Unconscious Interventions for Spider Phobia: Investigating Very Brief Counterconditioning vs. Very Brief Exposure Using Backward Masking

Susanne Schotanus

s5162610

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group number: 22

Supervisor: Dr. Irina Masselman, Prof. Peter de Jong

Second evaluator: Dr. Nienke Jonker

In collaboration with: Imme Arendsen, Ilse Vellinga, Lieke Schipper, Niamh Walsh-Doherty

June 25, 2025

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Because of delays in data collection, fictitious data has been used for this bachelor thesis. The results and conclusions in this bachelor thesis are thus not based on real data of participants and therefore should also not be used as such.

Abstract

Spider phobia is highly prevalent and often difficult to treat with standard exposure therapy due to intense fear responses. Masked interventions such as very brief exposure (VBE) and very brief counterconditioning (VBC) reduce emotional responses without conscious distress. This study investigated the effectiveness of VBC – pairing masked spider images with positive animal stimuli – in altering affective evaluations of spiders, compared to VBE, where masked spider images were followed by neutral letter arrays, and a neutral control condition (CTL), which replaced spider images with a blank screen followed by the mask and positive animal stimuli. A total of 151 female university students rated spider valence before (T0), after the first (T1), and after the second block of trials (T2). A significant time-by-condition interaction showed modest valence increases in VBC and VBE groups, contrasted with a decline in CTL. This suggests the interaction effect was driven by both improvements in the intervention groups and a decline in the CTL. For a second, non-conditioned spider stimulus, only a main effect of time emerged, likely driven by VBE. These findings align with prior studies (Masselman et al., 2024; De Jong et al., 2000), showing no consistent advantage of VBC over VBE. Limitations include a non-clinical, homogeneous sample and brief masked procedures. Future research should explore longer or repeated interventions in clinical populations and incorporate unmasked presentations alongside behavioral or physiological measures. Overall, VBC may influence affective evaluations but is unlikely to produce meaningful clinical benefits without more intensive or sustained approaches.

Unconscious Interventions for Spider Phobia: Investigating Very Brief Counterconditioning vs. Very Brief Exposure Using Backward Masking

Spider phobia is one of the most common specific phobias, with a particularly high prevalence among women (Rosenbaum et al., 2020). A large-scale study found that 5.6% of women and only 1.2% of men met the diagnostic criteria, illustrating a substantial gender difference (Fredrikson et al., 1996). Individuals with severe spider phobia often experience intense physiological and cognitive distress when exposed to spider-related stimuli, leading to avoidance of everyday activities such as reading newspapers or watching television (Johanson et al., 1998). Symptoms may include increased heart rate, sweating, and trembling, and cognitive impairments such as feeling mentally blocked and unable to think clearly. These effects highlight the disruptive impact of spider phobia on daily functioning and underscore the need for effective treatment.

One of the most widely used behavioral treatments for specific phobias, including spider phobia, is exposure therapy (Richtlijnendatabase, 2023). This approach involves repeated, systematic exposure to fear-provoking stimuli to reduce fear responses over time. Although demonstrably effective, supported by a meta-analysis reporting significant fear reduction across various specific phobias (Wolitzky-Taylor et al., 2008), its application faces several challenges. Fear responses may be temporary, as fear can recur over time due to new stressors or contextual changes (Zbozinek & Craske, 2015). Furthermore, the success of exposure therapy heavily depends on individuals' willingness to confront their fears, which frequently hinders treatment engagement (Siegel et al., 2021). Many individuals are understandably reluctant to face their most prominent fears and abandon established avoidance behaviors (Miegel et al., 2025). These limitations underscore the need for alternative interventions that are less distressing and more accessible for individuals with high levels of phobic avoidance.

A promising alternative that addresses this challenge of treatment resistance is very brief exposure (VBE). In VBE, fear-inducing stimuli (e.g., images of spiders) are presented so briefly that they remain outside conscious awareness, thereby minimizing the emotional intensity of the exposure (Siegel & Weinberger, 2009). For this reason, VBE is often referred to as a form of unconscious exposure. Research shows that individuals with spider phobia undergoing VBE approached a live tarantula more closely than those exposed to clearly visible spider images (CVE) (Siegel & Weinberger, 2009). Further neuroimaging research demonstrated that VBE activated the amygdala – a key region involved in fear processing – and engaged automatic fear-processing and regulatory circuits, without eliciting conscious fear (Siegel et al., 2017). In contrast, CVE elicited stronger subjective fear and reduced activation in brain areas responsible for emotion regulation. These findings suggest that initially avoiding direct confrontation with feared stimuli may reduce distress and thereby increase treatment engagement (Siegel & Weinberger, 2009). Building on this, subsequent research found that VBE led to reduced avoidance and self-reported fear of a live tarantula compared to CVE, without the physiological arousal typically observed in CVE (Siegel et al., 2021). Notably, reduced arousal during VBE strongly predicted fear reduction, emphasizing the potential benefits of unconscious exposure in managing phobic responses. However, VBE did not completely eliminate fear of the tarantula, highlighting a key limitation shared by both traditional and masked exposure: although effective in reducing fear, these methods often fail to address the deeper, underlying emotional associations individuals hold toward phobic stimuli (Baeyens et al., 1989). For instance, one study involving 30 students showed that while extinction procedures reduced fear-related responses, negative evaluative associations persisted (Baeyens et al., 1989). In their experiment, neutral facial images were paired with liked or disliked stimuli during acquisition, and later presented without reinforcement. Although fear-related responses decreased, participants continued to rate negatively paired faces as less pleasant, indicating that exposure alone may not fully eliminate underlying negative emotional associations. This points to a key challenge in treatment: addressing both the immediate fear response and the more persistent negative emotional associations linked to phobic stimuli.

One way to address this limitation is by targeting the deeper affective evaluations that influence responses to feared stimuli (Baeyens et al., 1992; Baeyens et al., 1989). These evaluations often contribute to the recurrence of fear and avoidance, hindering long-term treatment success. A promising technique that aims to modify such evaluations is counterconditioning (CC), which involves pairing a previously threatening conditioned stimulus (CS) with an unconditioned stimulus (US) of opposite valence, thereby altering the emotional valence of the CS (Baeyens et al., 1989; Meulders et al., 2015). For example, pairing a fear-related CS with a positive US can weaken or reverse the emotional response, while pairing neutral CSs with aversive USs can induce negative evaluations, demonstrating the bidirectional nature of CC. Unlike extinction, which involves presenting the CS without reinforcement, CC involves new associative learning that can actively overwrite or reverse the negative emotional evaluation. Empirical studies support this: CC reduced fear and improved affective evaluations more effectively than extinction in patients with pain-related fear, as shown by Meulders et al. (2015). Similarly, an experimental study using spider images as CSs found that although fear expectancy at the end of the intervention was similar for both CC and extinction, the CC group exhibited significantly less return of fear at follow-up (Kang et al., 2018). These findings indicate that CC may produce more durable changes in emotional evaluations, thereby reducing the likelihood of relapse.

One of the main challenges in CC is that strong negative emotional responses to feared stimuli often interfere with forming new positive associations (Keller et al., 2020). Intense fear can overshadow learning, making it difficult to replace negative valence with positive affect. Similar to its role in exposure therapy (Siegel & Weinberger, 2009; Siegel et al., 2018; Siegel et al., 2021), masking the feared stimulus during CC, by presenting it below conscious awareness, may reduce distress and improve outcomes compared to non-masked CC (Breitmeyer & Ogmen, 2006). This tempering of immediate emotional responses allows the positive US to more effectively alter the emotional valence of the CS, facilitating associative learning by minimizing interference from intense negative emotions (Keller et al., 2020). Attention also plays a crucial role, as humans naturally focus on threatening stimuli, which can cause negative information to dominate and hinder the formation of positive associations (Baumeister et al., 2001). Masking reduces conscious awareness of threat, potentially reducing this attentional bias. This aligns with Pessoa and Adolphs' (2010) framework, which shows that conscious perception enhances attentional prioritization of emotional stimuli. In this way, masking acts as an emotional "filter", dampening negative reactions and increasing the likelihood that positive emotional learning will be encoded more effectively and with longer-lasting effects.

While masked exposure has been studied (e.g., Siegel & Weinberger, 2009; Siegel et al., 2018; Siegel et al., 2021), masked CC remains relatively underexplored. To date, only one study directly compared these two approaches by examining whether masked CC offers distinct advantages over masked exposure in reducing affective evaluations of spiders (Masselman et al., 2024). In this study, 259 female participants were randomly assigned to four conditions: masked CC, masked exposure, non-masked CC, and non-masked exposure. Participants rated their emotional evaluation of spider images before and after exposure to masked or non-masked spider stimuli, paired with smiling faces (CC) or not (exposure). Both masked exposure and masked CC led to more positive spider evaluations, with no significant differences between techniques or between masked and non-masked CC. These findings raise important questions about the added value of masked CC.

One key factor to consider in understanding why Masselman et al. (2024) found no significant differences is the masking method they used. The study utilized continuous flash suppression (CFS), which presents a fearful stimulus (e.g., a spider) to one eye while showing a dynamic, non-fearful stimulus to the other, suppressing awareness of the fearful stimulus (Tsuchiya & Koch, 2005). CFS allows for longer presentation of the masked stimulus, potentially facilitating associative learning. However, suppression may have been suboptimal because affectively salient stimuli like spiders tend to break through CFS more easily than neutral stimuli due to their enhanced attentional capture (Gayet et al., 2016; Padmala & Pessoa, 2008; Phelps et al., 2006), making them harder to suppress effectively (Tsuchiya & Koch, 2005). Consequently, participants might have been more aware of the spider images than intended, possibly compromising the effectiveness of both masked exposure and masked CC. In contrast, backward masking may offer a more effective method for masking emotional stimuli, as shown in several studies by Siegel et al. (2009, 2011, 2012, 2018). Backward masking occurs when a visual stimulus is rapidly followed by another stimulus (the mask), making the first stimulus difficult to consciously perceive (Breitmeyer & Ogmen, 2006). This technique may be more effective than CFS in minimizing conscious awareness of emotional stimuli (e.g., the spider), thereby potentially increasing the efficacy of CC.

Given the challenges in effectively reducing negative emotional evaluations of feared stimuli, especially when conscious awareness interferes with learning, this study aimed to test whether very brief counterconditioning (VBC) using backward masking can more effectively reduce spider dislike. Due to the high prevalence of spider fears among students, particularly women (Fredrikson et al., 1996; Arrindell, 2000; Seim & Spates, 2009; Rosenbaum et al., 2020), our sample consisted exclusively of female students. Participants were randomly assigned to one of three conditions: (1) a VBC condition, where brief presentations of spider images (CS) were followed by a masked letter array and then a positive animal unconditioned stimulus (US); (2) a VBE condition, in which spider images were followed by a masked letter array and then a neutral letter array; and (3) a non-spider exposure control (CTL) condition, where a blank screen replaced the spider image, followed by a masked letter array and a positive animal US. We predicted that VBC would produce a greater increase in the positive evaluations of the spider images compared to VBE and CTL (hypothesis 1). We also expected VBE to produce more positive evaluations than the CTL condition (hypothesis 2). To evaluate broader emotional change, we also examined whether these effects generalized to a second, non-conditioned spider image, which was not subjected to these manipulations. Generalization is important, as individuals rarely encounter the exact same feared stimulus in daily life (Dunsmoor & Murphy, 2015). We predicted that VBC would produce larger positive changes for this non-exposed spider image compared to VBE and CTL (hypothesis 3), with VBE also outperforming CTL (hypothesis 4).

Methods

Participants

A total of 152 female participants (*M* age = 19.86, *SD* = 1.79) were recruited via the University of Groningen SONA and Paid Participant Pool (PPP) systems. SONA is an online platform that consists of primarily first year psychology students who receive course credits for their participation, while the PPP system involves a broader demographic of students who instead receive monetary compensation. Students were eligible to participate if they gave prior consent via an online prescreening and obtained a score of 77.2 or higher on the Fear of Spiders Questionnaire (FSQ; Szymanski & O'Donohue, 1995), similar to treatment seeking samples (e.g., De Jong & Peters, 2007). Participants were randomly assigned to one of three conditions: Control (CTL: *n* = 51), Very Brief Counterconditioning (VBC: *n* = 52), and Very Brief Exposure (VBE: *n* = 48). Prior to statistical analyses, data from one participant in the VBE condition were excluded due to missing data at T0 for all the affective ratings. Informed consent was obtained from all individuals, and the study received ethical approval from the University of Groningen Ethical Committee under the code PSY-2425-S-0008. The study was preregistered: https://aspredicted.org/9w9n-67gs.pdf.

Materials

Stimuli

The conditioned stimuli (CSs) consisted of 25 frontal-view images of house spiders, taken from Siegel en Weinberger (2012), which were presented in random order during the intervention in the VBC and VBE conditions. An additional spider image, sourced from Masselman et al. (2024), was used as a generalization spider in the questionnaires and was therefore not shown during the experiment blocks. This stimulus served to assess whether the effects of the manipulation would extend to a similar, yet distinct, feared stimulus. To increase stimulus variability and obscure the true focus of the study, ten filler stimuli were included in the affective ratings. These consisted of neutral or negatively valenced animals (e.g., a rat, a snake, or a fly). A scrambled image was used as a visual mask in all conditions. In the VBE condition, this was followed by a neutral array of letters, which served as a non-affective stimulus.

The unconditioned stimuli (USs) were 25 positively valenced animal images (e.g., a puppy or a duckling), which were systematically paired with the target spider images during the experimental blocks in the VBC condition. This was aimed at creating positive associations and reducing fear responses toward spiders. These same images were also presented in the CTL condition, but without preceding spider stimuli, serving as a control for exposure to positive stimuli alone. A stimulus validation study was conducted prior to the main experiment. Approximately 50 participants, recruited via thesis students' social networks, rated 50 positive animal images (sourced online) on a 0–100 valence scale using an anonymous, voluntary, and uncompensated online Qualtrics survey. The highest-rated images were selected as USs to ensure strong positive affective value.

Manipulation

The experiment was conducted using OpenSesame 3.3.6 (Mathôt et al., 2012). Each participant completed two blocks of 25 trials, with each block lasting approximately one minute. The task was presented on a neutral grey background. Participants were instructed to maintain focus on a centrally presented fixation cross throughout the task and were not informed about the nature, content, or sequence of the stimuli.

In each trial, a fixed three-part sequence was presented (see Figure 1): a 33.4 ms spider image (in the VBC and VBE conditions) or a neutral blank screen (CTL condition), followed by a 117 ms scrambled mask, and finally a 600 ms image. This final image was either a positively valenced animal (in the VBC and CTL conditions) or a neutral letter array (in the VBE condition). While the spider and animal images were randomized across trials, the neutral letter array in the VBE condition remained constant. Participants' subjective distress was assessed using the Subjective Units of Distress (SUD) scale, which was programmed directly into the task and presented immediately before and after each block.

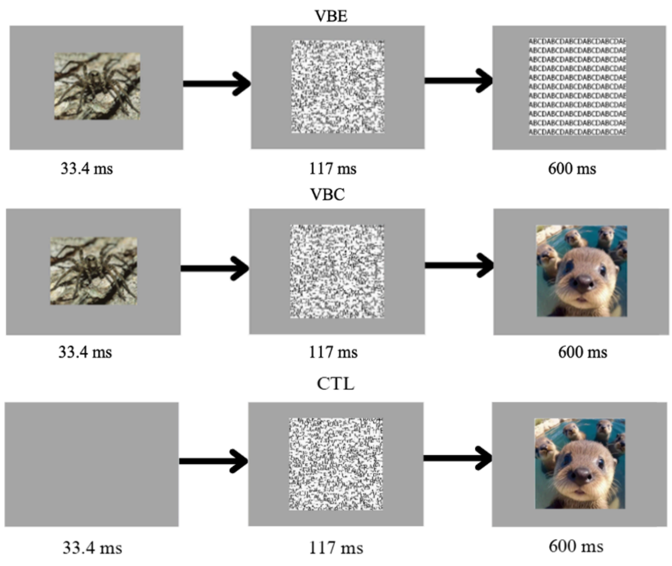


Figure 1

Example of a single intervention trial for the VBE, VBC, and CTL condition

Measures

Affective Ratings

Valence ratings were used to assess participants' subjective evaluation of the animal stimuli. The stimuli in the surveys were presented sequentially using Qualtrics, with a Visual Analogue Scale (VAS) displayed below each image (0 = *Very negative*, 100 = *Very positive*). Participants were shown an image of the spider CS, a generalization spider, followed by US animal and filler animal images. The spider CS and the generalization spider were presented first and in a fixed order (spider CS first, generalization spider after) to ensure their evaluation was not influenced by the other animal stimuli in the survey. The USs and filler animals followed after and were presented in a randomized order. Ratings for the USs were averaged. Out of the 25 spider images used in the experimental blocks, only one spider was rated as the spider CS in the surveys to avoid drawing attention to spider-related stimuli. The generalization spider was also rated but was not part of the experimental blocks. Higher scores indicated more positive evaluations of the animals (*range*: 0-100).

Mental Behavioral Approach Test (BAT)

Participants completed a mental behavioral approach test (BAT) to assess their mental willingness to approach the CS spider. Participants were shown the spider CS and were asked to imagine that the spider was placed in a glass jar on a table before them. They were instructed to indicate to which step of the BAT they would be willing to proceed. Steps ranged from minimal (0) to maximal engagement (8), with minimal engagement being ''none of the above'', and maximal engagement being ''let the spider walk over my hands''. The steps were based on standardized BAT steps used in prior research (de Jong et al., 2000), and are included in Appendix A. Higher scores indicated a stronger (mental) approach to the spider.

After completing the mental BAT, participants' subjective fear and disgust were assessed. They were instructed to imagine having to execute the final step of the mental BAT. First, they were presented with the spider CS and were asked "Imagine that this spider walks over your hands. How fearful would you be?". Afterwards, they were asked "Imagine that this spider walks over your hands. How disgusted would you be?". Scores for both variables were rated on a VAS ranging from 0 (*not at all*), to 100 (*extremely*), with higher scores indicating greater levels of fear and disgust.

Subjective Distress

Participants' level of subjective distress was assessed using the Subjective Unit of Distress (SUD) scale. At the top of the experimental manipulation page participants were asked "How do you feel right now?". Scores were rated on a VAS and ranged between 0 (*no fear whatsoever*) to 100 (*an extremely high, unbearable amount of fear*), with higher scores indicating a greater level of subjective distress.

Spider Fear

Spider-specific fear was measured using the 18-item FSQ (Szymanski & O'Donohue, 1995). The items assessed avoidance behaviors, fear, and cognitive preoccupation with spiders (e.g., "If I came across a spider now, I would get help from someone else to remove it" or "If I came across a spider now, I would leave the room"). Responses were rated on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*). Total FSQ scores were calculated by summing all items. Scores ranged from 18 to 126, with higher scores indicating a greater degree of spider fear. Participants also completed the FSQ during the prescreening, allowing for a baseline FSQ score. The internal consistency of the FSQ was high in the current sample with a Cronbach's alpha of .95 for both the baseline FSQ and the FSQ measured at T2.

Contingency Awareness and Confidence

The participants' contingency awareness in the VBC condition was assessed via a VAS in which participants reported their estimation of the CS-US contingency. Specifically, participants were asked to estimate the proportion of non-spider animal images that were preceded by a spider image, with responses ranging from 0% (*none*) to 100% (*all*). Higher percentages reflected a greater level of perceived contingency awareness. Participants were also asked to indicate how confident they were in their contingency estimate, using a separate VAS ranging from 0% (*not at all confident*) to 100% (*extremely confident*). Higher percentages indicated greater confidence in the participants' own contingency estimates.

Other Awareness and Manipulation Checks

A complete set of funneled awareness questions, designed to assess participants' attention to the task, stimulus awareness, contingency awareness and perceived study purpose, is provided in Appendix B. To clarify the intended reference stimulus in each question, the wording of these questions included slight variations across conditions.

Procedure

This study was advertised as an investigation into how viewing animals that one likes or dislikes affects emotional responses. On arrival at the laboratory, participants were given a printed information letter, and asked to give written informed consent. Participants were informed that participation was voluntary and that they could withdraw at any time without negative consequences. The experiment took place in a quiet, dimly lit testing room to minimize distractions. All participants were seated in front of a computer screen (1920 x 1080 IIYAMA ProLite G2773HS, refresh rate = 100 Hz) and asked to position their head in a chin rest to maintain a standardized viewing distance during the experimental tasks. The procedure consisted of three survey time points (T0, T1, T2) and two short experimental blocks. Prior to the start of the procedure, participants were randomly assigned to one of the three experimental conditions in a between-subject design. Participants began with the first questionnaire (T0), in which they evaluated the spider CS, followed by the generalization spider and subsequently the 25 USs and 10 filler animals. Next and directly prior to the first experimental block, participants' SUD was measured. Subsequently, participants completed the first experimental block of 25 trials in alignment with the allocated condition. Thereafter (T1a), the SUD of the participants was re-assessed. The second questionnaire (T1) was then conducted, with participants rating all previously assessed animals in a comparable manner to the first survey. Furthermore, the participants completed the questions that measured the mental BAT, subjective fear and subjective disgust. Baseline assessments at T0 for these measurements were omitted to avoid drawing attention to the spider-related stimuli. After this, participants' SUD was reassessed prior to the start of the second experimental block (T1b). The following second experimental block employed the same condition-specific manipulation as the earlier block. Immediately following the second block and before the start of the final survey, the SUD was assessed again (T2). The final survey (T2) included the same measurements as the second, along with the FSQ and a funneled awareness interview. In the final stage of the questionnaire, participants reported their age. Following the submission of the final survey, participants were asked about their overall experience and were given the opportunity to share any discomfort or distress they may have encountered during the procedure. The participants were then compensated with course credits or monetary compensation. To minimize the risk of information spreading that could influence the ongoing data collection, a full debriefing was postponed until after the study concluded.

Statistical analyses

To test the hypotheses, two repeated measures analyses of variance (ANOVA) were conducted. The first analysis examined whether the increase in positive evaluation of the CS spider was greater in the VBC condition compared to the VBE and CTL, with valence rating of the CS spider as the dependent variable (testing hypothesis 1 and 2). The second analysis followed the same procedure, but focused on the generalization spider's valence rating as dependent variable (testing hypothesis 3 and 4). In both analyses, Condition (CTL, VBE, VBC) was the between-subjects factor, and Time (pre-manipulation ratings T0 vs. post-Block 2 ratings at T2) the within-subject factor. To assess an overall effect of two sessions of the same manipulation and optimize statistical power by reducing the number of comparisons, both analyses focused on T0 and T2, excluding T1. We expected a significant Time × Condition interaction in both analyses, indicating that changes in spider evaluation over time would differ between conditions, with the largest increase in the VBC condition. Main effects of Time and Condition were also examined, to identify overall changes across time and differences between groups, providing a fuller understanding of the data. Within-group changes were explored using post hoc paired-sample t-tests. These tests assessed whether the change in CS and generalization spider valence within each condition was statistically significant, providing insight into the direction and magnitude of the change over time. For all analyses, a significance level of $\alpha = .05$ was used.

Power calculation

An a priori power analysis for the RM ANOVA was conducted using G*Power 3.1 (Faul et al., 2009), which indicated that a sample of 159 participants would provide a power of .80 to detect a medium effect ($f = .25$) at an alpha level of .05. The final sample consisted of 151 participants.

Results

Participants

An overview of the condition means and standard deviations for the descriptive variables is provided in Table 1. There were no significant age differences between the three groups, $F(2, 148) = 1.34$, $p = .27$. FSQ pre-scores, measured before T0, also did not differ significantly between the groups, $F(2, 148) = 0.26$, $p = .78$. These results indicate that the groups were comparable at baseline in terms of age and fear of spiders, suggesting that any differences observed at later time points are unlikely to be due to pre-existing group differences.

Table 1

Condition Means and Standard Deviations per Descriptive Variable

	CTL (n = 51)	VBE (n = 48)	VBC (n = 52)
	M (SD)	M (SD)	M (SD)
Age	20.20 (2.08)	19.69 (1.48)	19.69 (1.76)
FSQPre	96.51 (16.50)	95.12 (17.69)	97.52 (23.58)

Note. FSQ = Fear of Spider Questionnaire, administered prior to T0 as part of the prescreening. Scores range from 18 to 126, with higher scores indicating greater levels of spider fear.

Valence Ratings of the CS Spider

For an overview of the means and standard deviations for the outcome variables at T0, T1, and T2 per condition, see Table 2.

Table 1

Means and Standard Deviations for Outcome Variables at T0, T1, and T2 per Condition

		CTL (n = 51)			VBE (n = 48)			VBC (n = 52)		
		T0	T1	T2	T0	T1	T2	T0	T1	T2
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Valence	Spider CS	21.18 (17.38)	23.98 (19.84)	15.24 (19.27)	15.50 (17.85)	21.84 (16.98)	20.24 (17.56)	15.44 (15.64)	20.48 (20.16)	21.58 (17.62)
	Spider 2	20.06 (17.21)	21.69 (17.70)	21.69 (17.70)	16.35 (16.00)	21.31 (18.69)	21.31 (18.69)	15.15 (15.58)	17.58 (17.85)	17.58 (17.85)
	USs	70.02 (11.48)	68.48 (12.01)	68.48 (12.01)	69.87 (12.45)	67.96 (12.60)	67.96 (12.60)	70.91 (12.53)	68.07 (13.52)	68.07 (13.52)
Fear of Spider		—	65.14 (2C3.74)	63.18 (23.63)	—	63.22 (28.92)	61.31 (28.74)	—	68.31 (22.73)	66.31 (22.73)
Mental BAT		—	4.16 (2.34)	3.96 (2.05)	—	3.88 (2.53)	4.51 (2.26)	—	3.69 (2.59)	3.29 (2.51)
SUD		10.10 (5.89)	11.12 (6.03)/ 11.57 (5.39)	9.33 (6.67)	70.20 (18.05)	69.29 (18.04)/ 67.33 (18.40)	65.16 (20.30)	68.96 (16.24)	68.77 (16.99)/ 70.44 (17.85)	65.31 (19.87)
FSQPost		—	—	70.84 (23.84)	—	—	69.73 (25.10)	—	—	72.92 (23.58)

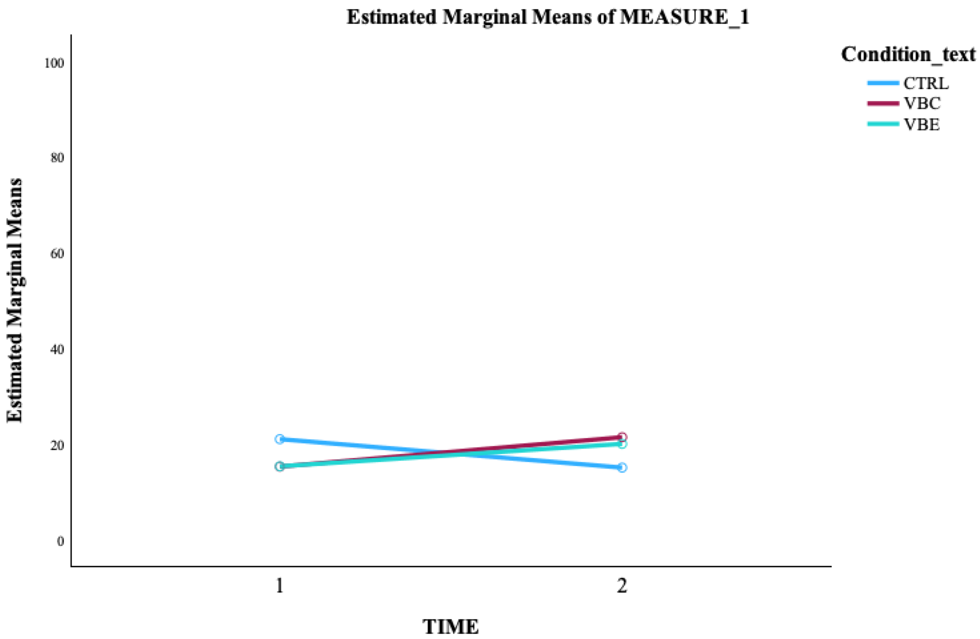
Note. Valence (range 0-100; higher scores indicate a more positive evaluation of the CS, spider 2, and USs); spider 2 = generalization spider, which was not included in the manipulation; USs = mean valence of 25 positively valenced animal stimuli; fear of spiders (range 0-100; higher scores indicate greater spider fear); mental BAT (range 0-8; higher scores indicate greater willingness to approach the spider); SUD = Subjective Units of Distress (range 0-100; higher scores indicate greater current fear); at T1, SUD was assessed twice: before and after manipulation; FSQ post= Fear of Spider Questionnaire, administered after the second manipulation block (range 18-126, higher scores indicating greater levels of spider fear).

Prior to data analysis, assumption checks were conducted to ensure the appropriateness of the statistical tests. Shapiro-Wilk tests indicated significant deviations from normality for CS spider scores at T0 and T2 across all three conditions ($p < .05$). However, given the relatively large sample size (n = 151), these violations were not considered problematic, as large samples tend to minimize the impact of non-normality (Field, 2024). Visual inspection of Q-Q plots and boxplots suggested approximate normality, although several outliers were identified at T0 (one in CTL and one in VBE) and T2 (three in CTL and three in VBE). To assess their potential influence, analyses were performed both with and without these cases. Levene's tests indicated no clear violations of homogeneity of variance at T0, $F(2,148) = 0.62$, $p = .54$, and at T2, $F(2,148) = 0.25$, $p = .78$. Box's M test indicated no significant violation of homogeneity of covariance matrices, *Box's M* = 2.49, $F(6, 529143.35) = 0.41$, $p = .88$. As only two time points (T0 and T2) were analyzed, sphericity was not assessed. Linearity between T0 and T2 was supported by scatterplots. Therefore, repeated measures ANOVA was deemed appropriate for analyzing changes in valence over time.

A repeated measures ANOVA was conducted to examine the effect of condition (VBC, VBE, and CTL) on valence ratings of the CS spider from T0 and T2. There was no significant main effect of time, $F(1,148) = 0.62$, $p =.43$, $\eta_p^2 = .004$, indicating no overall change in valence ratings across all conditions. Similarly, no significant main effect of condition was found, F (2, 148) = 0.04, $p = .96$, $\eta_p^2 = .00$, suggesting that overall valence ratings were comparable between groups regardless of time. Thus, condition alone did not influence how positively the CS spider was perceived. However, a significant interaction between time and condition emerged, $F(2,148) = 3.46$, $p = .03$, $\eta_p^2 = .05$, indicating that change in valence over time differed between conditions (see Figure 2). To evaluate the impact of outliers, the analysis was repeated excluding these outliers. The results remained consistent: no main effect of time, $F(1, 140) = 0.20$, $p = .65$, $\eta_p^2 = .001$; no main effect of condition, $F(2, 140) = 0.06$, $p = .95$, $\eta_p^2 = .001$; and the Time × Condition interaction remained significant, $F(2, 140) = 3.24$, $p = .04$, $\eta_p^2 = .04$. This confirms that outliers did not substantially affect the overall findings.

Figure 2

Average Valence Ratings of the CS spider at each Time Point per Condition



To further explore the significant Time × Condition interaction, paired-samples t-tests were conducted within each condition to examine whether the mean difference in valence ratings between T0 and T2 was statistically significant. Although directional hypotheses predicted increased valence in the active conditions (VBC and VBE) compared to CTL, two-tailed tests were used for a conservative approach. Assumption checks were performed on the difference scores (T2 – T0) for each condition using Shapiro-Wilk test, Q-Q plots, and boxplots. In the CTL group, the Shapiro-Wilk test indicated a significant deviation from normality, $W = .94$, $p = .02$. However, the Q-Q plot suggested approximate normality, with two outliers identified in the boxplot. In the VBC group, normality was supported by both the Shapiro-Wilk test, $W = .98$, $p = .47$, and visual inspections, with no outliers detected. Similarly, the VBE group showed no violation of normality according to the Shapiro-Wilk test, $W = .99$, $p = .83$, with normal Q-Q plot patterns, but one outlier identified in the boxplot. Although paired-samples t-tests are generally robust to such deviations (Field, 2024), sensitivity analysis was conducted by re-running the tests excluding these cases. Given that T0 and T2 ratings were obtained from the same participants, paired-samples t-tests were appropriate to compare these dependent measurements.

Results indicated no significant mean differences in valence ratings for the CTL group, $t(50) = 1.65$, $p = .11$, Cohen's $d = 0.23$; the VBE group, $t(47) = -1.20$, $p = .24$, Cohen's $d = -0.17$; or in the VBC group, $t(51) = -1.91$, $p = .06$, Cohen's $d = -0.27$. Despite the significant interaction, these within-group changes were small and non-significant, suggesting that while the pattern of change over time differed between groups, the direction or magnitude of change within each group was modest. These findings partially align with hypothesis 1 and 2, which predicted greater positive valence increases in VBC and VBE than in CTL, though effects were small and non-significant within individual groups. Excluding outliers in the CTL group revealed a significant increase in valence from T0 to T2, $t(48) = 2.91$, $p = .006$, Cohen's $d = 0.42$. Excluding the single outlier in the VBE group did not change the results, $t(46) = 0.86$, $p = .39$, Cohen's $d = -0.13$. As no outliers were identified in the VBC group, the results of the paired samples t-tests remained unchanged. These sensitivity analyses suggest that outliers influenced the CTL results but not those of the VBC and VBE conditions. Moreover, the significant increase observed in the CTL group after outlier removal indicates that factors other than the experimental manipulation may have contributed to valence changes.

Valence Ratings of the Generalization Spider

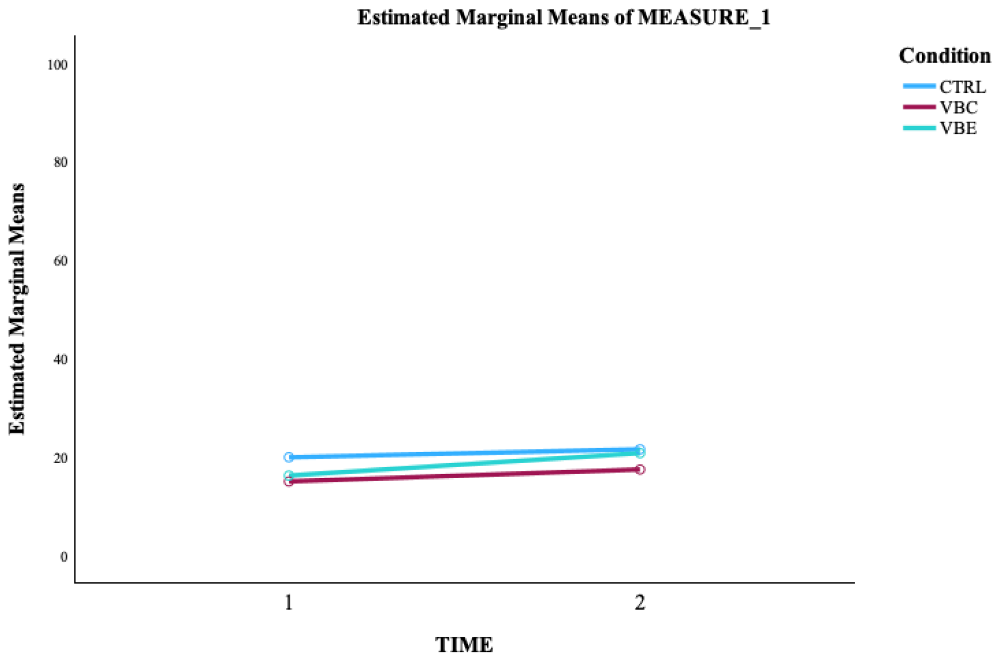
To ensure the suitability of analyses of the generalization spider (spider 2) valence ratings, relevant statistical assumptions were evaluated. Shapiro-Wilk tests indicated significant deviations from normality at T0 and T2 across all conditions ($p < .05$). Nonetheless, given the relatively large sample size (n = 151), these deviations were considered acceptable, as parametric tests are less affected by non-normality when the sample size is large (Field, 2024). Visual inspection of Q-Q plots and boxplots suggested approximately normal, despite two outliers in the CTL condition at T2. To assess their impact, analyses were performed both with and without these cases. Levene's tests indicated no significant violation of homogeneity of variances at T0, $F(2,148) = 0.48$, $p = .62$, or at T2, $F(2,148) = 0.64$, $p = .53$. Box's M test showed no significant violation of homogeneity of covariance matrices, *Box's M* = 1.91, $F(6, 529143.35) = 0.31$, $p = .93$. Since only two time points were analyzed, sphericity was not tested. Furthermore, scatterplots indicated a linear relationship between scores at T0 and T2, supporting the linearity assumption.

A repeated measures ANOVA was conducted to examine changes in the valence ratings of the generalization spider (spider 2) from T0 to T2 across the three experimental conditions (CTL, VBE, VBC), The analysis revealed a significant main effect of time, $F(1,148) = 8.11$, $p = .005$, $\eta_p^2 = .05$, indicating that participants overall rated the generalization spider more positively at T2 compared to T0. The main effect of condition was not significant, $F(2, 148) = 1.01$, $p = .37$, $\eta_p^2 = .01$, suggesting that valence ratings did not differ significantly between the conditions regardless of time. Furthermore, the interaction between time and condition was not significant, $F(2,148) = 0.73$, $p = .48$, $\eta_p^2 = .01$, indicating that changes in valence over time did were comparable across between the conditions. Figure

3 displays the mean valence scores for the generalization spider across time for each condition. Although a slight increase in valence over time was visually observable in the VBE group, these patterns should be interpreted with caution given the absence of a statistically significant interaction. To ensure that the results were not driven by the identified outliers, the analysis was repeated excluding these cases. The pattern of results remained consistent: the main effect of time remained consistent, $F(1, 146) = 7.54$, $p = .01$, $\eta_p^2 = .05$; the main effect of condition remained non-significant, $F(2, 146) = 1.05$, $p = .35$, $\eta_p^2 = .01$; and the Time × Condition interaction remained non-significant, $F(2, 146) = 0.68$, $p = .51$, $\eta_p^2 = .01$. This confirms that outliers did not substantially influence the overall findings.

Figure 3

Average Valence Ratings of the generalization spider at each Time Point per Condition



To further examine the significant main effect of time, paired-samples *t*-tests were conducted within each condition to compare generalization spider ratings from T0 to T2. Although directional hypotheses predicted greater increases in valence in the active conditions (VBC and VBE) compared to CTL, two-tailed tests were used to maintain a conservative approach. Prior to analysis, the normality of the difference scores (T2 – T0) was assessed per condition using the Shapiro-Wilk test, Q-Q plots, and boxplots. The Shapiro-Wilk tests indicated deviations from normality in the CTL group ($W = .93$, $p = .004$), the VBC group ($W = .96$, $p = .046$), and the VBE group ($W = .92$, $p = .002$). Nevertheless, Q-Q plots suggested approximate normality, and boxplots identified several outliers: nine in CTL, three in VBC, and nine in VBE. Although paired-samples t-tests are generally robust to such deviations (Field, 2024), sensitivity analyses excluding these outliers were conducted to assess their impact. Because the T0 and T2 valence ratings were obtained from the same individuals, paired-samples t-tests were appropriate given the dependent structure of the data.

No significant change in valence ratings was found in the CTL group, $t(50) = -0.91$, $p = .37$, Cohen's $d = 0.13$, or in the VBC group, $t(51) = -1.52$, $p = .14$, Cohen's $d = -0.21$. In contrast, the VBE group showed a statistically significant increase, $t(47) = -2.46$, $p = .02$, Cohen's $d = -0.35$, reflecting a small to moderate effect size. These findings help clarify the significant main effect of time in the repeated measures ANOVA: although participants overall rated the generalization spider more positively at T2 compared to T0, this effect appears to have been primarily driven by the increase observed in the VBE group. However, the non-significant interaction effect suggests this group-specific change cannot be confidently distinguished from those in the other groups. Sensitivity analysis excluding outliers revealed notable differences. In the CTL group, removing nine outliers did not alter the result; the change remained non-significant, $t(41) = -0.46$, $p = .65$, Cohen's $d = -0.07$. However, in the VBC group, excluding three outliers led to a significant increase in valence ratings from T0 to T2, $t(48) = -2.30$, $p = .03$, Cohen's $d = -0.33$, which had not been significant in the initial analysis. For the VBE group, excluding of nine outliers did not change the outcome, $t(38) = -2.10$, $p = .04$, Cohen's $d = -0.34$. These findings suggest that in the VBC group, outliers may have masked a significant effect, while in the CTL and VBE groups, the results remained consistent, indicating minimal influence of extreme values in those conditions.

Discussion

The study examined whether a VBC intervention, in which masked spider images were consistently followed by unconditioned animal stimuli, would lead to more positive affective evaluations of the CS spider and a generalization spider, compared to a VBE procedure or a CTL condition without spider-related stimuli. It was further hypothesized that the VBE intervention would result in more positive affective valence ratings than the CTL condition.

The way participants' affective evaluations of the CS spider changed over time differed by condition. Although no strong differences were found when examining groups or time points separately, combining both showed notable differences. This suggests that the intervention influenced not only how people felt towards the spider, but also how those feelings developed over time. Descriptively, the CTL group showed an initial increase in valence ratings from baseline (T0 = 21.18) to the first post-manipulation (T1 = 23.98), followed by a sharp decrease at the second post-manipulation (T2 = 15.24) (see Table 2). This unexpected drop after T1 is noteworthy, as one might assume that without any intervention, affective evaluations would remain stable. The initial increase may reflect a temporary familiarity or task-related effect due to repeated exposure or the presence of surrounding positive stimuli. However, the subsequent drop could indicate an affective contrast effect: repeated exposure to positive stimuli without pairing them with the spider may have made the spider seem more negative by comparison. Additionally, because the CTL group received no active intervention, boredom or disengagement may have contributed. In contrast, both the VBE and VBC groups showed relatively stable or slightly increasing valence ratings over time, with the VBC group showing the most consistent upward trend. These trends are visually reflected in Figure 2. However, these improvements were modest and not statistically robust. Importantly, the decline in the CTL group remained clear even after excluding extreme cases, suggesting that the effect was not driven by a few outliers. Together, these findings suggest that passive exposure to positive stimuli – when not meaningfully paired with the feared object – may not necessarily be neutral and may even worsen affective evaluations over time. The decline in the CTL group likely contributed to the overall differences between conditions, indicating that the observed interaction reflects both modest improvements in the VBE and VBC groups and a negative shift in the CTL group. This complicates interpretation, as the interaction is not solely driven by positive intervention effects.

Further examination of valence changes within each group further confirmed this pattern: when including all participants, no clear changes were found, but after excluding outliers, the decline in the CTL group became evident. The VBE and VBC groups showed slight, though uncertain improvements, with the VBC group showing the most promising trend ($p = .06$). These patterns provide partial support for the CS-related hypotheses. Hypothesis 1 (VBC > VBE/CTL) was not strongly supported, as the VBC group showed only a modest, non-significant upward trend, and did not significantly outperform the VBE group. Hypothesis 2 (VBE > CTL) received limited support. Although the VBE group showed slightly more positive affective evaluations over time compared to the CTL group, these differences were not statistically significant. Therefore, while trend is in the expected direction, the evidence is not strong enough to draw firm conclusions.

From a clinical perspective, these results cautiously suggest that VBC or VBE interventions might influence how people feel about feared stimuli. However, given the small magnitude of change and the lack of clear within-group improvements, such short interventions are unlikely to produce meaningfully clinical effects on their own. At the same time, the decline in the CTL group also indicates that ‘neutral’ control conditions may have unintended adverse effects, highlighting that the absence of an intervention does not always imply the absence of an effect.

Affective evaluations of the generalization spider (Spider 2) became more positive over time, regardless of condition. This suggests, on average, participants rated the generalization spider more positively at the end of the experiment compared to baseline. This pattern remained even after excluding outliers: the improvement persisted, but there were still no clear differences between the conditions or in how their responses changed over time.

When looking more closely within each group, only participants in the VBE condition showed a meaningful increase in valence ratings over time. After excluding outliers, the VBC group showed a similar trend, suggesting that initial variability may have obscured this effect. The CTL group, by contrast, showed no substantial change after outlier removal. These patterns, reflected in Figure 3, suggest that the overall improvement was primarily driven by the VBE group, with some contribution from VBC. However, since the groups did not significantly differ in how their ratings evolved, we cannot conclude that either intervention was more effective than the CTL condition in altering evaluations of the generalization spider. As such, hypothesis 3 (VBC > VBE/CTL) was not supported for the generalization stimulus, while hypothesis 4 (VBE > CTL) received limited support. Overall, these results suggest that VBE may lead to some generalization effects, but the added value of CC over exposure remains unclear.

From a clinical perspective, the general improvement indicates that affective evaluations of unexposed but similar stimuli (i.e., generalization spider) can shift positively over time. However, because this change occurred across groups rather than being specific to one intervention, it is uncertain whether this effect can be truly attributed to the intervention conditions. The VBE group contributed most to this shift, suggesting it may have facilitated some generalization, although the absence of a significant interaction limits the strength of this conclusion. The VBC group showed a similar trend after outlier removal. These findings align with Masselman et al. (2024), who also observed that masked exposure alone was sufficient to improve affective evaluations not only of the target spider, but also of a second, non-exposed spider. This supports the idea that masked exposure may promote generalization even without conscious awareness or explicit learning. However, this pattern contrasts with earlier findings (e.g., Baeyens et al., 1989), which suggested that pairing CS with positive USs typically results to stronger reductions in negative evaluations than exposure alone. Consequently, while brief interventions like VBE may support small positive shifts in affective generalization, their clinical relevance remains uncertain without more robust or sustained effects.

Our findings align with those of Masselman et al. (2024) and De Jong et al. (2000), who also found no clear superiority of VBC over VBE in altering affective evaluations of spiders. Like our study, Masselman et al. (2024) observed modest increases in valence ratings following exposure alone, raising doubts about the added value of pairing feared stimuli with positive USs under masked conditions. They further suggested that masking may interfere with contingency learning, limiting the effectiveness of CC procedures. Additionally, their non-clinical sample may have contributed to the weak effects observed, a factor that may similarly apply to our predominantly female student sample. These converging findings highlight the difficulty of achieving strong clinical effects through brief, masked interventions and underscore the need for longer, repeated, or more ecologically valid approaches.

In addition, several limitations should be acknowledged. The sample consisted solely of female university students aged 18 to 22 years, recruited via the University of Groningen's SONA system or Paid Participant Pool. Therefore, participants' motivation may have been driven by course requirements or compensation than genuine interest. Moreover, participants likely varied in spider fear severity and were not formally diagnosed, which limits the generalizability of the findings to clinical populations. The intervention itself was also very brief, which may not have been sufficient or elicit strong or lasting changes in affective evaluations. These factors limit the generalizability and clinical relevance of the findings, especially to more diverse, treatment-seeking, individuals from different educational or cultural backgrounds.

Further studies should include more diverse and clinically relevant populations, such as individuals formally diagnosed with spider phobia, to evaluate whether these very brief interventions hold therapeutic potential. Additionally, increasing the duration or frequency of the intervention could clarify whether longer or repeated exposure and conditioning phases produce more robust and lasting effects. Future research might also compare masked and unmasked versions of counterconditioning to better understand the role of awareness in affective learning, as some findings (e.g., Masselman et al., 2024) suggest that masked presentations may impair contingency awareness. Finally, incorporating behavioral and physiological outcome measures could provide a more comprehensive understanding of intervention effects beyond self-reported valence ratings. However, it is important to note that such measures have limitations: for instance, physiological responses like increased heart rate can be triggered by various non-specific factors and may not always accurately reflect fear or emotional change (Mauss et al., 2005). Therefore, a multi-method approach combining self-report, behavior, and physiological data is recommended to more reliably capture the complexity of affective responses.

Taken together, our findings indicate that brief, masked interventions can modestly influence affective evaluations of feared stimuli over time. The changes depended not only on the intervention type but also on temporal patterns, including an unexpected decline in the CTL group, suggesting that neutral conditions are not always inert. Although VBE showed some potential in generalizing effects to related stimuli, neither VBE nor VBC clearly outperformed the other. These findings underscore the importance of further research to optimize brief interventions and explore their therapeutic potential in more diverse and clinical populations.

References

Arrindell, W. A. (2000). Phobic dimensions: IV. The structure of animal fears. *Behaviour Research and Therapy*, *38*(5), 509–530. https://doi.org/10.1016/s0005-7967(99)00097-2

Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, *30*(2), 133-142. https://doi.org/10.1016/0005-7967(92)90136-5

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1989). Acquired affective-evaluative value: Conservative but not unchangeable. *Behaviour Research and Therapy*, *27*(3), 279-287. https://doi.org/10.1016/0005-7967(89)90047-8

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review Of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Breitmeyer, B., & Ogmen, H. (2006). *Visual masking: Time Slices through Conscious and Unconscious Vision*, *2nd ed.* Oxfort University Press. https://doi.org/10.1093/acprof:oso/9780198530671.001.0001

De Jong, P. J., & Merckelbach, H. (2000). Phobia-relevant illusory correlations: The role of phobic responsivity. *Journal of Abnormal Psychology*, *109*(4), 597–601. https://doi.org/10.1037/0021-843x.109.4.597

De Jong, P. J., & Peters, M. L. (2007). Contamination vs. harm-relevant outcome expectancies and covariation bias in spider phobia. *Behaviour Research and Therapy*, 45(6), 1271–1284. https://doi.org/10.1016/j.brat.2006.09.007

De Jong, P. J., Vorage, I., & Van Den Hout, M. A. (2000). Counterconditioning in the treatment of spider phobia: effects on disgust, fear and valence. *Behaviour Research*

And Therapy, *38*(11), 1055–1069. https://doi.org/10.1016/s0005-7967(99)00135-7

Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: How humans generalize fear. *Trends in Cognitive Sciences*, *19*(2), 73-77. https://doi.org/10.1016/j.tics.2014.12.003

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Field, A. (2024). Discovering statistics using IBM SPSS statistics (6th ed.). Sage.

Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour Research and Therapy*, *34*(1), 33 39. https://doi.org/10/1016/0005-7967(95)00048-3

Gayet, S., Paffen, C. L., Belopolsky, A. V., Theeuwes, J., & Van Der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, *149*, 77–83. https://doi.org/10.1016/j.cognition.2016.01.009

Johanson, A., Gustafson, L., Passant, U., Risberg, J., Smith, G., Warkentin, S., & Tucker, D. (1998). Brain function in spider phobia. *Psychiatry Research: Neuroimaging*, *84*(2-3), 101-111. https://doi.org/10.1016/S0925-4927(98)00051-1

Kang, S., Vervliet, B., Engelhard, I. M., Van Dis, E. A., & Hagenaars, M. A. (2018). Reduced return of threat expectancy after counterconditioning versus extinction. *Behaviour Research And Therapy*, *108*, 78–84. https://doi.org/10.1016/j.brat.2018.06.009

Keller, N. E., Hennings, A. C., & Dunsmoor, J. E. (2020). Behavioral and neural processes in counterconditioning: Past and future directions. *Behaviour Research and Therapy*, 125(1). https://doi.org/10.1016/j.brat.2019.103532

Masselman, I., Glashouwer, K. A., & de Jong, P. J. (2024). What you don’t know, can’t hurt you: The differential effect of masked versus non-masked counterconditioning and mere exposure to spider pictures on women's affective evaluation of spiders. *Journal of Experimental Psychopathology*, *15*(1), 20438087231224338. https://doi.org/10.1177/20438087231224338

Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavioral Research Methods, 44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

Meulders, A., Karsdorp, P. A., Claes, N., & Vlaeyen, J. W. (2015). Comparing counterconditioning and extinction as methods to reduce fear of movement-related pain. *Journal Of Pain*, *16*(12), 1353–1365. https://doi.org/10.1016/j.jpain.2015.09.007

Miegel, F., Lindner, P., Schultz, J., Lohse, L., Yassari, A. H., & Jelinek, L. (2025). Breaking barriers? Targeting willingness to improve engagement in exposure therapy through a disorder-nonspecific exposure session in virtual reality. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues.* https://doi.org/10.1007/s12144-024-07194-9 1-17.

Padmala, S., & Pessoa, L. (2008). Affective learning enhances visual detection and responses in primary visual cortex. *Journal Of Neuroscience*, *28*(24), 6202–6210. https://doi.org/10.1523/jneurosci.1233-08.2008

Pessoa, L., & Adoplhs, R. (2010). Emotion processing and the amygdala: From a 'load road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773-782. https://doi.org/10.1038/nrn2920

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, *17*(4), 292–299. https://doi.org/10.1111/j.1467-9280.2006.01701.x

Richtlijnendatabase. (2023). *Psychotherapie algemeen bij angststoornissen.*

https://richtlijnendatabase.nl/richtlijn/angst_en_dwangstoornissen/psychotherapie_bij_angst_en_dwangstoornissen_2023/psychotherapie_algemeen_bij_angststoornissen_2023.html

Rosenbaum, D., Leehr, E. J., Kroczek, A., Rubel, J. A., Int-Veen, I., Deutsch, K., Maier, M. J., Hudak, J., Fallgatter, A. J., & Ehlis, A. (2020). Neuronal correlates of spider phobia in a combined fNIRS-EEG study. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-69127-3

Seim, R. W., & Spates, C. R. (2009). The prevalence and comorbidity of specific phobias in college students and their interest in receiving treatment. *Journal Of College Student Psychotherapy*, *24*(1), 49–58. https://doi.org/10.1080/87568220903400302

Siegel, P., & Weinberger, J. (2009). Very brief exposure: The effects of unreportable stimuli on fearful behavior. *Consciousness And Cognition*, *18*(4), 939–951. https://doi.org/10.1016/j.concog.2009.08.001

Siegel, P., Anderson, J. F., & Han, E. (2011). Very brief exposure II: The effects of unreportable stimuli on reducing phobic behavior. *Consciousness and Cognition, 20*(2), 181-190. https://doi.org/10.1016/j.concog.2010.09.003

Siegel, P., & Weinberger, J. (2012). Less is more: The effects of very brief versus clearly visible exposure. *Emotion*, *12*(2), 394–402. https://doi.org/10.1037/a0026806

Siegel, P., Warren, R., Wang, Z., Yang, J., Cohen, D., Anderson, J. F., Murray, L., & Peterson, B. S. (2017). Less is more: Neural activity during very brief and clearly visible exposure to phobic stimuli. *Human Brain Mapping*, *38*(5), 2466–2481. https://doi.org/10.1002/hbm.23533

Siegel, P., Warren, R., Jacobson, G., & Merritt, E. (2018). Masking exposure to phobic stimuli reduces fear without inducing electrodermal activity. *Psychophysiology*, *55*(5). https://doi.org/10.1111.psyp.13045

Siegel, P., Cohen, B., & Warren, R. (2021). Nothing to fear but fear itself: A mechanistic test of unconscious exposure. *Biological Psychiatry*, *91*(3), 294–302. https://doi.org/10.1016/j.biopsych.2021.08.022

Szymanski, J., & O'Donohue, W. (1995). Fear of spiders questionnaire. *Journal of behavior therapy and experimental psychiatry, 26*(1), 31–34. https://doi.org/10.1016/00057916(94)00072-t

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096–1101. https://doi.org/10.1038/nn1500

Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review*, *28*(6), 1021-1037. https://doi.org/10.1016/j.cpr.2008.02.007

Zbozinek, T. D., & Craske, M. G. (2017). Positive affect predicts less reacquisition of fear: Relevance for long-term outcomes of exposure therapy. *Cognition and Emotion*, *31*(4), 712-725. http://doi.org/10.1080/02699931.2016.1142428

Appendix A

Mental Behavioral Approach Test

The steps used in the Mental Behavioral Approach Test (BAT) are as follows:

0 = none of the above

1 = walk towards the spider as near as I can

2 = touch the jar

3 = open the jar

4 = take the jar in my hands

5 = touch the spider with a pencil

6 = hold the jar upside-down to put the spider in a washing bowl

7 = touch the spider with a finger

8 = let the spider walk over my hands

Appendix B

Funneled Awareness Questionnaire

This questionnaire was administered at the end of the study to assess participants' awareness of the stimuli, task contingencies, and their engagement during the experiment.

Motivation Check

Did you fill out the questions seriously?

Demand Awareness

What is the purpose of this study, according to you?

Stimulus Awareness

When you were watching the computer screen, what do you remember happening?

Flash Perception Check (Condition-dependent)

In between the X and the black-and-white pattern presented on the screen, did you see something?

- If you were in the VBE condition:

☐ No ☐ A flash ☐ Something else

- If you were not in the VBE condition:

We don't mean the animal pictures that were presented after the black-and-white pattern.

☐ No ☐ A flash ☐ Something else

Flash Content Guess

Something was flashed on the screen between the X and black-and-white pattern. If you had to guess what the flash was, what would you say it was?

Flash Frequency Estimate

How many times did you see this?

Contingency Awareness

Is there anything else that you noticed with regard to order of these pictures?

☐ No ☐ Yes

Contingency Awareness - Detail

If Yes: What did you notice?

Incidental Spider Awareness

During the task you were exposed to pictures of spiders and bugs. How many spiders did you see?

Contingency Estimation

How many of the other animals were preceded by a spider picture do you think? Please indicate the percentage.

Confidence Rating

Please indicate how confident you are in your estimation:

Attention Check

Did you pay attention to the screen during the entire computer task? If you didn't (e.g., looked away purposefully), please select 'no' below.

☐ Yes ☐ No

Appendix C

Figure C1

QQ-Plot of Spider CS at T0 for CTL-Condition

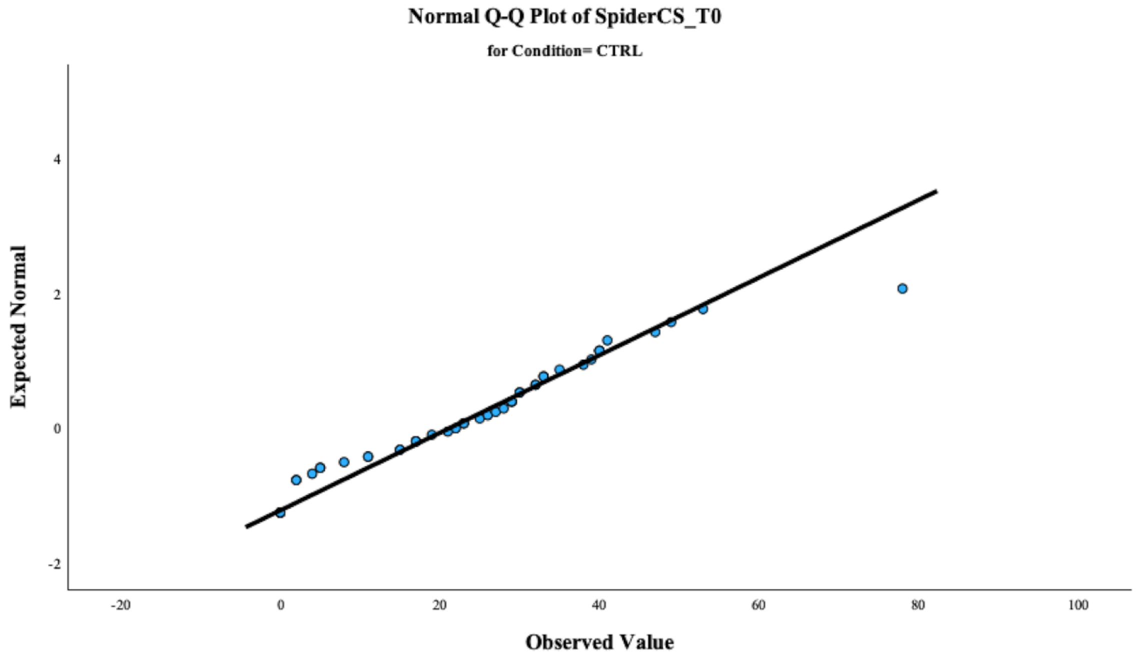


Figure C2

QQ-Plot of Spider CS at T0 for VBE-Condition

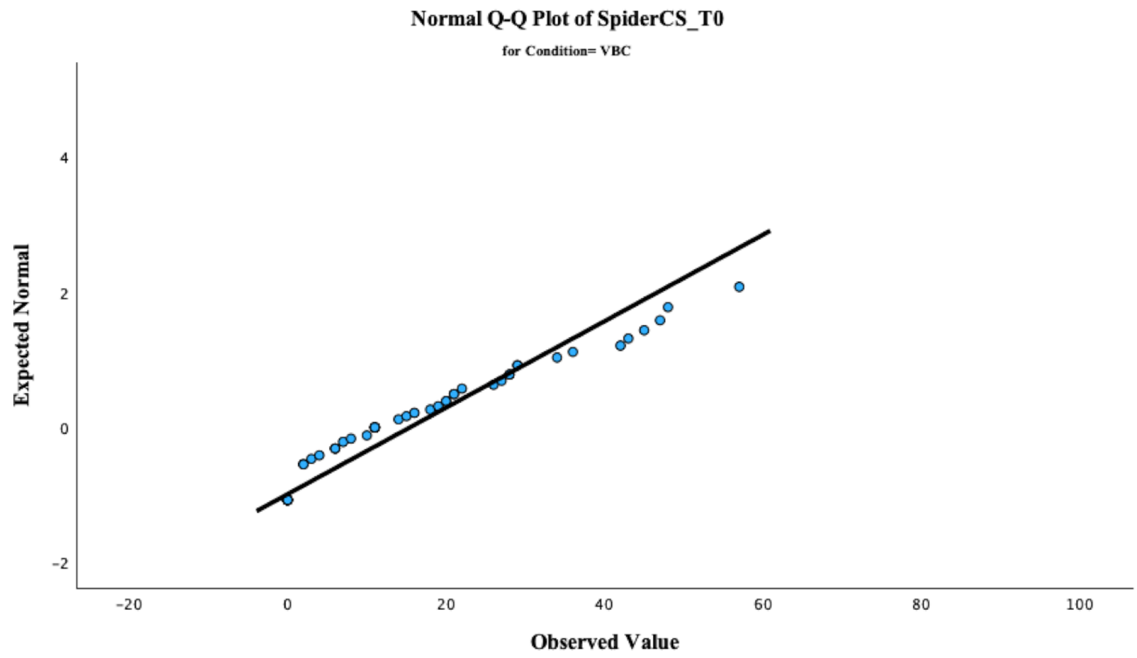


Figure C3

QQ-Plot of Spider CS at T0 for VBC-Condition

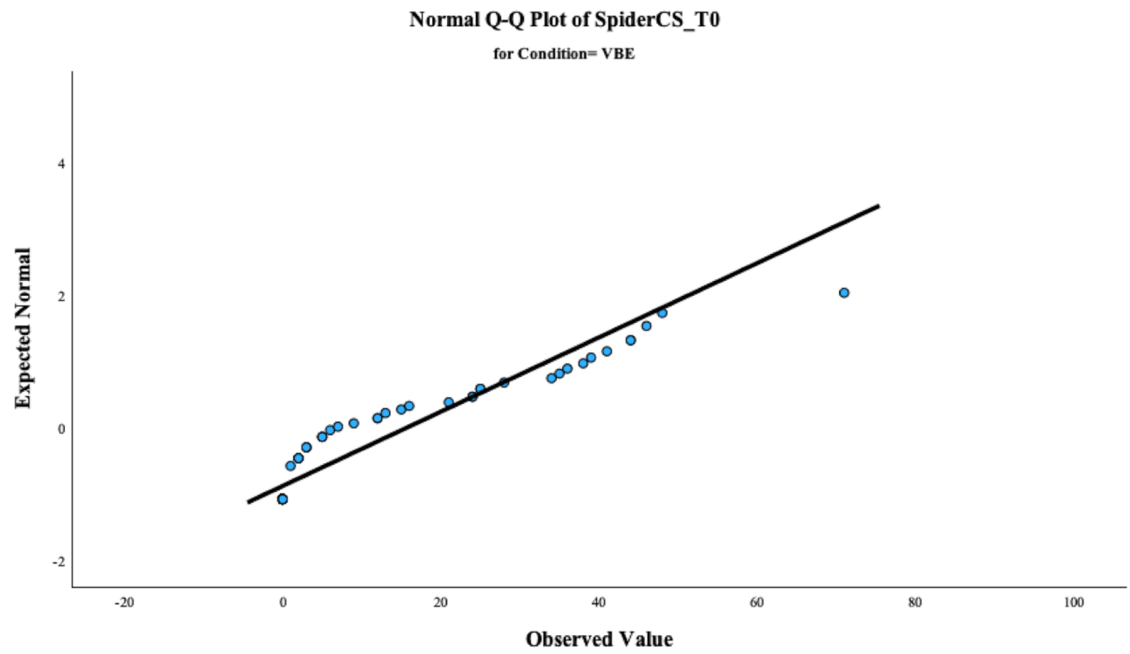


Figure C4

Boxplots of Spider CS at T0 for each Condition

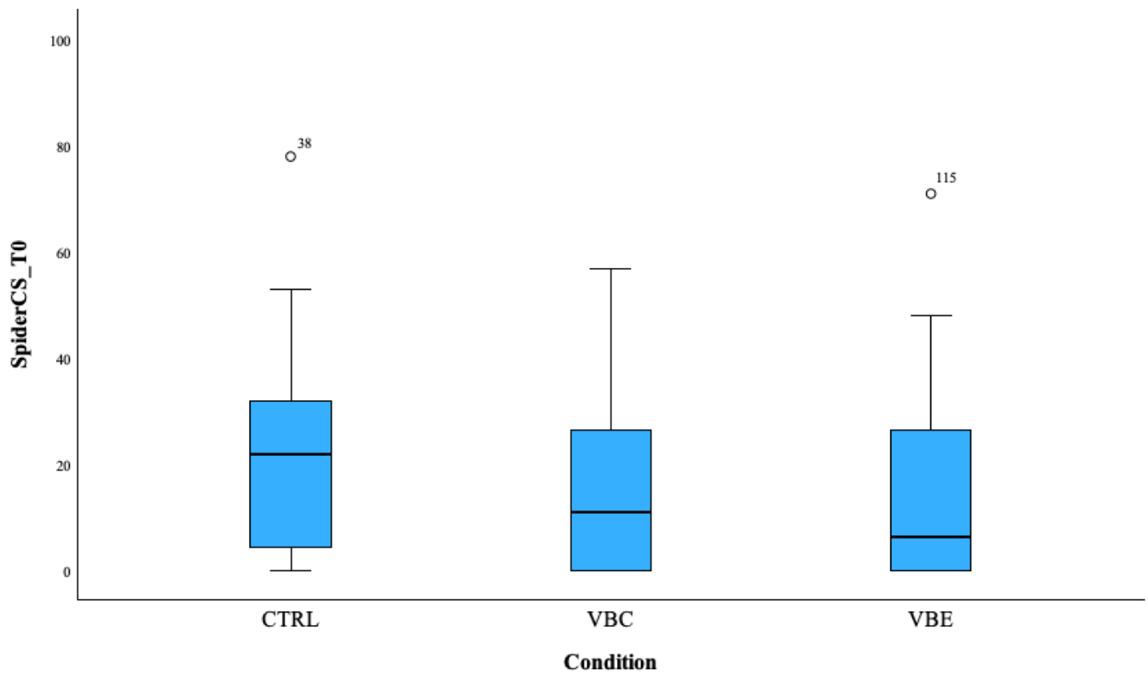


Figure C5

QQ-Plot of Spider CS at T2 for CTL-Condition

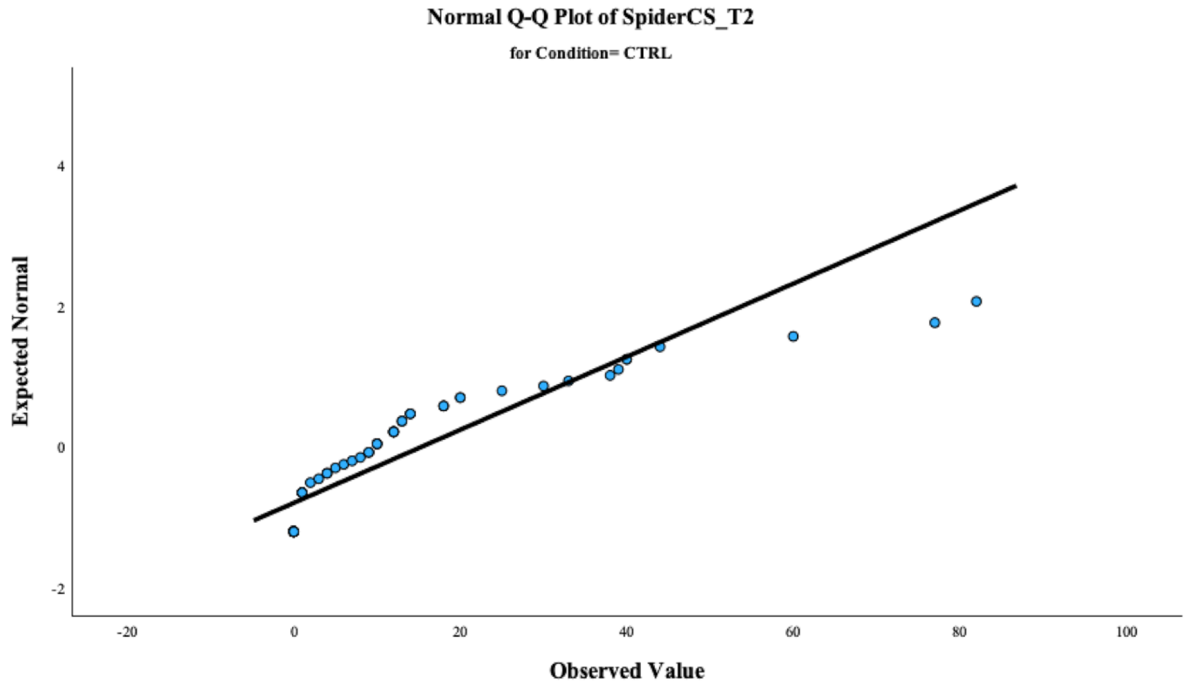


Figure C6

QQ-Plot of Spider CS at T2 for VBE-Condition

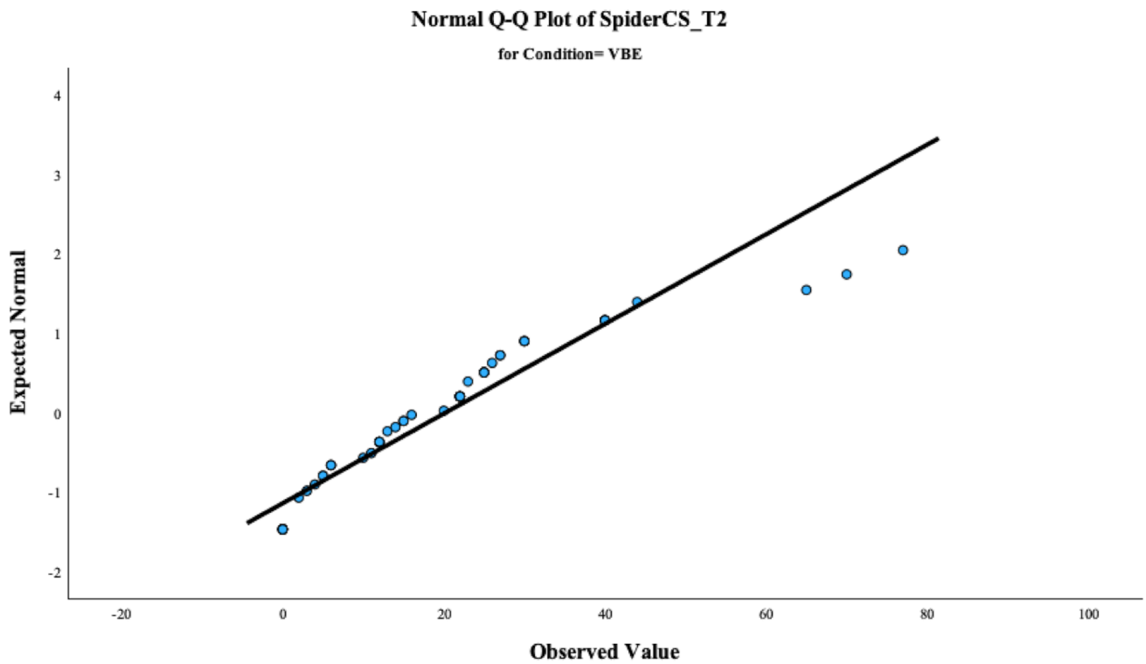


Figure C7

QQ-Plot of Spider CS at T2 for VBC-Condition

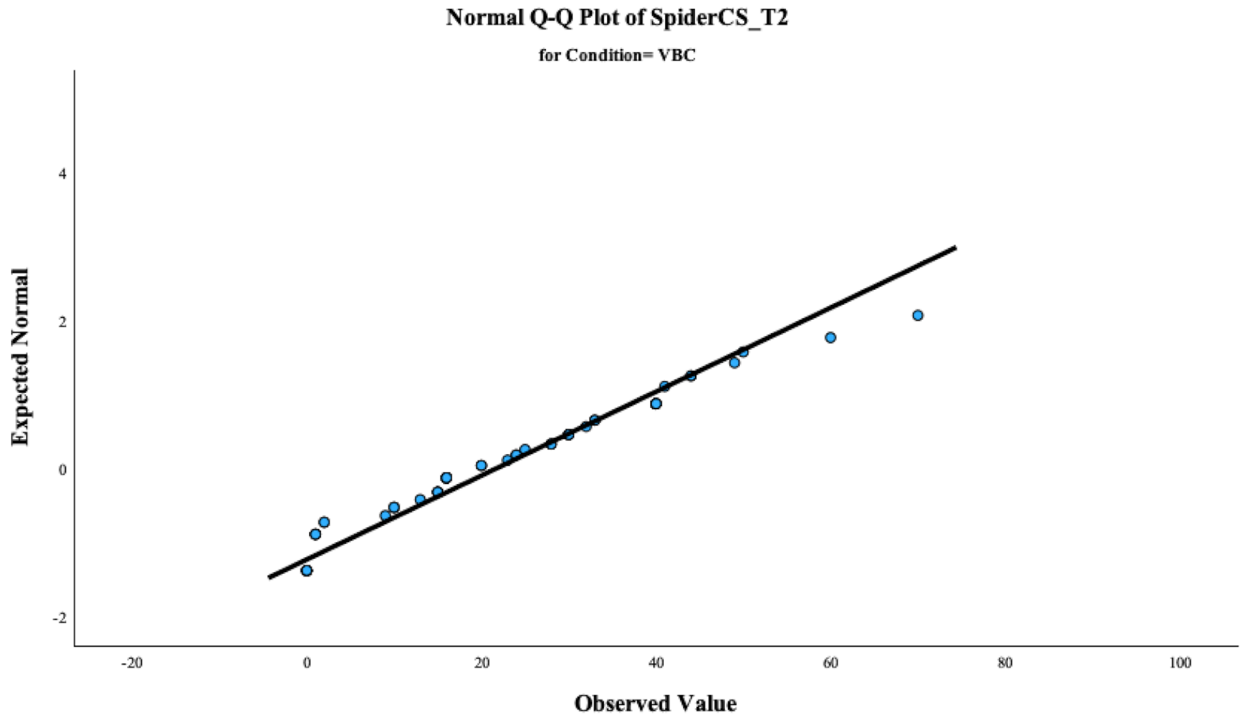
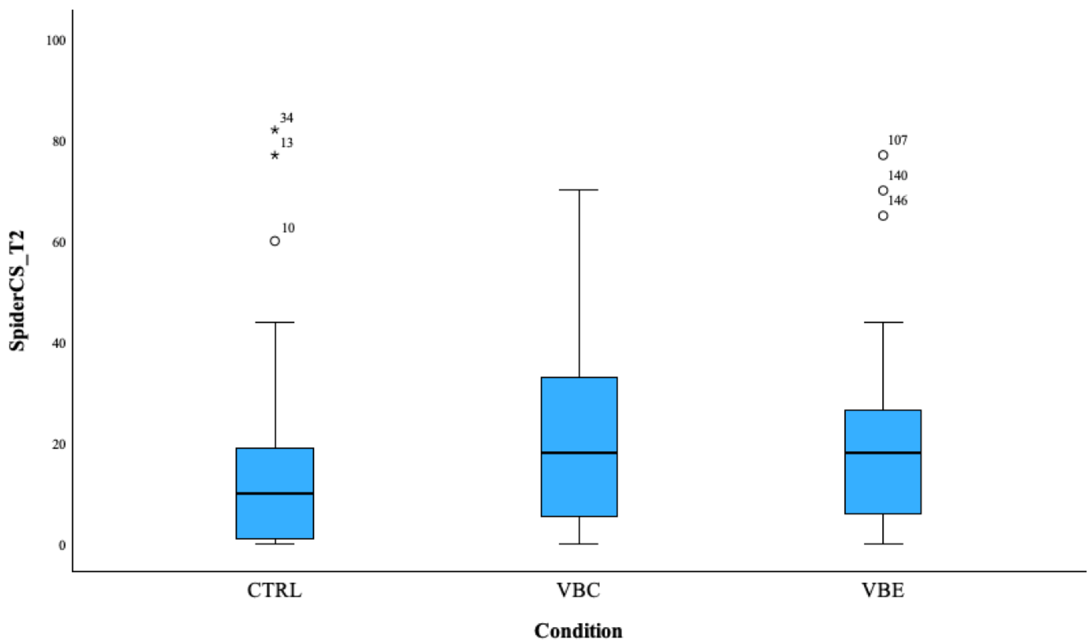


Figure C8

Boxplots of Spider CS at T2 for each Condition



Appendix D

Figure D1

Scatterplot of Valence Ratings of the CS Spider for CTL-Group

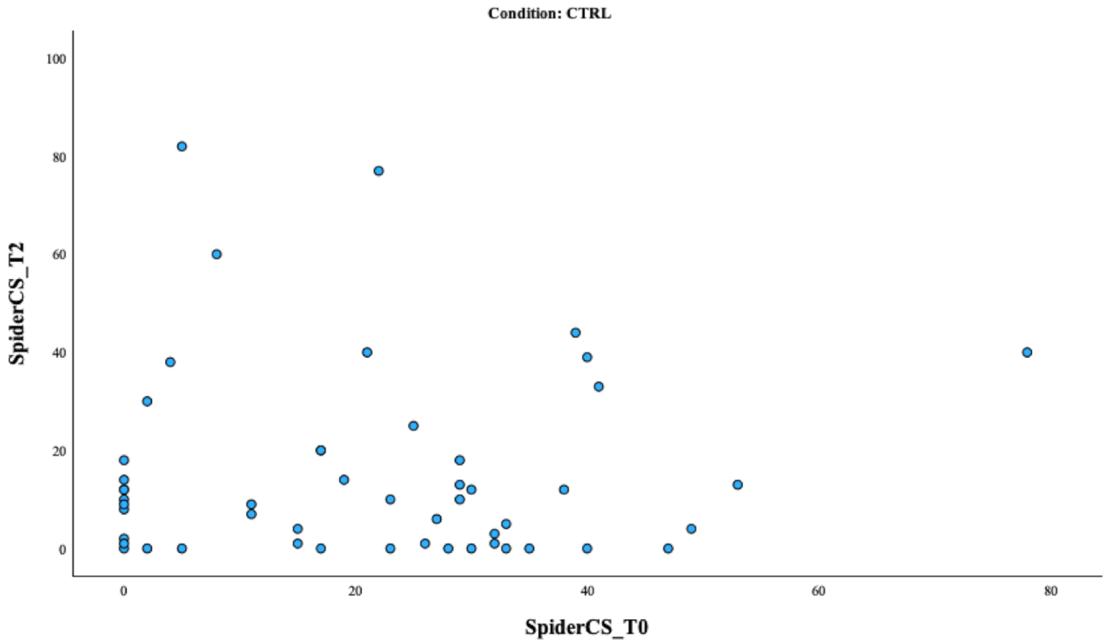


Figure D2

Scatterplot of Valence Ratings of the CS Spider for VBE-Group

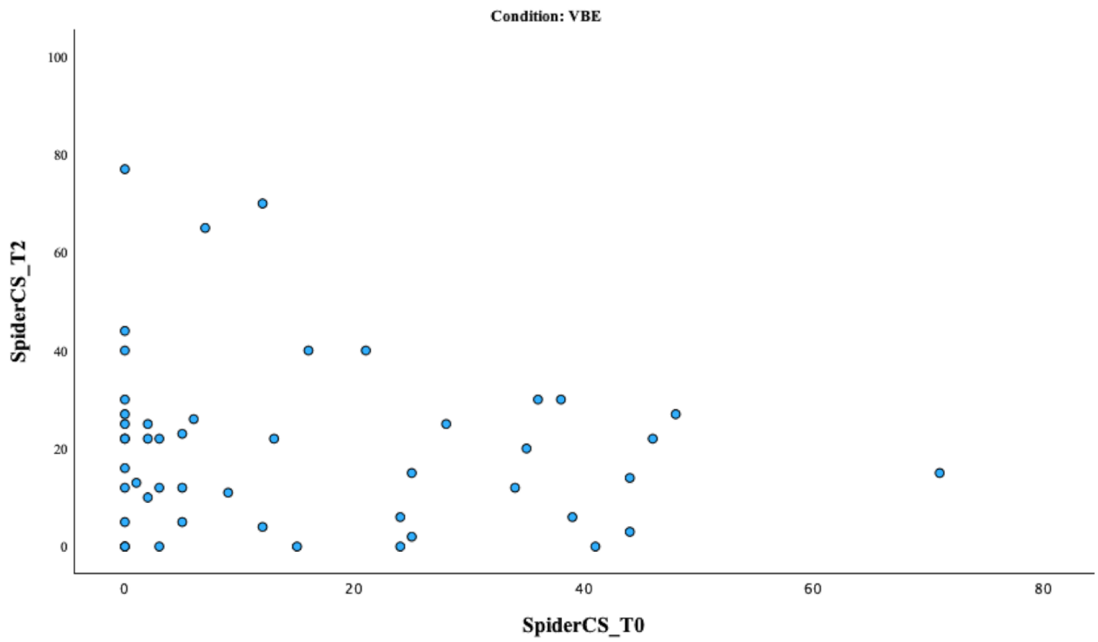
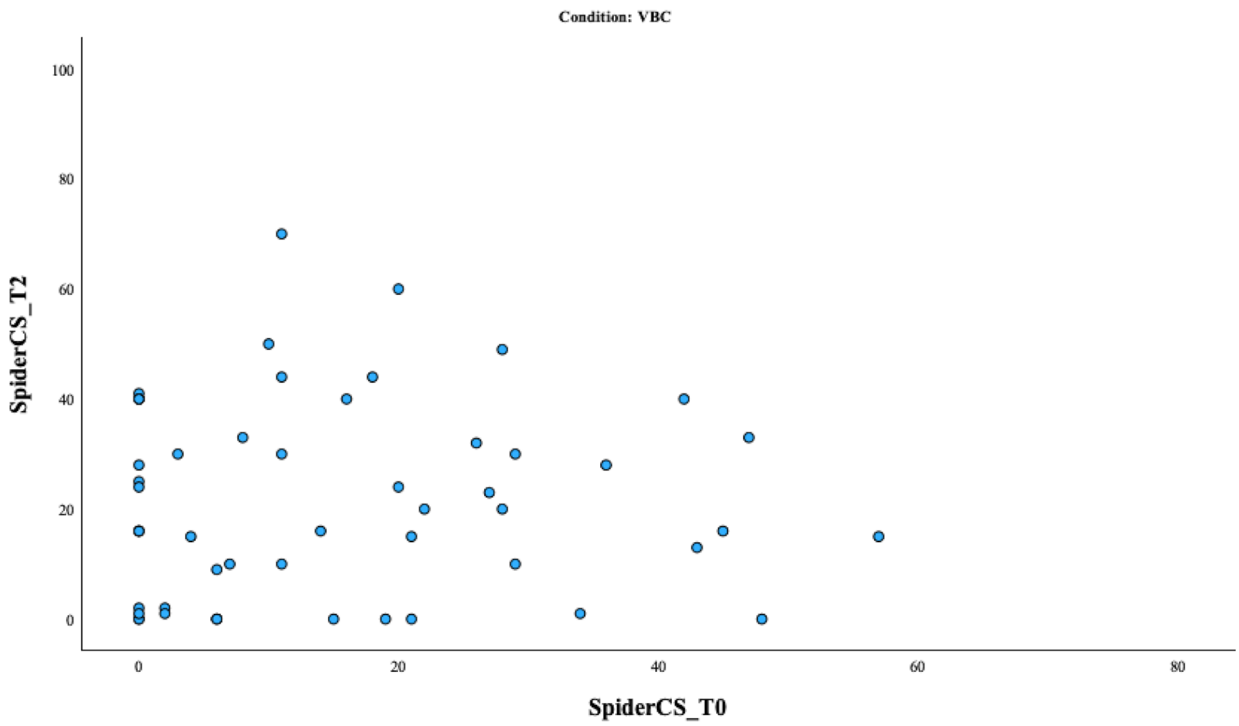


Figure D3

Scatterplot of Valence Ratings of the CS Spider for VBC-Group



Appendix E

Figure E1

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for the CS Spider in the CTL Condition

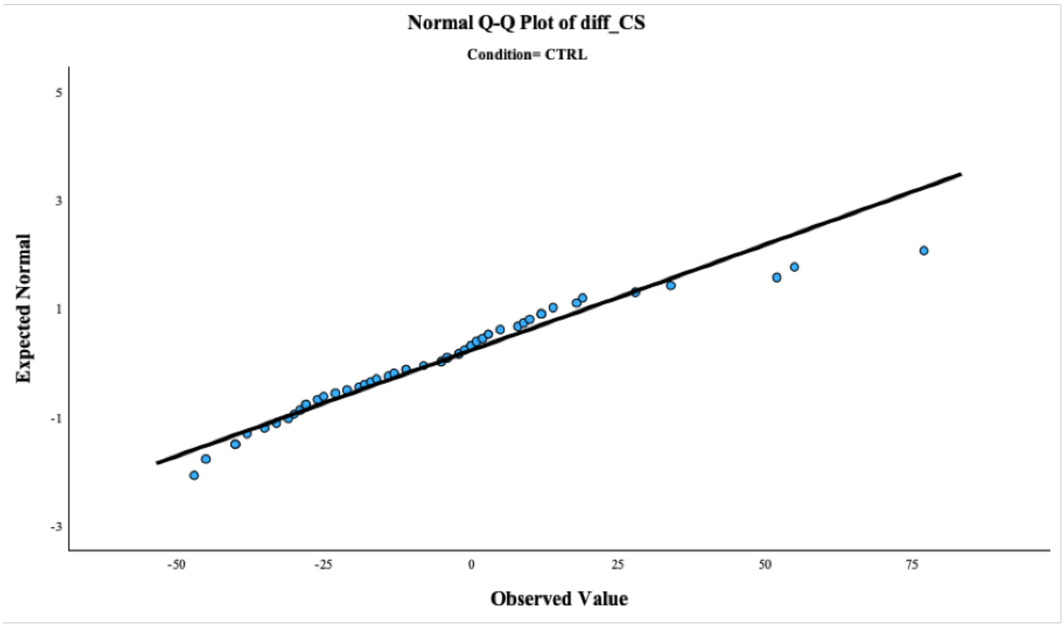


Figure E2

Boxplot of the Difference Scores (T2-T0) for the CS Spider in the CTL Condition

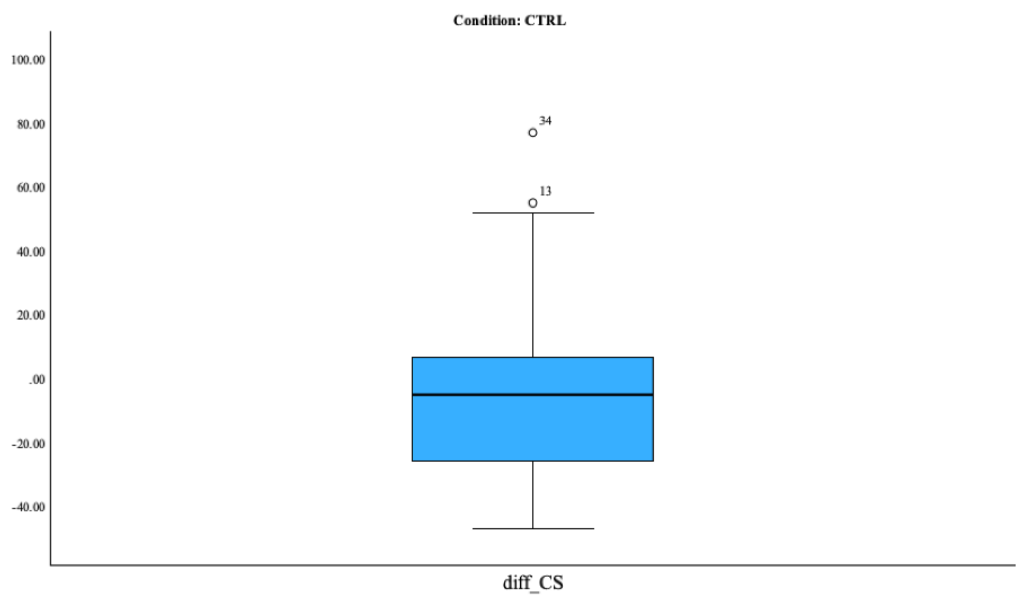


Figure E3

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for the CS spider in the VBC Condition

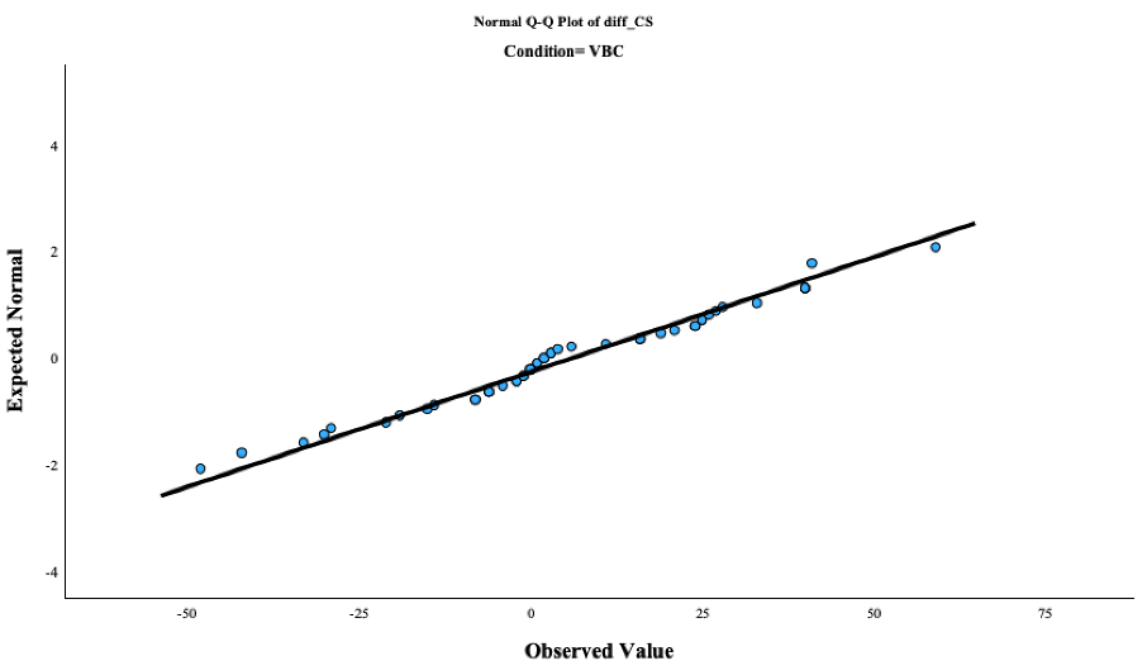


Figure E4

Boxplot of the Difference Scores (T2-T0) for the CS Spider in the VBC Condition

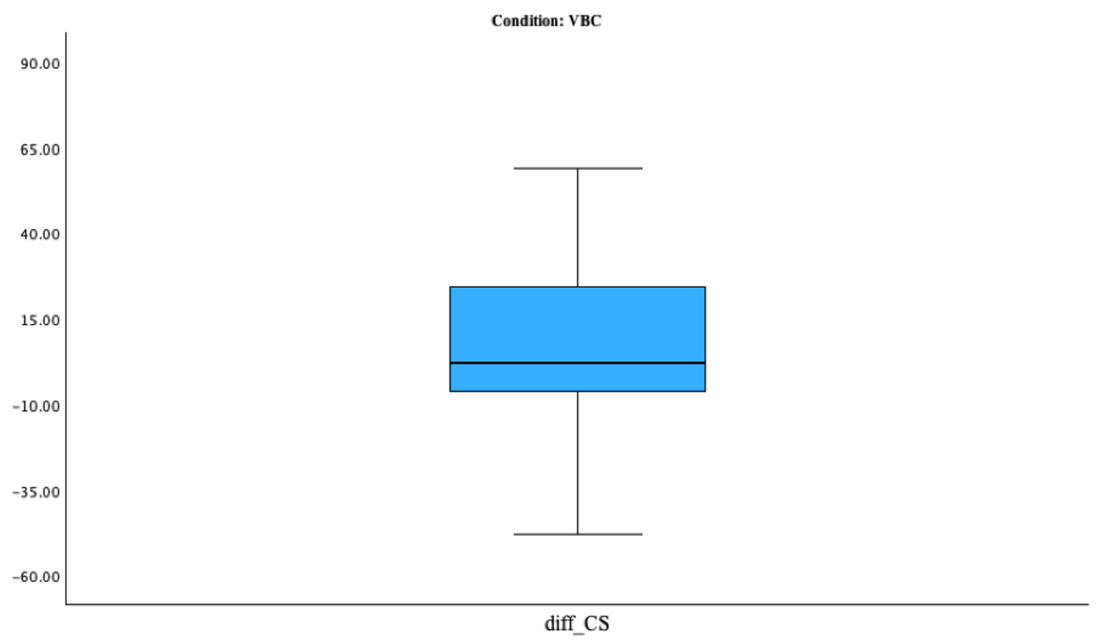


Figure E5

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for the CS Spider in the VBE Condition

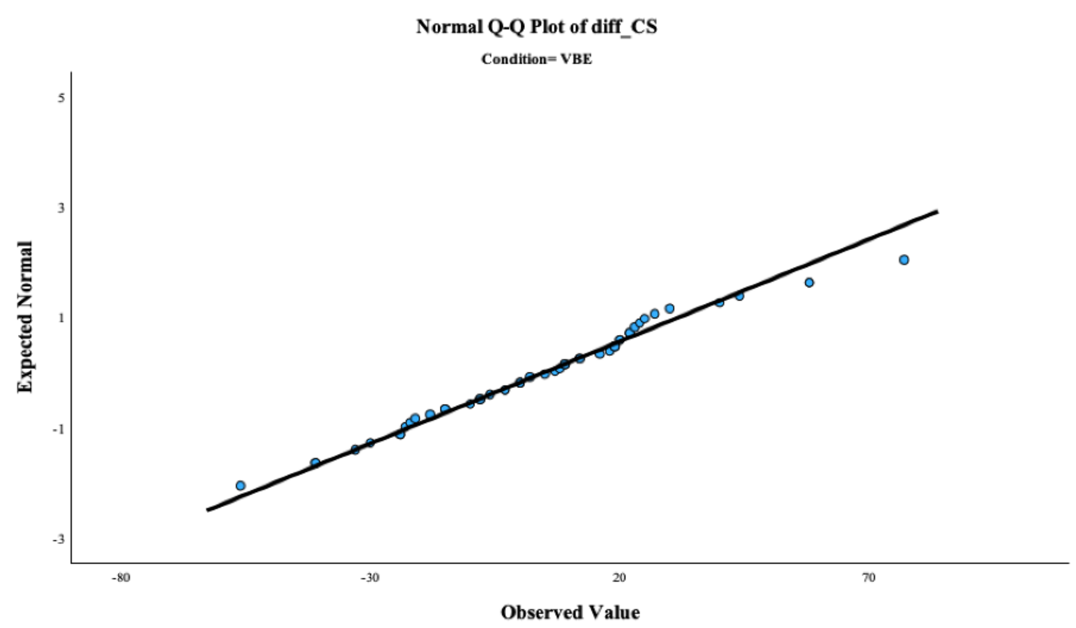
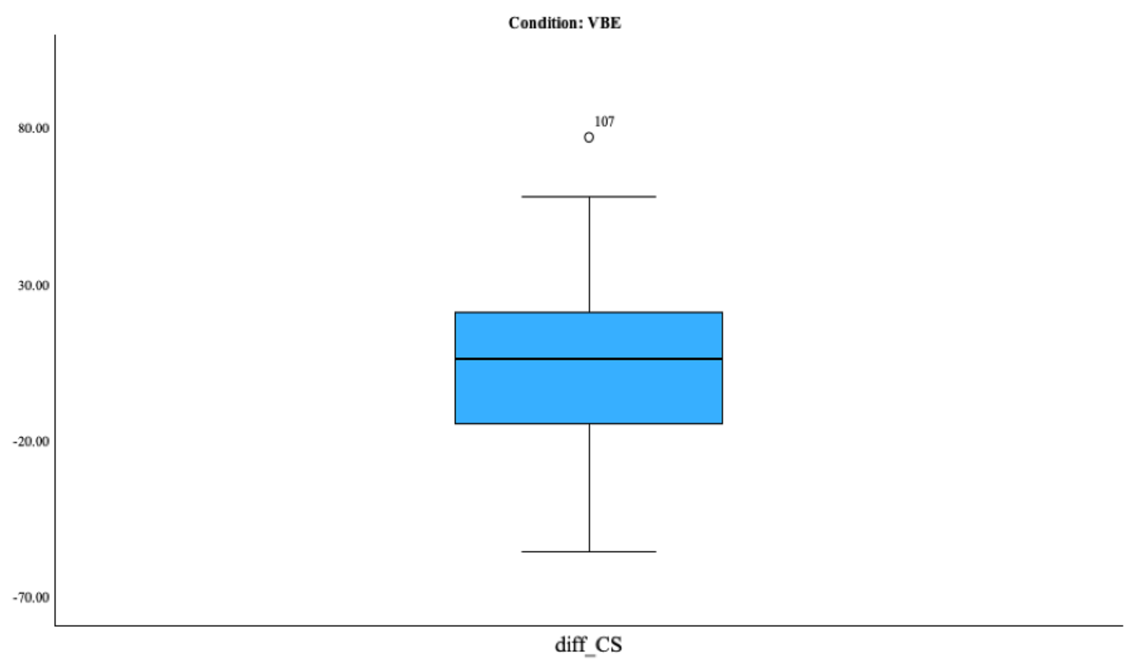


Figure E6

Boxplot of the Difference Scores (T2-T0) for the CS Spider in the VBE Condition



Appendix F

Figure F1

QQ-Plot of Spider 2 at T0 for CTL-Condition

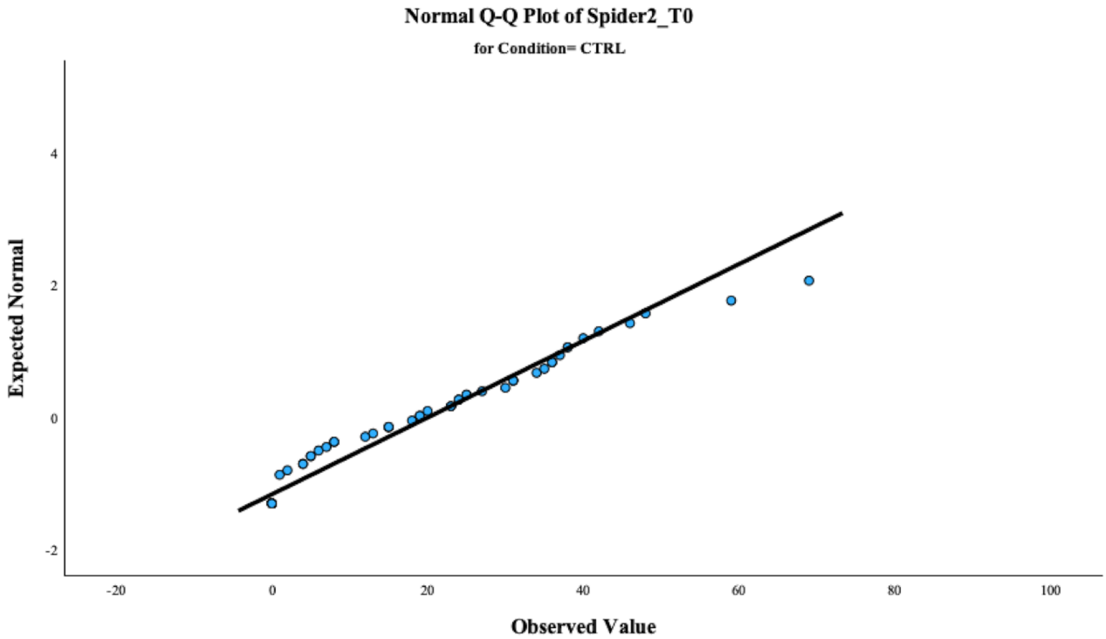


Figure F2

QQ-Plot of Spider 2 at T0 for VBE-Condition

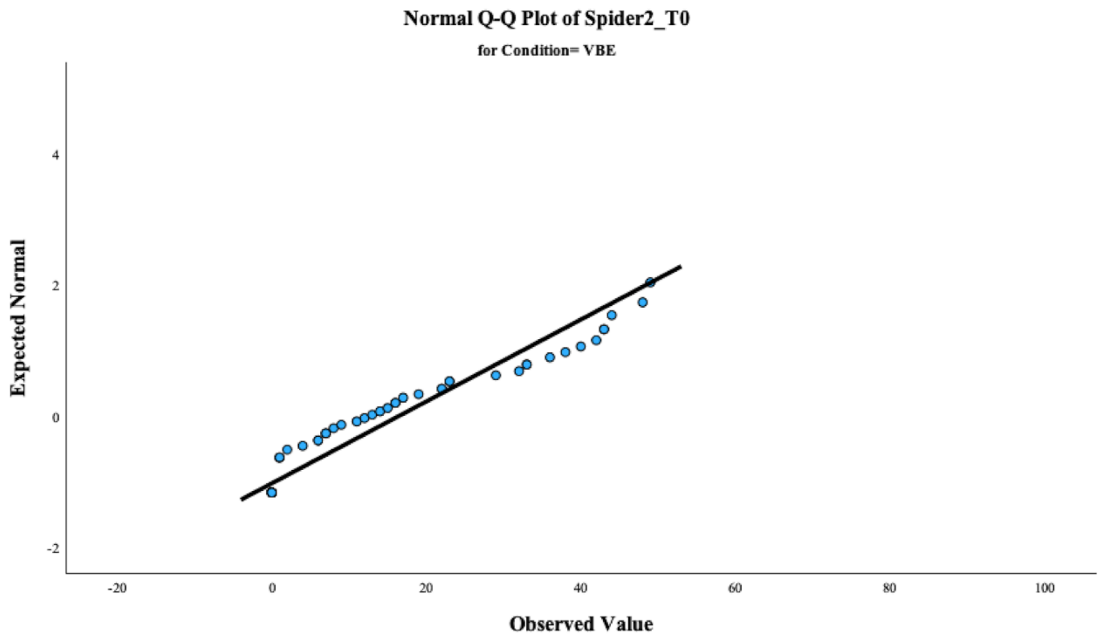


Figure F3

QQ-Plot of Spider 2 at T0 for VBC-Condition

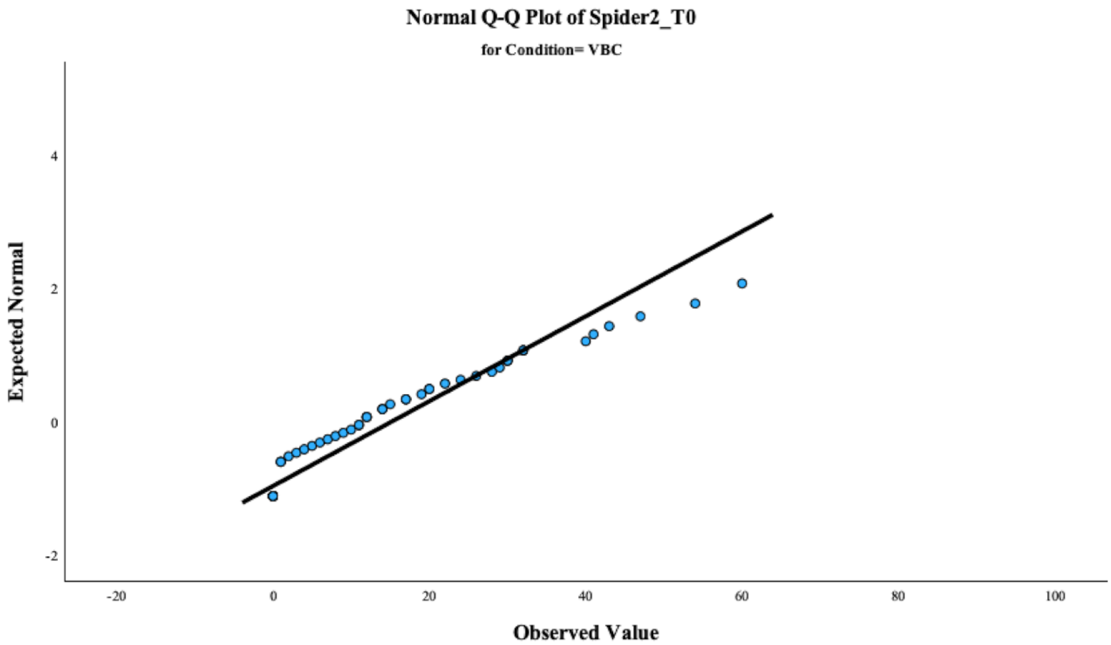


Figure F4

Boxplots of Spider 2 at T0 for each Condition

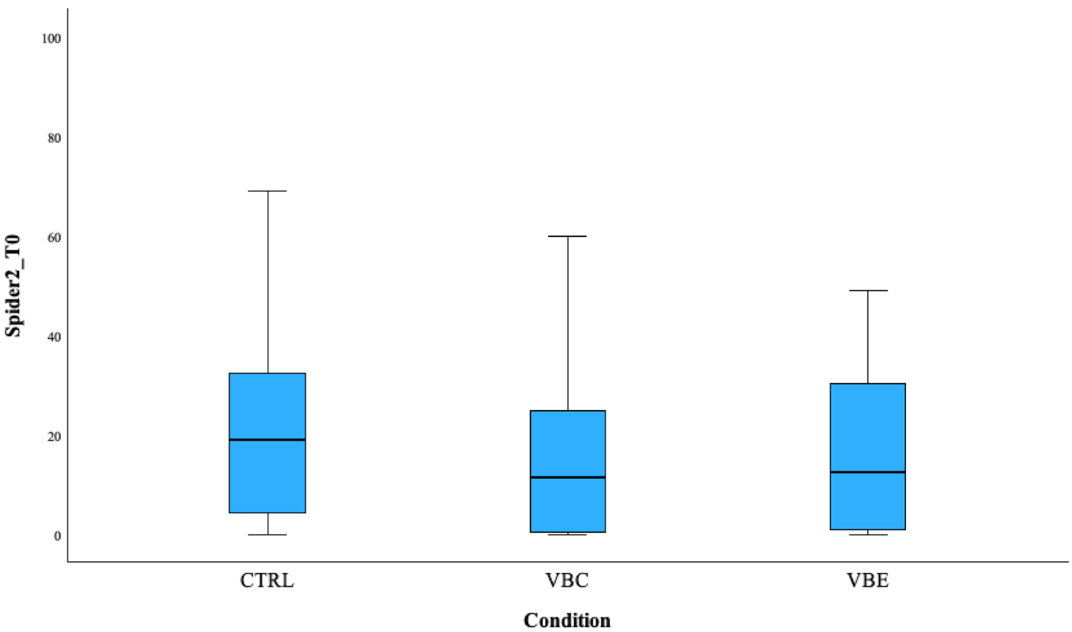


Figure F5

QQ-Plot of Spider 2 at T2 for CTL-Condition

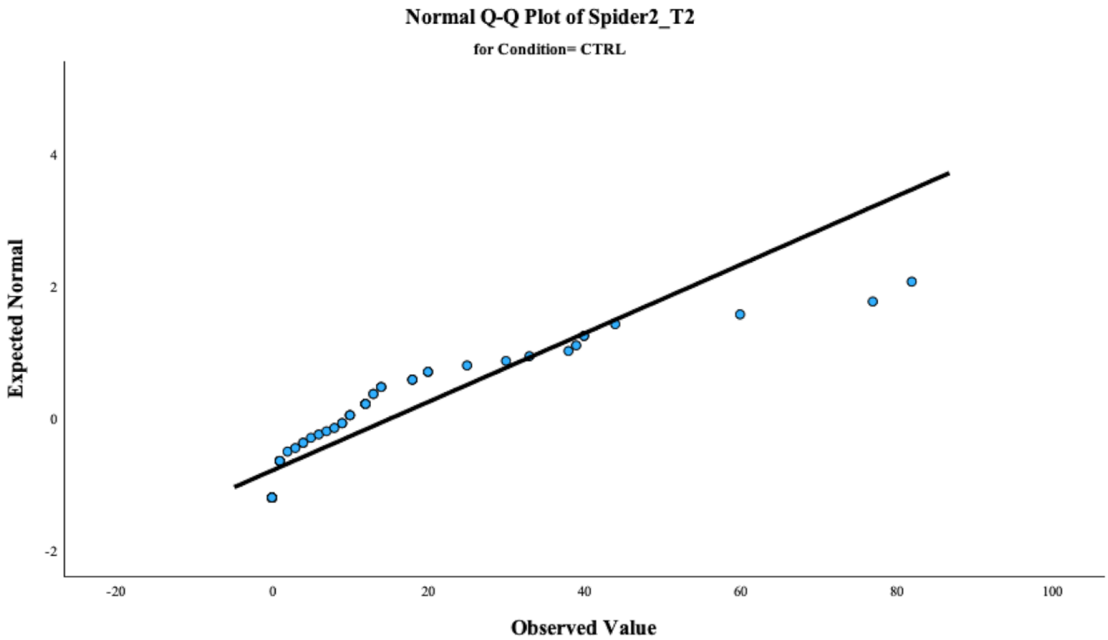


Figure F6

QQ-Plot of Spider 2 at T2 for VBE-Condition

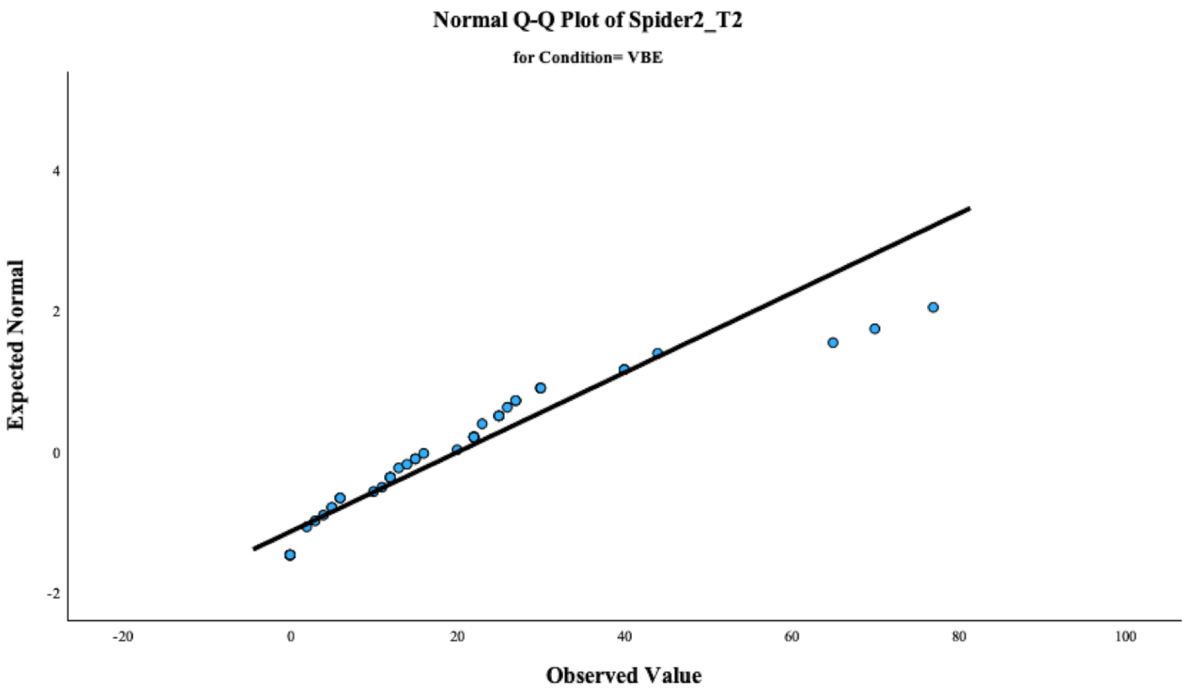


Figure F7

QQ-Plot of Spider 2 at T2 for VBC-Condition

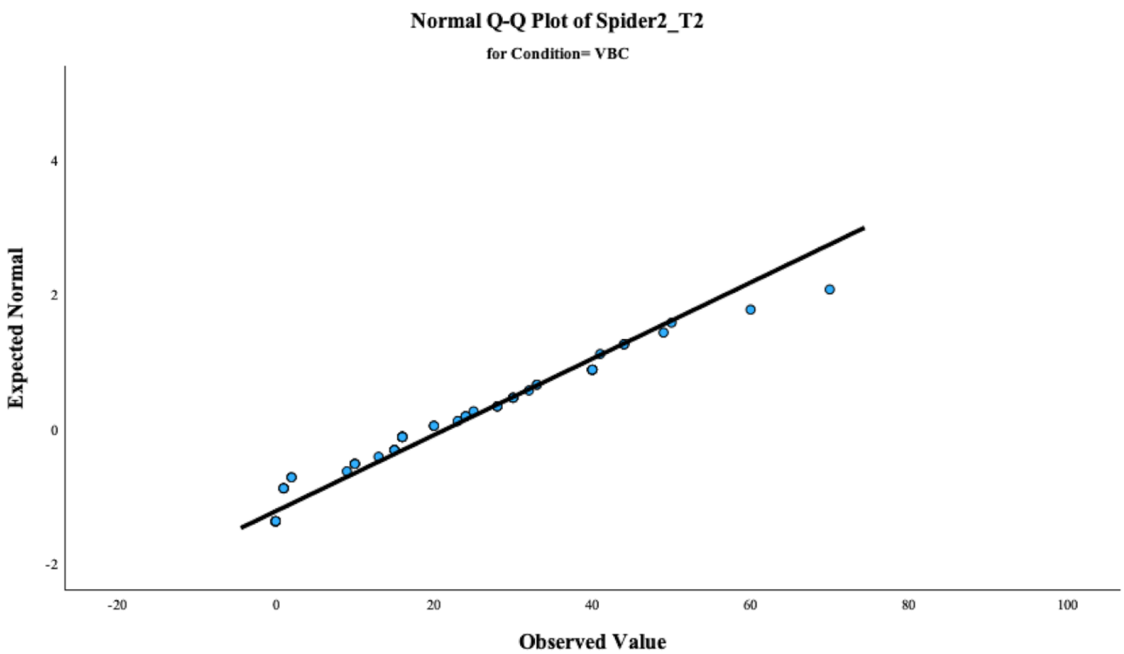
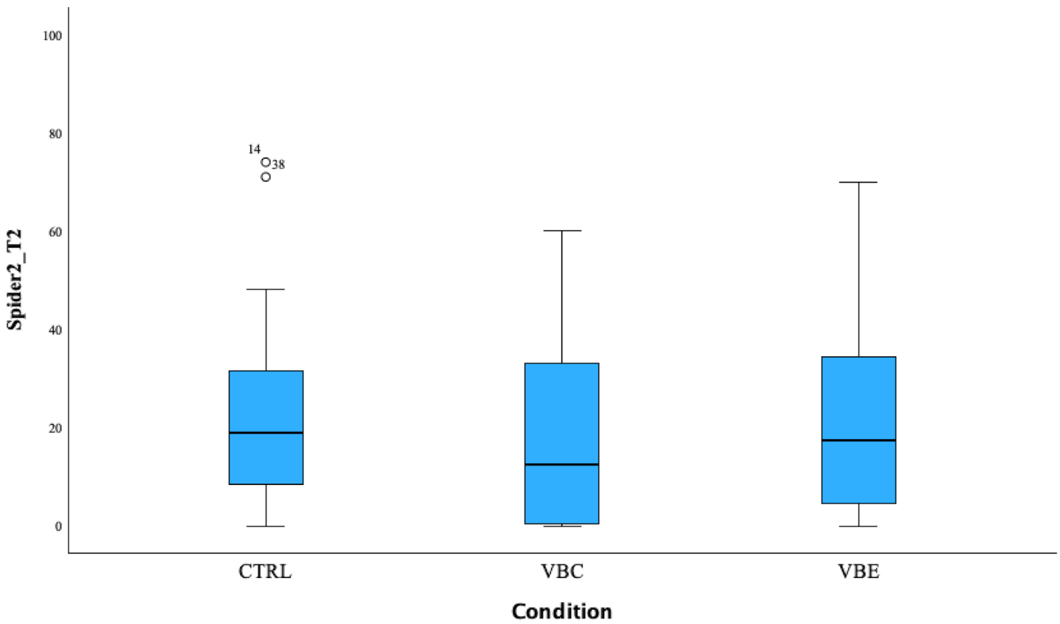


Figure F8

Boxplots of Spider 2 at T2 for each Condition



Appendix G

Figure G1

Scatterplot of Valence Ratings of the Spider 2 for CTL-Group

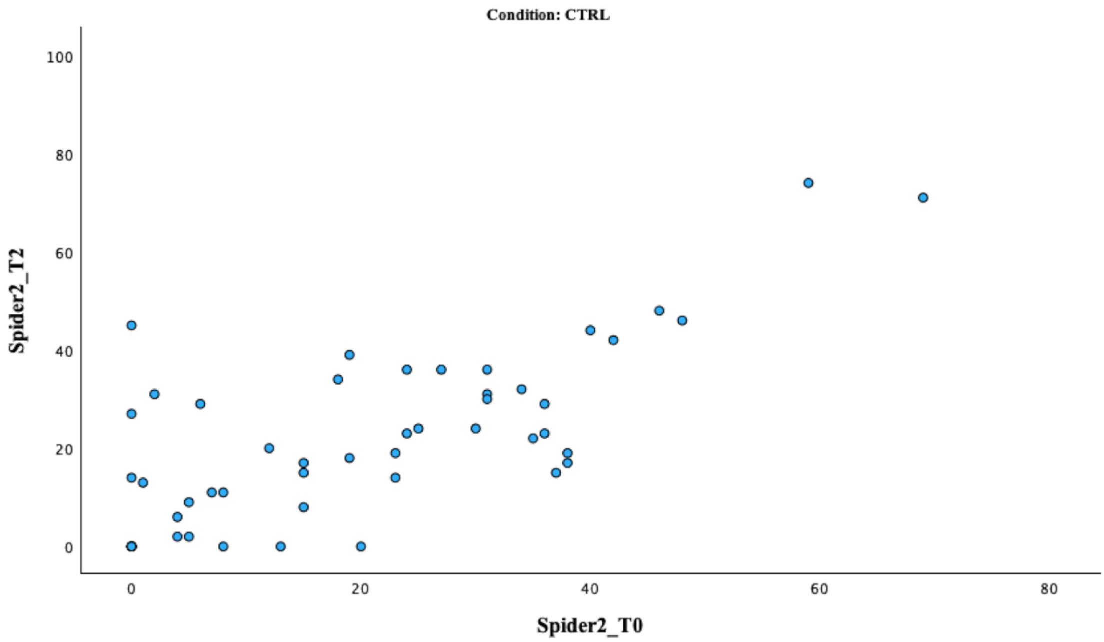


Figure G2

Scatterplot of Valence Ratings of the Spider 2 for VBE-Group

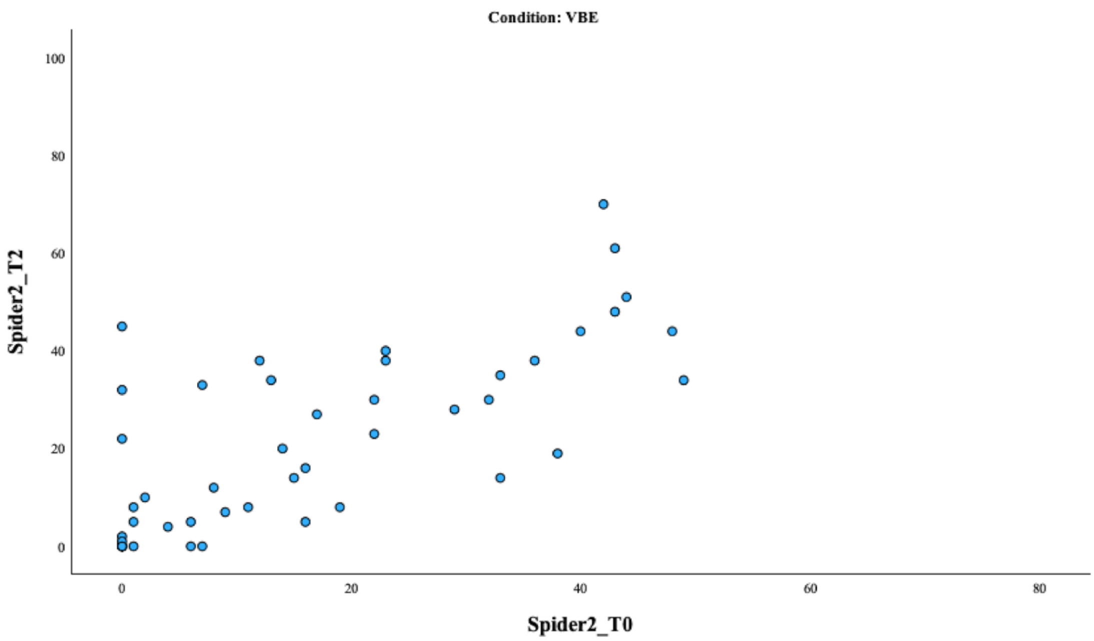
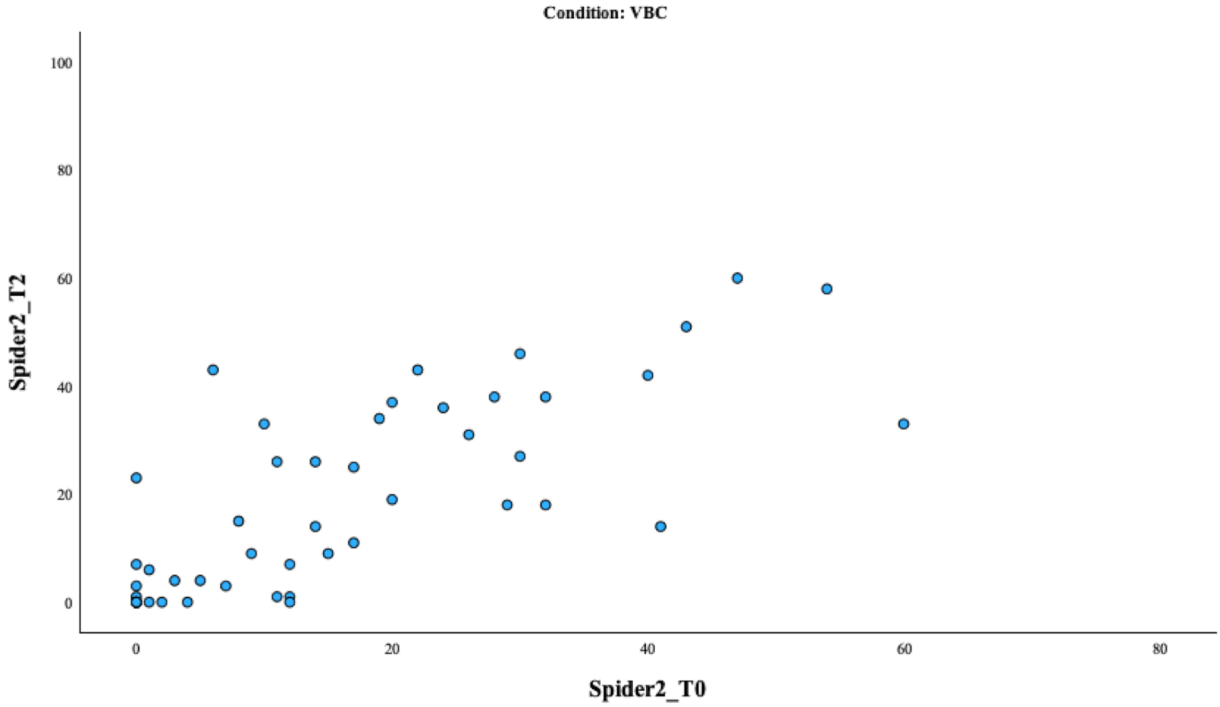


Figure G3

Scatterplot of Valence Ratings of the Spider 2 for VBC-Group



Appendix H

Figure H1

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for Spider 2 in the CTL Condition

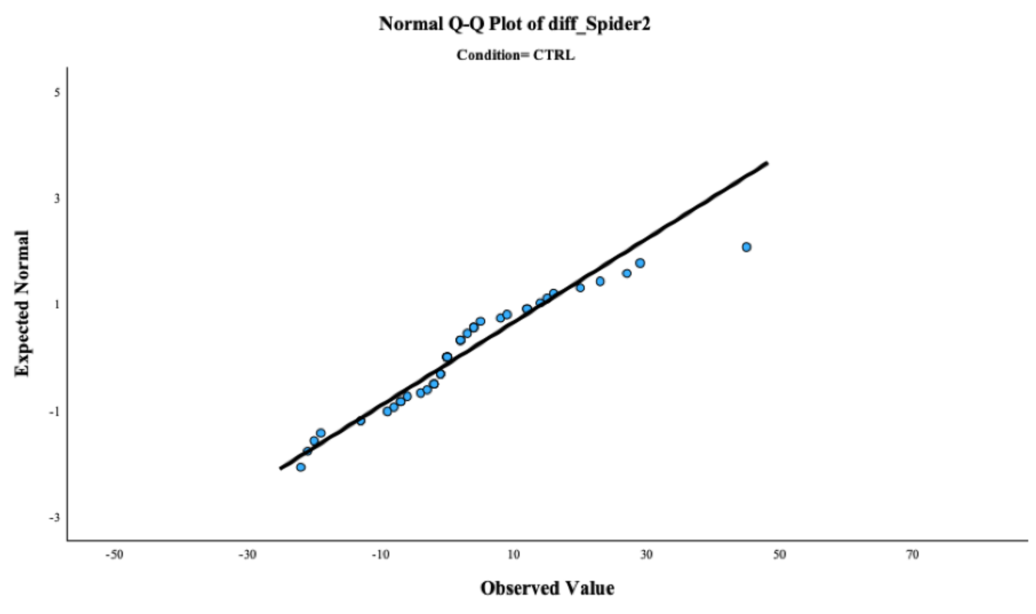


Figure H2

Boxplot of the Difference Scores (T2-T0) for Spider 2 in the CTL Condition

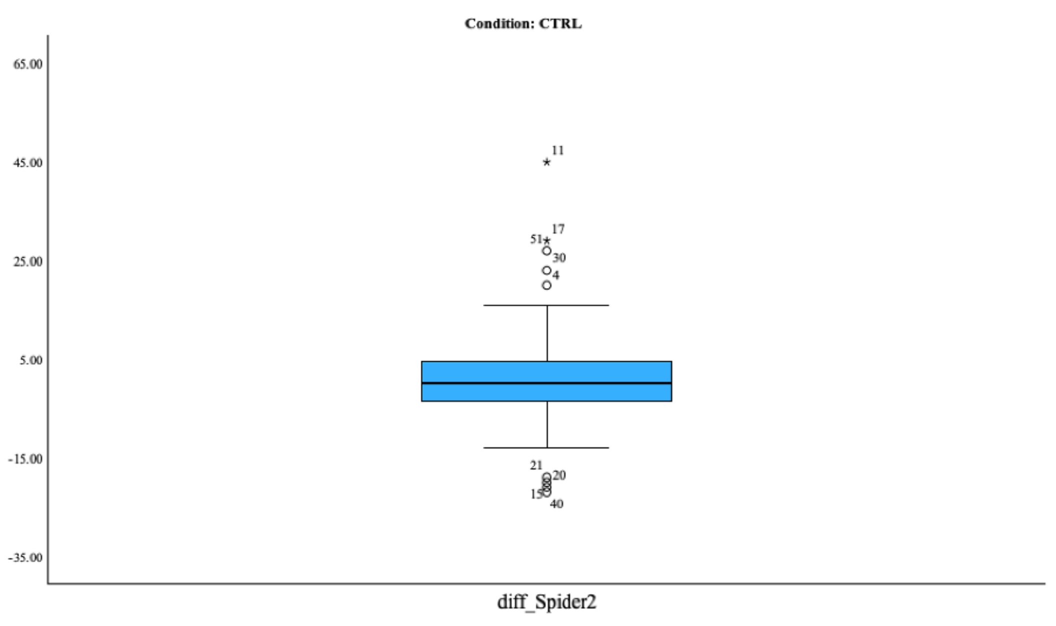


Figure H3

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for Spider 2 in the VBC Condition

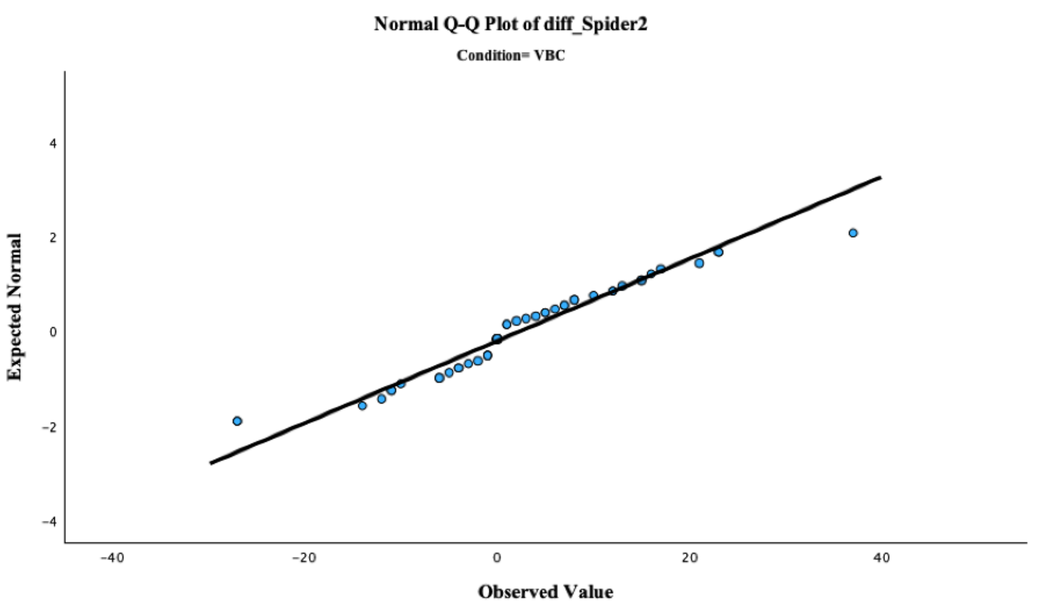


Figure H4

Boxplot of the Difference Scores (T2-T0) for Spider 2 in the VBC Condition

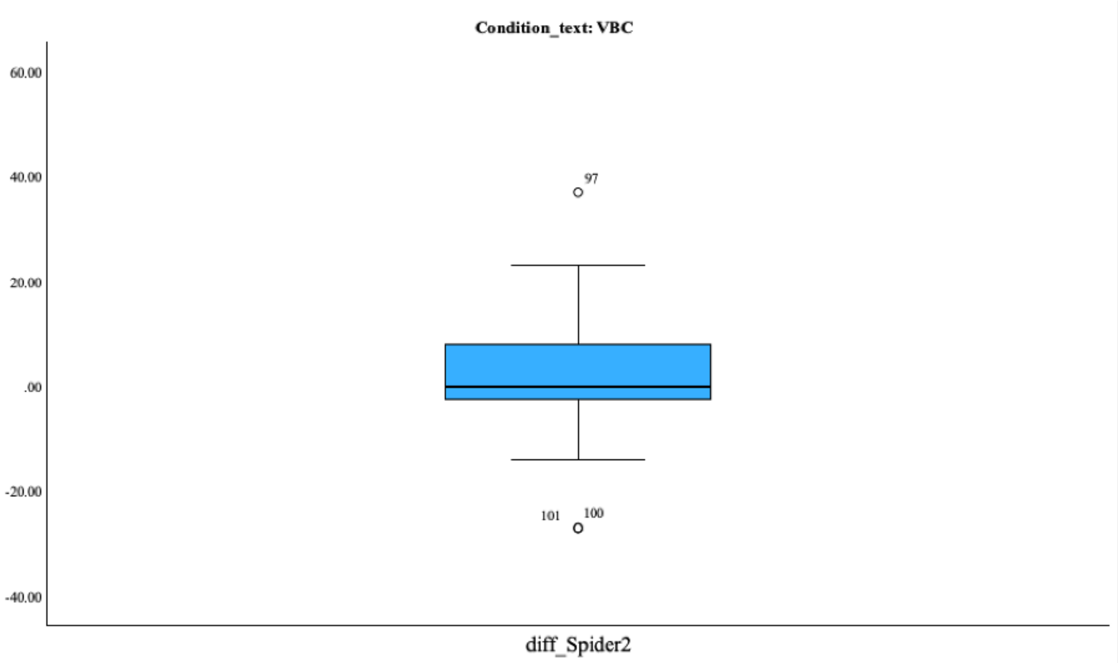


Figure H5

Q-Q Plot assessing the Normality of the Difference Scores (T2-T0) for Spider 2 in the VBE Condition

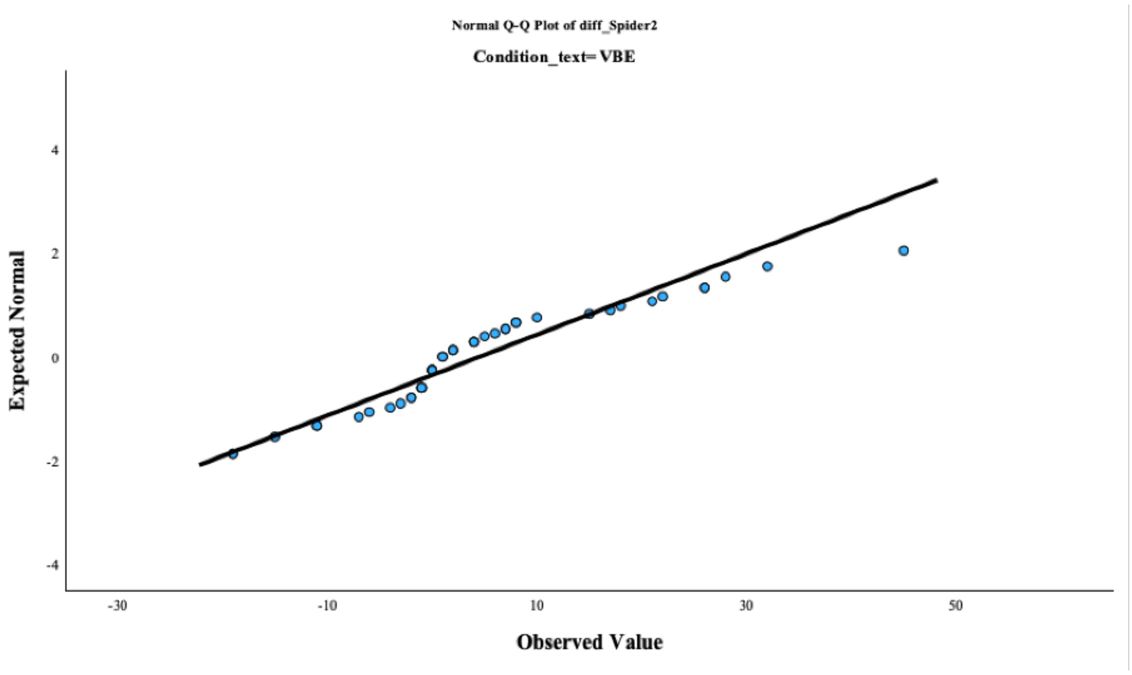


Figure H6

Boxplot of the Difference Scores (T2-T0) for Spider 2 in the VBE Condition