craske, M. G. (2011). Clinical relevance of retrieval cues for attenuating context renewal of fear. *Journal of Anxiety Disorders, 25*(2), 284–292. https://doi.org/10.1016/j.janxdis.2010.10.002.
Eisenberg, M., Kobilo, T., Berman, D. E., & Dudai, Y. (2003). Stability of retrieved memory: Inverse correlation with trace dominance. *Science, 301*(5636), 1102–1104. https://doi.org/10.1126/science.1086881.
Engelhard, I. M., Macklin, M. L., McNally, R. J., Van den Hout, M. A., & Arntz, A. (2001). Emotion- and intrusion-based reasoning in Vietnam veterans with and without chronic posttraumatic stress disorder. *Behaviour Research and Therapy, 39*(11), 1339–1348. https://doi.org/10.1016/S0005-7967(00)00101-7.
Eysenck, H. J., & Eysenck, S. B. G. (1991). *Manual of the Eysenck Personality Scales (EPS Adult)*. Hodder and Stoughton.
Fava, G. A., Grandi, S., Rafanelli, C., Ruini, C., Conti, S., & Belluardo, P. (2001). Long-term outcome of social phobia treated by exposure. *Psychological Medicine, 31*(05), 899–905. https://doi.org/10.1017/s0033291701004020.
First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2016). *Structured Clinical Interview for DSM-5 Disorders: SCID-5-CV clinician version*. American Psychiatric Association.
Georgiou, K., Larentzakis, A. V., Khamis, N. N., Alsuhaibani, G. I., Alaska, Y. A., & Giallafos, E. J. (2018). Can wearable devices accurately measure heart rate variability? A systematic review. *Folia Medica, 60*(1), 7–20. https://doi.org/10.2478/folmed-2018-0012.
Hagenaars, M. A., Oitzl, M., & Roelofs, K. (2014). Updating freeze: Aligning animal and human research. *Neuroscience and Biobehavioral Reviews, 47*, 165–176. https://doi.org/10.1016/j.neubiorev.2014.07.021.
Hollenstein, T., & Lanteigne, D. (2014). Models and methods of emotional concordance. *Biological Psychology, 98*(1), 1–5. https://doi.org/10.1016/j.biopsycho.2013.12.012.
JASP Team (2020). JASP (Version 0.12.2.0) [Computer software].
Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford University Press.
Kearns, M., & Engelhard, I. M. (2015). Psychophysiological responsivity to script-driven imagery: An exploratory study of the effects of eye movements on public speaking flashforwards. *Frontiers in Psychiatry, 6*(115), 1–9. https://doi.org/10.3389/fpsyt.2015.00115.
Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21*(3), 169–184. https://doi.org/10.1002/mpr.1359.
Landkroon, E., Mertens, G., Sevenster, D., Dibbets, P., & Engelhard, I. M. (2019). Renewal of conditioned fear responses using a film clip as the aversive unconditioned stimulus. *Journal of Behavior Therapy and Experimental Psychiatry, 65*(101493), 1–8. https://doi.org/10.1016/j.jbtep.2019.101493.
Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review, 42*, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004.
Mansell, W., & Clark, D. M. (1999). How do I appear to others? Social anxiety and processing of the observable self. *Behaviour Research and Therapy, 37*(5), 419–434. https://doi.org/10.1016/S0005-7967(98)00148-X.
McCroskey, J. (1970). Measures of communication bound anxiety. *Speech Monographs, 37*, 269–277.
Mertens, G., Braem, S., Kuhn, M., Lonsdorf, T. B., van den Hout, M. A., & Engelhard, I. M. (2018). Does US expectancy mediate the additive effects of CS–US pairings on contingency instructions? Results from subjective, psychophysiological and neural measures. *Behaviour Research and Therapy, 110*, 41–46. https://doi.org/10.1016/j.brat.2018.09.003.
Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience, 9*(660), 1–9. https://doi.org/10.3389/fnhum.2015.00660.
Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the

prediction of fearfulness. *Behaviour Research and Therapy, 24*(1), 1–8.

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin and Review, 21*(2), 301–308. https://doi.org/10.3758/s13423-014-0595-4.

Sanderman, R., Arrindell, W. A., Ranchor, A. V., Eysenck, H. J., & Eysenck, S. B. G. (2012). *Het meten van persoonlijkheidskenmerken met de Eysenck Personality Questionnaire (EPQ), een handleiding* (2nd ed.). Research Institute SHARE.

Scheveneels, S., Boddez, Y., & Hermans, D. (2018). Learning mechanisms in fear and anxiety: It is still not what you thought it was. In B. Olatunji (Ed.), *The Cambridge handbook of anxiety and related disorders*. Cambridge University Press (pp. 13–40). Cambridge University Press. https://doi.org/10.1017/9781108140416.002.

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy, 86*, 87–94. https://doi.org/10.1016/j.brat.2016.08.015.

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.

Shin, K. E., & Newman, M. G. (2018). Using retrieval cues to attenuate return of fear in individuals with public speaking anxiety. *Behavior Therapy, 49*(2), 212–224. https://doi.org/10.1016/j.beth.2017.07.011.

Stopa, L., & Clark, D. M. (1993). Cognitive processes in social phobia. *Behaviour Research and Therapy, 31*(3), 255–267. https://doi.org/10.1016/0005-7967(93)90024-O.

Tsao, J. C., & Craske, M. G. (2000). Timing of treatment and return of fear: Effects of massed, uniform-, and expanding-spaced exposure schedules. *Behavior Therapy, 31*(2), 479–497. https://doi.org/10.1016/s0005-7894(00)80026-x.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03.

van Dis, E. A. M., Hagenaars, M. A., Bockting, C. L. H., & Engelhard, I. M. (2019). Reducing negative stimulus valence does not attenuate the return of fear: Two counterconditioning experiments. *Behaviour Research and Therapy, 120*(103416), 1–10. https://doi.org/10.1016/j.brat.2019.103416.

Van Dis, E. A., Van Veen, S. C., Hagenaars, M. A., Batelaan, N. M., Bockting, C. L., Van Den Heuvel, R. M., Cuijpers, P., & Engelhard, I. M. (2020). Long-term outcomes of cognitive behavioral therapy for anxiety-related disorders: A systematic review and meta-analysis. *JAMA Psychiatry, 77*(3), 265–273. https://doi.org/10.1001/jamapsychiatry.2019.3986.

van Veen, S. C., Zbozinek, T. D., Engelhard, I. M., van Dis, E. A. M., & Craske, M. G. (2020). *Positive mood induction does not reduce return of fear: A virtual reality exposure study for public speaking anxiety* (submitted for publication).

Vasey, M. W., Harbaugh, C. N., Buffington, A. G., Jones, C. R., & Fazio, R. H. (2012). Predicting return of fear following exposure therapy with an implicit measure of attitudes. *Behaviour Research and Therapy, 50*(12), 767–774. https://doi.org/10.1016/j.brat.2012.08.007.

Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2013). Extinction, generalization, and return of fear: A critical review of renewal research in humans. *Biological Psychology, 92*(1), 51–58. https://doi.org/10.1016/j.biopsycho.2012.01.006.

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology, 9*(1), 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542.

Vujanovic, A. A., Arrindell, W. A., Bernstein, A., Norton, P. J., & Zvolensky, M. J. (2007). Sixteen-item Anxiety Sensitivity Index: Confirmatory factor analytic evidence, internal consistency, and construct validity in a young adult sample from the Netherlands. *Assessment, 14*(2), 129–143. https://doi.org/10.1177/1073191106295053.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., & Meerhoff, F. (2018). Bayesian inference for psychology, Part II. Example applications with JASP. *Psychonomic Bulletin and Review, 25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7.

Received: July 27, 2020
Accepted: January 17, 2021
Available online: 2 February 2021