Replicating the Suppression-Induced Forgetting Effect in Online and In-Person Settings: Interactions with Emotion Regulation and Psychopathy

Hendrik Herholz

Master Thesis – Clinical Forensic Psychology & Victimology

S4014944
June 2024
Department of Psychology
University of Groningen
Examiner/Daily supervisor:
Ineke Wessel

Abstract

The ability to suppress unwanted memories, known as suppression-induced forgetting (SIF), is crucial for mental well-being and is often studied using the Think/No-Think (TNT) task. While prior research has consistently demonstrated less recall of No-Think words compared to Baseline (SIF effect) in laboratory settings, the first online implementation of the TNT task by Wiechert et al. (2023) yielded non-significant findings. Wiechert et al.'s (2023) in-lab replication reported statistically significant results, questioning the SIF effects' generalizability across different testing conditions. Additionally, past literature suggested a possible relationship of both cognitive control and psychopathy with memory suppression, but exact interactions remain unclear. This study aimed to replicate and extend previous findings by comparing the SIF effect in an online and laboratory context, and investigate its interactions with emotion regulation, as a cognitive control facet, and psychopathy. We implemented Wiechert et al.'s (2023) TNT procedure with n = 108 undergraduate psychology students, evenly split between the online or laboratory condition. Hypotheses predicted a SIF effect in both testing conditions, a positive correlation of the SIF effect with emotion regulation, and a negative correlation with psychopathy. Results of the same probe (SP) test indicated a significant SIF effect only in the laboratory condition, and no significant correlation of the SIF effect with either emotion regulation or psychopathy. These findings question the generalizability of the SIF effect, and we recommend both replication and further examination of our findings to ensure the TNT tasks' external validity and clarify its relationship with cognitive control and psychopathy.

Keywords: think/no-think; suppression-induced forgetting; cognitive control; online testing; psychopathy

Replicating the Suppression-Induced Forgetting Effect in Online and In-Person Settings: Interactions with Emotion Regulation and Psychopathy

During the course of our lives, we collect both positive and negative memories. While positive experiences are gladly remembered, certain negative experiences might understandably not be as desired. In order to reduce the occurrence of negative memories, one might try suppressing them to reduce the chance of future recall. In this context, suppressing is likely important for many people, as it may allow for higher general well-being. The ability to suppress is also relevant in a general mental health context (Costanzi et al., 2012), as intrusive memories are related to multiple mental disorders (Brewin et al., 2010; Harvey et al. 2004). Furthermore, having reduced control over one's memories also lies at the heart of certain disorders such as PTSD (McTeague et al., 2016; Brewin, 2011; Catarino et al., 2015), making memory suppression all the more relevant in a broader psychological setting. In general, there seem to be large individual differences in suppression ability, likely caused by a multitude of mental processes such as executive functioning (Levy & Anderson, 2008), and similarly, cognitive control capacity[1] (Chen et al., 2022). Furthering our understanding of which factors can influence the effectiveness of memory suppression will help improve mental well-being in both clinical, as well as more general contexts (e.g., Hertel & Calcaterra, 2005).

Emotion regulation and memory suppression

When considering that cognitive control capacity has an influence on memory suppression (Chen et al., 2022), a resulting question would be whether certain aspects of cognitive control have different effects. Cognitive control includes abilities such as attentional

[1] Note: The terms "executive functioning" and "cognitive control" are often used interchangeably in the literature. In this case, we use Mackie et al.'s (2013) definition of cognitive control, and relatedly cognitive control capacity, as "the broader construct of information prioritization for goal-driven decision-making [, compared to executive control] as a specific component of attention for conflict processing" (p.301).

control, working memory, and emotion regulation, among others (Mackie et al., 2013). Of these aspects, emotion regulation lends itself as an easily testable facet of cognitive control (Pruessner, 2020), as it is usually assessed by a short self-report measure (Gross & John, 2003). For clarification, emotion regulation can be defined as the "conscious or unconscious processes of monitoring, […] and managing emotional experiences" (Kok, 2017, p.1), including control over memory. In fact, memory suppression is a fundamental mechanism of emotion regulation (Richards & Gross, 2000), and while memory suppression tests generally use neutral stimuli, emotion regulation is also positively related to learning enhancement (Martin & Ochsner, 2017). Therefore, individuals with greater emotional control could potentially learn and perform better in memory suppression tasks, possibly achieving higher scores. These ideas are however speculative, and there is currently little research investigating this specific relationship. A further exploration into if and how emotion regulation correlates with memory suppression could subsequently help in developing more effective emotion regulation strategies, as well as specify its relationship with cognitive control.

A link to psychopathy

Interestingly, emotion regulation is also related to a very relevant aspect in the forensic setting, namely psychopathy (Garofalo et al., 2020). Psychopathy is one of the most researched clinical constructs in forensic psychology (Hare & Neumann, 2008; Miller et al., 2008), made up of a multitude of subtypes and possible personality traits (Murphy & Vess, 2003; Falkenbach & Zappale, 2021). The general interest in this research is not at the least because psychopathy serves as a predictive item for a multitude of criminal behaviors (DeLisi, 2009). Psychopathic traits include impulsivity, skillful manipulation, lack of remorse, and pathological lying, among others (Hare et al., 1989). Due to its complexity, many different aspects of psychopathy that could be relevant to memory suppression have already been explored. This includes the relationship with memory fabrication (Battista et al.,

2023), emotional memory (Ragbeer & Burnette, 2012), as well as a reduced susceptibility for memory intrusions (Moul & Caroline, 2017). Despite these findings, no studies have yet tested if psychopathic traits are related to actual memory suppression ability, and research into this area would provide a novel insight into how psychopathic traits may relate to memory.

Our main prediction for such an investigative study comes from Gillespie et al., (2022), who found that psychopathy was modestly associated with response inhibition problems in go/no-go tasks. They further note that there is a lack of research specifically examining the relationship between psychopathic traits and response inhibition, despite evidence suggesting a general impairment in cognitive tasks in psychopathic individuals (see Morgan & Lilienfeld, 2000; Ogilvie et al., 2011). Considering that inhibitory processes might be related to memory suppression (Anderson et al., 2004), psychopaths might therefore struggle with suppressing their responses. However, a counterargument could be that a reduced susceptibility for intrusive memories might improve memory suppression, as there are fewer mental distractions. While a literature review did not reveal exact support for this argument, a study by Castiglione and Aron (2020) offers a nuanced perspective, using the idea that memory retrieval may utilize a similar mechanism to motor suppression. They found that participants reporting fewer intrusions in a TNT task also showed a rapid motor suppression. This could either suggest a more efficient stopping mechanism for unwanted memories, or a less focused state that potentially leads to more mistakes. It also stands to reason if these assumptions can be generalized to people with psychopathic traits, yet no specific studies investigating this idea were found. Given the multifaceted nature of psychopathy, and the complex relationship between intrusive memories and suppression, further research is needed to directly assess memory suppression ability in individuals with psychopathic traits, as well as clarify underlying mechanisms.

The TNT task

Memory suppression ability is generally tested through the so-called Think/No-Think (TNT) task, introduced by Anderson and Green (2001) as an extension of the go/no-go task for use in memory retrieval. In the TNT task, participants were first shown a certain number of word pairs (e.g., WAFFLE – MAPLE), wherein left-hand words served as the cue and right-hand words as the response in the later tasks. Once participants had successfully completed a certain minimum retention threshold (originally 50% in Anderson & Green, 2001), the main TNT task began. Individual cue words were instructed to be either part of think or no-think trials, shown in green or red respectively. For the green think cues, participants should respond with the correct response word out loud as fast as possible. For the red no-think cues, response words were supposed to be thought of or recalled. Some word pairs were left out of this phase and served as baseline items. Subsequently, all cue words, including baseline items, were presented. Participants were instructed to now respond to all cue words, regardless of any previous association during the think or no-think phase. Anderson & Green (2001) found that think items, when compared to baseline items, were recalled more frequently. More importantly, however, no-think items were recalled less often when compared to baseline items. This suggested a suppression-induced forgetting (SIF) effect with no-think items, wherein the "impairment [also] increased linearly with suppression practice" (Anderson & Green, 2001). The TNT task has generally provided significant SIF results throughout a multitude of testing variations (Anderson & Huddleston, 2012; Clark, 2021; Stramaccia et al., 2021).

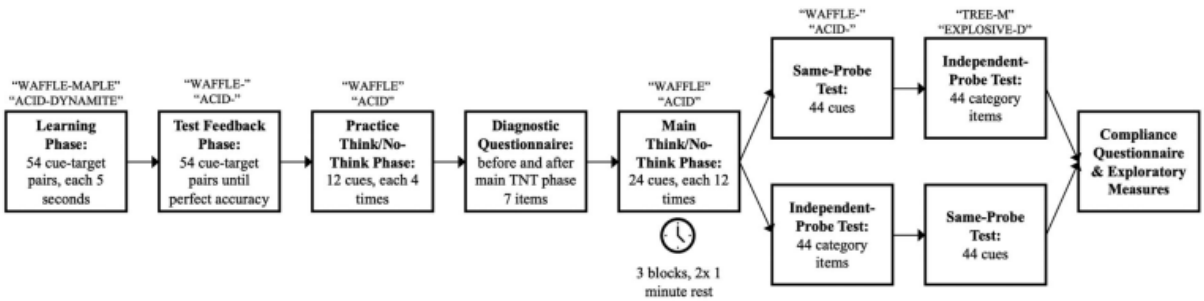
IP and SP testing

Importantly, there are two different types of tests within the TNT task, namely the same probe (SP) and independent probe (IP) test (see Figure 1 for a full procedural overview from Wiechert et al., 2023). The SP test presents the same original cues (e.g., "rug") and asks participants to recall the associated response (i.e., "wool"). On the other hand, participants in

the IP test see words that are semantically related to the expected response word, as well as the first letter of the actual response word (e.g., “fabric - w”). Anderson & Green (2001) originally introduced these two types to differentiate between inhibitory and non-inhibitory suppression effects. In a SIF task, two potential mechanisms could explain the SP test (SIF-SP). First, a repeated presentation of the cue (e.g., “rug”) without the target word (i.e., “wool”) during the suppression phase may simply weaken the word pair association in memory. Second, the participant actively suppresses retrieval of the target word and subsequently cannot recall the correct answer, even if the connection to the cue word is still somewhat associated with it in memory. The IP test (SIF-IP) solves this problem by removing the initial cue word and therefore only focuses on a possible reduction of memory traces. While Anderson & Green (2001) did not find a notable difference between both tests, more recent studies skew towards SIF-SP being more effective (Anderson & Huddleston, 2012; Stramaccia et al., 2021). However, as Wessel et al. (2020) note, IP tests are also not as common as SP tests, possibly due to sample sizes being based on SP effect sizes, which results in insufficient statistical power, null findings, and therefore fewer publications of IP tests.

Figure 1

Overview of the TNT procedure flow



Note. Reprinted with permission from “Suppression-induced forgetting: A pre-registered replication of the think/ no-think paradigm.”, by Wiechert et al., 2023, *Memory*, 1-14.

Grounds for replication

In general, the TNT literature provides few null findings (e.g., Bulevich et al., 2009; Wessel et al., 2020; Wiechert et al., 2023), and replicating these studies is relevant to assessing the effectiveness of the SIF phenomenon. A study by Wiechert et al. (2023), while well-powered (n = 150), reported non-significant results of both SP and IP tests in the first online TNT testing environment with international participants. In general, results regarding some cognitive tasks (e.g., reaction time, Stroop tasks, priming tasks) have indeed been shown to be comparable when done online and in-lab (Semmelmann & Weigelt, 2017; Sauter et al., 2022). Following Wiechert et al.'s study, Wessel et al. (2023) set to replicate their results with undergraduate psychology students. To find out if the SIF effect might be set-up dependent, an in-person laboratory session was chosen. As null findings of SP testing are generally rare, and because of resource constraints, they focused on only the SP test in their replication. Their analysis suggested a strong SIF-SP effect, failing to replicate Wiechert et al.'s (2023) results. Wessel et al. (2023) acknowledged the exclusion of an IP test as a minor change but framed the testing environments and sample characteristics as the main reason for the differing results.

Research goals and Hypotheses

Given the differing findings regarding the SIF effect in online and in-person settings (Wiechert et al., 2023; Wessel et al., 2023), our study closely replicated Wessel et al.'s (2023) experiment, using undergraduate psychology students evenly divided between the online and in-person laboratory condition. Additionally, the relationship between the SIF effect and emotion regulation, as well as psychopathy will be explored. As discussed earlier, emotion regulation is a component of cognitive control and may influence memory suppression (Chen et al., 2022; Richards & Gross, 2000). Furthermore, despite the potential relationship of psychopathy and memory suppression, no studies have directly investigated this interaction.

Building on the research suggesting a link between psychopathy and impaired response inhibition (Gillespie et al., 2022), psychopathic traits and their correlation with the SIF effect will be examined. To our knowledge, this is the first time psychopathy and emotion regulation will be studied specifically with the TNT task. To replicate Wessel et al. (2023) closely, and because of similar resource constraints, it was decided to exclude the IP test from our study. We hypothesize that in our SP test, (1) both in-person and online settings will result in a SIF effect, (2) lower scores on emotional regulation correlate with a reduced SIF effect, and (3) higher scores of psychopathy correlate with a reduced SIF effect.

Method

Transparency statement

The hypotheses, method, and data analysis plan were preregistered, and are available on the Open Science Framework (OSF). Two deviations from the pre-registration occurred. First, the preregistered inclusion criteria were initially set to exclude participants with any past mental disorders. However, this criterion was revised to include participants that were no longer diagnosed with a mental disorder. Second is the additional analysis of the Counterbalancing Conditions (A, B, C) with no-think and baseline scores, using a repeated measures ANOVA.

Sample

Eligible participants were 108 (74 female, 2 non-binary) first-year psychology bachelor students at the University of Groningen, from both the English and Dutch language tracks. Students received course credits for their participation. The mean age was 20.27 (SD = 3.01, n = 108, range 18-36), with 36 Dutch, 20 German, 14 English, and 38 other native language speakers. Taken from Wessel et al. (2023), the inclusion criteria for both online and in-lab experiments were: ages between 18 and 45, normal or corrected to normal eyesight, no colorblindness, no dyslexia, no long-term health conditions or disabilities, and no current

mental illnesses. For the online version of the experiment, stable internet, a working webcam, a working laptop/computer, and a quiet place without distractions were also required.

Originally, 130 participants signed up for the experiment. Four out of these 130 (3.07%) did not complete the experiment, because of technical problems or mental illness exclusions. As follows, 126 participants completed the experimental session. Post-experiment exclusions included 1 case of missing data (0.79% of $n = 126$), 2 insufficient environmental settings (1.58% of $n = 126$), and 3 participants (2.38% of $n = 126$) taking longer than 25 minutes to correctly recall all words once. For these cases, continuing with the experiment was done as study credits were involved. Furthermore, 6 exclusions occurred due to scores higher than 4 in the compliance questionnaire (4.76% of $n = 126$). Lastly, 6 (4.76% of $n = 126$) participants with possibly influencing mental illnesses still completed the experiment but were excluded afterward[2]. The majority of exclusions occurred in the online setting (78.95% of $n = 19$). No other exclusions took place (i.e., reacting to at least 75% of targets in the TNT main phase, and no indications of "very much" regarding the distraction questions in the session evaluation questionnaire). Exclusions were compensated for by new participants until $n = 108$ total completions were reached for the analysis.

The experiment was approved by the Ethics Committee of the University of Groningen's psychology department (Research code PSY-2324-S-0182).

Design and power analysis

A within-subject design with 3 instruction conditions including the levels of Think, No-Think, and Baseline was used. The dependent variable was the percentage of correctly recalled items (SIF score), measured by the recall difference between No-Think and Baseline items. Given the directional hypothesis that there would be a decreased recall of no-think

[2] Some experimenters erroneously continued with the testing of participants with a relevant mental illness. As noted, these participants were still excluded afterwards.

items compared to baseline items, and in line with the replication of Wessel et al. (2023), the same a priori G*power analysis (Faul et al., 2009) with a one-sided $\alpha = .025$ was used. This revealed that $n = 54$ would be required for detecting a medium effect size (Cohen's $d = 0.5$) with a power of .95. Considering an online and in-person setting, this resulted in a total of 108 participants.

Materials

Materials were adopted from the Wiechert et al. (2023; see https://osf.io/e75a6/) and Wessel et al. (2023; see https://osf.io/gsa58/) studies, excluding relevant exploratory measures. For the TNT task, which includes the Diagnostic Questionnaire, Inquisit lab (version 6.6.1) was utilized. The Qualtrics platform (November 2023) was used for all other questionnaires.

Experimental control questionnaire

Before the TNT task started in an online session, experimenters used the experimental control questionnaire to assess whether the technical setup was functional, and the session environment was adequate to commence the session. These 11 questions included set-up aspects such as "Has the participant switched off their phone?", and distraction checks like "Is it busy in the background of the participant?".

Think/No-Think task

Like Wessel et al. (2023), only Same Probe (SP) recall was assessed in the TNT task. The TNT task includes 54 cue-target pairs (e.g., WAFFLE – MAPLE), subdivided into three 12-word pairs and 18 filler pairs. These three 12-word pairs resulted in three versions (A, B, C) that counterbalanced the word pairs across the Think, No-Think, and Baseline conditions.

Diagnostic Questionnaire (see Anderson et al., 2004)

During the computer task, instruction comprehension and adherence were tested by seven items, such as "When you looked at the RED hint word, how often did you read and understand it?".

Compliance Questionnaire (see Levy & Anderson, 2012)

Three questions with five-point answer scales (0 = *Never*, 4 = *Very frequently*) determined participants' instruction compliance. Combined scores of 5 and above (out of a maximum of 12) meant participant exclusion.

Strategy check

Two questions with five-point answer scales (1 = *Strongly agree*, 5 = *Strongly disagree*) determined instruction compliance specifically regarding the participant's continued attention towards the red words.

Session Evaluation Questionnaire (Zwaan et al., 2018)

Nine questions with three-point Likert scales (1 = *None at all*; 3 = *Very much*) assessed participants' feelings of the experiment. Selecting "Very much" on either of the two distraction questions (i.e., "There are a lot of distractions here", and "I was distracted during the experiment") meant participant exclusion.

Emotion Regulation Questionnaire

In order to assess Emotion Regulation ability, the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003) was included as the first part of our exploratory purposes (see Appendix A for the full questionnaire). The ERQ is the most common self-report measure for testing emotion regulation strategies, including emotional experience as well as emotional expression, and contains 10 items measured by a 7-point Likert scale (1 = *Strongly disagree*; 7 = *Strongly agree*).

Levenson Self-Report Psychopathy Scale

To estimate our second exploratory measure of possible Psychopathy, the Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995) was utilized, which measures psychopathic traits in non-institutionalized populations and encompasses primary and secondary psychopathy (see Appendix B for the full questionnaire). The 26-item self-report inventory measures likely psychopathic attributes such as deception tendency and lack of remorse with a 5-point Likert scale (1 = *Disagree*; 5 = *Agree*). Both the LSRP and the aforementioned ERQ were chosen as they were the most easily accessible and did not require purchase.

IES-R, THQ, and RES Questionnaires

As this thesis project included other students, 3 other questionnaires were also included. Tested aspects were traumatic events and resilience, assessed by the Impact of Events Scale Revised (IES-R; Weiss & Marmar, 1997), Trauma History Questionnaire (THQ; Hooper et al. 2011), and Resilience Evaluation Scale (RES; van der Meer et al., 2018), respectively. This is included here to clarify all materials used in the experimental session.

Demographics

Participants were asked about their gender identity, utilizing male, female, non-binary/ third gender, and non-disclosure answer options. Age was assessed by a write-in option. Answer options for the participant's native language included English, Dutch, German, and a write-in option.

Procedure

The study was conducted in English. The experiment sessions were conducted by four students of the MSc. Program in Clinical Forensic Psychology & Victimology program, and one second-year Honors student of the BSc. Psychology program. All experimenters scored the equivalent of at least a C1 proficiency on either a recent TOEFL test, or an online LexTale English test (Lemhöfer & Broersma, 2021). Training was conducted across multiple meetings

by two senior experimenters, namely the PI in Wiechert et al. (2023), and an experimenter in both Wessel et al. (2023) and Wiechert et al. (2023). Every new experimenter passed training, requiring 100% correct recall of all word pairs, as well as completing a mock experiment session. Furthermore, the online testing sessions were evaluated by a senior experimenter about halfway through each experimenter's testing progression.

The local participant recruitment system (SONA) was used for participant sign-ups. Descriptions for both the online and in-lab experiments were identical, not mentioning memory but rather framing the study as attention-related. To avoid sign-ups based on preference, the last two digits of a participant's phone number (provided through a SONA prescreening) were used to decide which version they would be shown in SONA. Participants with ending phone numbers from 01 to 20, and 41 to 60 could only see and sign up for the laboratory condition. The remaining possible participants with ending phone numbers from 21 to 40, 61 to 80, and 81 to 00 could only see and sign up for the online condition. Participants were tested individually either online or in a laboratory room. At the beginning of the session, the experimenter obtained the informed consent, checked inclusion criteria, and explained the TNT task while in the same room as the participant. For the online experiment, participants downloaded the Research Information and Informed consent through Qualtrics and then consented within the same questionnaire. Subsequently, the TNT task was administered by the experimenter from the adjacent room with the door open. The experimenters read the same standardized script for every participant (see Wiechert et al., 2023; https://osf.io/456tk). In both the online and in-lab settings, the screen could be seen by both the experimenter and participant. Specifics of what could be seen are detailed in the individual phase descriptions below. The experimenter read the instructions on screen out loud, and the participant was instructed to read along. The participant responded verbally to the presented stimuli while the experimenter coded their responses. In the online condition,

the experimenter disabled their camera and muted themselves during the trials. They unmuted themselves only before reading instructions, or for important notices during trials (e.g., addressing a reoccurring noise). The participant was asked to leave their camera and microphone on for the entire duration of the experiment. After completion of the main TNT task in the laboratory, the participant filled in the post-experimental questionnaires on their own. The counterbalancing condition order was randomly assigned for each experimenter in advance by utilizing Google's AI chatbot (formerly Bard, now Gemini). Conditions were assigned by asking for a random order of the letters A, B, and C, 36 times in a row, resulting in 108 equally distributed conditions. Progress tracking was done in a shared Google Sheets document. The TNT task consisted of five phases detailed below, after which the Qualtrics questionnaire was administered.

Learning phase

Participants were instructed by the experimenter to learn word pairs for an upcoming recall. It was explained that when presented with the left-hand word, the right-hand word was the desired response word. Participants then saw all cue-target pairs being presented for 5 seconds one after another in a semi-random order. The first and last words were always baseline items. Unless stated otherwise, all word pairs (and singular words in the following phases) were presented in the middle of the screen in white text, with a black background.

Test-feedback phase

Participants were asked to recall the response words as fast as possible, and their abilities were tested and strengthened using a drop-off procedure (Levy & Anderson, 2012). Cue words were shown for 5 seconds. Correctly recalled words would not appear again in this phase. Failure to correctly recall a word resulted in the correct word being shown in blue for 2.5 seconds. Testing lasted until either all pairs were recalled correctly once, or the 25-minute

limit was exceeded. In the latter case, the experimenter skipped forward to the next phase and excluded the participant afterward.

Practice Think/No-Think Phase

Participants were instructed that certain words would now appear in green, and some in red. Green words were supposed to be recalled as before, and red words were now neither to be thought of or recalled out loud. It was emphasized that this suppression should occur solely while focusing on the red word, rather than, for example, thinking of unrelated items to forget the response word. All cues were now presented for 3.5 seconds. Failure to correctly recall green words again resulted in the correct word being shown in blue for 2 seconds. Recalling anything for red words resulted in an error message in red (i.e., "Do not recall no-think items"). This practice phase consisted of 48 trials including 12 filler baseline pairs being shown in a random order. Both the six Think and No-Think items appeared four times. After this practice phase, the first diagnostic questionnaire appeared. The participant indicated their responses, and, if necessary, the experimenter clarified and emphasized certain instructions (e.g., emphasizing the importance of not thinking of the response for the red words).

Think/ No-Think phase

This phase consists of the same setup as the previous practice phase. Now, 24 cues divided equally into Think and No-Think, without any baseline cues, were shown. Three blocks consisting of 96 trials each were utilized, and participants were given a 1-minute break after the first and second blocks. The total critical trial number was therefore 288, with each Think and No-Think cue being shown 12 times each (144 trials for each condition). Halfway through the second block, the diagnostic questionnaire was conducted the same way as in its first occurrence.

Same-Probe (SP) test

Participants were asked to now recall all words out loud, regardless of their previous color. All words (including baseline items) were presented in white in a random order, starting with 8 filler cues. If participants did not recall the correct word, they did not receive feedback.

End-of-Session Questionnaire & Debriefing

In the in-lab setting, the experimenter closed the Inquisit task and opened the link to the End-of-Session Qualtrics questionnaire on the participant's computer. In the online setting, the experimenter pasted the Qualtrics link in the Google Meet chat function. The questionnaire included the diagnostic questionnaire, compliance questionnaire, session evaluation, and exploratory measures. In the in-lab setting, the participants were informed that the researcher's screen would be turned off for privacy reasons, and that the separating door would be closed until the participant finished. Once the questionnaire was done, the end-of-study screen asked the participant to notify the researcher. After the completion, participants were thanked, and debriefed minimally to avoid influencing future participants. A full debriefing was clarified to be received by April 15th, 2024. Lastly, the experimenter granted compensation in SONA.

Statistical analysis

After 108 total completions, the data was anonymized by the supervisor and subsequently analyzed in SPSS (Version 29.0.1.0). A new variable representing the SIF effect was created by subtracting the no-think item scores from the baseline item scores. Additionally, both the ERQ and LSRP scores were separately summed to create a *Total ERQ score* and a *Total LSRP score* variable. Outliers were determined by a preregistered rule using the interquartile range (IQR), i.e. data points exceeding 3 times the IQR below the first quartile (Q1) or above the third quartile (Q3). In line with previous studies, the distribution of the differences for the *t*-test (i.e., the SIF effect variable) was expected to be normal (see

Wiechert et al., 2023, and Wessel et al., 2023), and, in accordance with our preregistration, no data transformation was subsequently needed. Furthermore, the total ERQ score and total LSRP score variables were examined in SPSS using histograms, box plots, and descriptives to assess normality and skewness (see Appendices C to E for all relevant distributions). To investigate the suppression-induced forgetting effect (comparing *no-think* against *baseline* targets) as well as the positive control effect (comparing *think* against *baseline* targets), two paired sample *t*-tests were conducted for each setting (online and in-lab). Paired sample *t*-tests were chosen in line with previous studies, but also due to their relative robustness to moderate deviations from normality. Pearson correlations comparing SIF scores against total ERQ and total LSRP scores were conducted for the exploratory analysis as a simple method to explore potential linear relationships. Unlike Wiechert et al. (2023) and Wessel et al. (2023), no Bayesian analyses were conducted for analysis simplicity.

Results

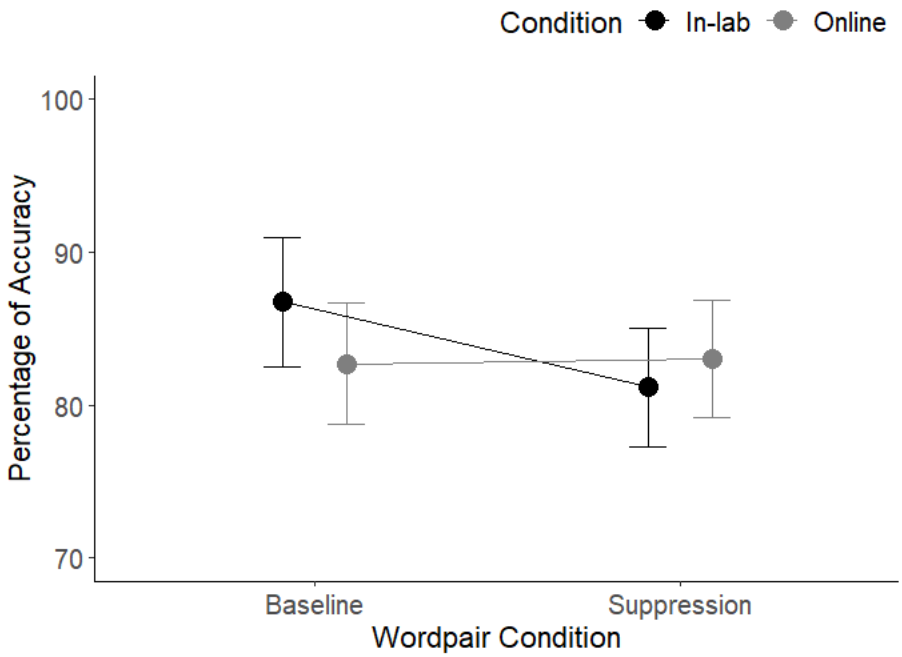
Suppression-Induced Forgetting effect

In order to examine the first hypothesis of a possible suppression-induced forgetting effect (i.e., worse recall of No-Think than Baseline items), two paired samples *t*-tests were conducted for each testing environment. Visual inspection of the total SIF score variable using descriptive statistics in SPSS showed a normal distribution as expected (see Appendix C). One extreme outlier (i.e., a 67% difference between baseline and suppression scores, outside the 3 IQR specification) in the online condition was excluded from the analysis. The exclusion did not impact the results. As can be seen in Figure 2, results from the in-lab condition suggested No-Think items (M = 81.17%, SD = 14.59%) being recalled less than Baseline items (M = 86.73%, SD = 10.43%), indicating a successful suppression-induced forgetting effect ($t(53) = 3.28$, $p < .001$; $d = 0.45$, [95% CI 0.02; 0.09]). In the online condition, suppression scores (M = 83.02%, SD = 14.52%) did not differ significantly from

Baseline scores (M = 82.70%, SD = 14.97%), suggesting no significant SIF effect ($t(52)$ = -0.12, p = .451; d = -0.02, [95% CI -0.05; 0.05]).

Figure 2

Differences between Baseline condition average accuracy and Suppression condition average accuracy divided by the testing condition



Note. Error bars represent the 95% Confidence Interval. This layout was inspired by Figures 3 and 4 in Wiechert et al. (2023).

Additionally, a positive control effect was found, suggesting a significant difference between Baseline and Recall scores for both the in-lab ($t(53)$ = 8.27, p < .001; d = 1.13, [95% CI 0.09; 0.15]) and online condition ($t(53)$ = 7.76, p < .001; d = 1.06, [95% CI 0.12; 0.2]).

Exploratory Analyses

To examine the second and third hypotheses, two Pearson correlations were conducted. The first one examined participants' total ERQ score with the SIF effect score, while the second one investigated the total LSRP score with the SIF effect score. Visual analyses of the data point distributions of the total ERQ and total LSRP scores did not show

any relevant deviations (see Appendix D). No extreme outliers were observed. Contrary to our hypotheses, our results suggested no significant positive correlation between SIF and ERQ scores either in-lab ($r(53) = -.12$, $p = .400$), or online ($r(53) = -.06$, $p = .680$; Total dataset: $r(107) = -.08$, $p = .418$). Similarly, no significant negative correlation between SIF and LSRP scores in-lab ($r(53) = -.02$, $p = .881$), or online ($r(53) = -.07$, $p = .636$; Total dataset: $r(107) = -.04$, $p = .716$) was found. Appendix E shows the scatterplots for these results. It should be noted that, as detailed in the *Materials section*, 3 further questionnaires were also included in our experimental sessions. These were however part of other students' thesis projects, and their questionnaires' results are subsequently not reported.

Furthermore, both Wiechert et al.'s (2023) and Wessel et al.'s (2023) data suggested a possible difference between the three counterbalancing conditions, with Condition B showing the largest difference. To examine this, a repeated measures ANOVA was conducted with the Instruction (Baseline vs. No-think) as the repeated measures factor, and the Counterbalancing Conditions (A vs. B vs. C) as the between participants factor, as well as a post hoc Tukey's range test. Results indicated no significant interaction between the Conditions and the Instruction ($F(2,105) = 1.12$, $p = .330$, $\eta^2 = .02$). While Condition B did show a higher average SIF mean ($M = 6.71\%$, $SD = 16.28\%$) than Condition A ($M = 1.39\%$, $SD = 13.87\%$) and C ($M = 1.62\%$, $SD = 20.4\%$; see Appendix F for a full boxplot), the post hoc Tukey HSD test indicated no statistically significant differences in any condition pair (A&B: $p = .951$; A&C: $p = .214$; B&C: $p = .352$).

Discussion

This study intended to replicate the findings by Wiechert et al. (2023) and Wessel et al. (2023) by conducting two separate TNT experiments in an online or laboratory environment. The well-powered study by Wiechert et al. (2023) found no significant suppression-induced forgetting (SIF) effect in their online environment sample, while using

international participants. Wessel et al. (2023) replicated this study in-lab with first-year university students and found a significant same-probe SIF effect (SIF-SP). Besides this replication, the exploratory analyses investigated the possible relationship of emotional regulation, as an easily testable part of cognitive control, and psychopathy, as a major aspect of forensic psychology, with the SIF phenomenon. Our hypotheses predicted a SIF-SP effect in both the online and in-lab environment, a positive correlation between emotion regulation and memory suppression, and a negative correlation between psychopathy and memory suppression. In line with our hypothesis, our results indicated a significant SIF-SP effect in the in-lab environment, successfully replicating Wessel et al.'s (2023) findings. Contrary to our expectations and similar to Wiechert et al. (2023), no significant SIF-SP effect for participants in the online environment was found. Furthermore, both of our exploratory hypotheses showed non-significant correlations between both emotion regulation and memory suppression scores, and psychopathy and memory suppression.

Previous findings

As our results regarding the SIF effect online and in-lab aligns with both Wiechert et al.'s (2023) and Wessel et al.'s (2023) findings, the testing environment may indeed be an important factor in the memory suppression effect. Previous studies have shown that some cognitive tasks (e.g., reaction time, Stroop tasks, priming tasks) done online and in-lab can be comparable (Semmelmann & Weigelt, 2017; Sauter et al., 2022), yet this study suggests different results for the SIF effect. The TNT task and the corresponding SIF effect might therefore be an example of online and in-lab data not being as easily equatable. In general, one could argue that the online condition is actually closer to real-life situations, where distractions are more common than in a laboratory environment. Therefore, the SIF effect's generalizability may not be as large as originally thought, challenging the previous

assumption that the SIF effect is a robust phenomenon across different contexts (Anderson & Huddleston, 2012; Clark, 2021; Stramaccia et al., 2021).

Considering the exploratory analyses regarding emotion regulation and psychopathy were of a novel nature, there are few findings to be referenced. Nonetheless, starting with emotion regulation, the general relationship as a part of cognitive control is well established (Pruessner, 2020), and cognitive control capacity does have an influence on memory suppression (Chen et al., 2022). Our results do not challenge Chen et al.'s (2022) findings but may simply suggest that emotion regulation is not one of the relevant aspects in cognitive control that influences memory suppression. Additionally, the findings by Gillespie et al. (2022) suggested that psychopathy is modestly associated with response inhibition problems in go/no-go tasks. Our study's findings do not contradict this related literature, and we rather propose the novel idea that psychopathy may simply not be correlated with the TNT task that was utilized for this replication.

Theoretical implications

As mentioned in the previous section, some cognitive tasks done online and in-lab can be comparable (Semmelmann & Weigelt, 2017; Sauter et al., 2022), and our results suggest that this might not apply to the Think/No-Think paradigm. Furthermore, the idea that the online condition may be closer to real-life situations could explain the discrepancy between the online and in-lab findings. More specifically, the online environment lends itself to more distractions, less control by the experimenter, and subsequently maybe even a less motivated participant (see *Methodological considerations* for more extensive limitations). Especially the factor of more distractions relates to the real-life setting, which is all but stimuli-free. Acknowledging the influence of context and environmental factors shows the need for a more nuanced understanding of memory suppression, beyond the sterile environment of the laboratory. More generally, a non-significant SIF effect could suggest that other factors such

as testing context, individual differences, or specific task demands play a more significant role in memory suppression results than previously thought.

Considering that our study did not find a significant correlation between emotion regulation and the SIF effect may simply mean that emotion regulation is not a part of cognitive control that influences memory suppression. Additionally, it is feasible that only emotionally taxing items result in a SIF effect, as the common definition of emotion regulation includes "managing emotional experiences" (Kok, 2017, p.1). The positive relationship between emotion regulation and learning enhancement (Martin & Ochsner, 2017) could then not be sufficient to offset this lack of emotional items.

On another note, while the findings of psychopathy and response inhibition by Gillespie et al (2022) were significant, they reported a relatively small effect size, and concluded that response inhibition might not be a central aspect of psychopathy. If response inhibition only occurs in certain subsections of psychopathic individuals, a strong correlation with the SIF effect would subsequently be unlikely. However, it is also possible that alternative mechanisms such as the reduced susceptibility for memory intrusions (Moul & Caroline, 2017) could compensate for any deficits in response inhibition (Gillespie et al., 2022). The different subtypes of psychopathy (see Murphy & Vess, 2003; Falkenbach & Zappale, 2021) may also correlate differently with the TNT task, if these certain subtypes exhibit greater deficits in response inhibition than others. To our knowledge, this has however not yet been investigated. As a last point, there is also the possibility that emotion regulation, psychopathy, or both simply do not have a significant influence on a person's ability to suppress certain memories, yet more research needs to be done to confirm these patterns.

Methodological considerations

In general, our study had several strengths, including a relatively large sample size, as well as the close replication of previous studies by (1) using the same materials, (2)

conducting the study at the same University, and subsequently (3) having very similar sample characteristics (i.e., first-year international psychology students fluent in English). It should however be noted that despite our student sample being randomly distributed between the two conditions, a non-significant SIF-SP effect was still observed for the online condition, leading us to believe that the main reason for the different findings is indeed the testing condition, rather than sample characteristics. Three possible limitations might explain our results of online participants scoring differently than in-lab participants, as well as the majority of the post-hoc exclusions occurring in the online condition. First, participants were not in a clearly defined "experimental" surrounding, but rather in their home or a quiet public space. This may have removed some of the associated formality that many participants, especially first-year students, likely feel in an actual laboratory environment. Subsequently, participants might be less concerned about their performance and exert less effort on the task. While one could implement even more stringent attention and motivation checks in the online environment to possibly observe a SIF effect, it is unclear if this is enough to observe a significant effect, and if too many compliance checks reduce external validity. Second, and relatedly, instruction compliance might not have been as successful in the online condition. For example, experimenters reported multiple instances where participants had notifications or alarms go off on their phone or laptop, despite previous assurance that all possible disturbances were disabled. While instruction compliance was regularly checked by experimenters, participants did not have to show their screen or the entirety of their surroundings, leaving the possibility of certain "invisible" distractions. Third, stimuli presented through an online session might differ from the more controlled laboratory setting (e.g., reduced text size and/or clarity, micro stutters, or audio issues), resulting in reduced participant engagement. While longer persisting internet problems or audio issues are possible

to detect for the researcher, shorter disturbances might not be, yet they still are distracting to the participant.

Both the less defined experimental setting as well as the reduced instruction compliance may also partially explain the non-significant correlations of the ERQ and LSRP with the SIF-SP score in the online condition. However, non-significant results were found regardless of condition, leading us to suggest some further possible limitations. In general, both the ERQ, but especially the LSRP, measure relatively broad constructs and might therefore not capture the specific psychopathic traits or emotion regulation strategies that could have an influence on the SIF effect. Regarding psychopathy, it is also possible that a social desirability bias may have also influenced the participants to underreport especially psychopathic tendencies. However, no floor effect was observed in our data, and both the ERQ and LSRP have shown acceptable internal validity (Gouveia et al., 2018; Psederska et al., 2020), leading us to be less concerned about this particular limitation.

Implications for future research

Considering the previous sections, we suggest that more research needs to be done in order to clarify the different relationships in the TNT paradigm. More specifically, replications of this study in a wider range of controlled settings, both online and in-lab, are recommended. Slight variations of, for example, cognitive load may provide insights into the robustness of the SIF effect under various levels of distraction, while still maintaining the core idea of suppression being effortful. Additionally, to expand the understanding of the relationship between cognitive control and memory suppression, future studies should explore different aspects of cognitive control, such as working memory, attentional control, or cognitive flexibility. Investigating these aspects could help find the specific mechanisms by which cognitive control influences memory suppression. For example, working memory might help maintain suppression goals, attentional control could aid in resisting interference

of unwanted memories, and cognitive flexibility could improve suppression strategies. Besides this, the different subtypes of psychopathy and their possible relationship with memory suppression may also be of interest in order to explore if there are differences between them. Furthermore, due to resource constraints, our study only used same-probe testing, neglecting the possible influence that independent probe testing may have. While Wiechert et al. (2023) did not find significant results regardless of the probe test, employing both methods would allow for a more "complete" picture when also utilizing online and in-lab testing.

Lastly, Wiechert et al. (2023) suggested a possible difference between the SIF effect and the three Counterbalancing Conditions (Conditions A, B, and C), Wessel et al. (2023) found statistically significant results investigating a possible difference between the SIF effect and the three Counterbalancing Conditions (Conditions A, B, and C). Slightly higher SIF means were found for Counterbalancing Condition B, but the findings suggested non-significance among the different conditions. The Counterbalancing Conditions (taken from Benoit & Anderson, 2012) have been successfully implemented in the past, and now in our own, yet the replicated studies showed different results. To ensure internal validity, and that conclusions of future TNT experiments are as accurate as possible, further exploration into these counterbalancing patterns is favored.

Conclusion

In the TNT literature, null findings of the corresponding suppression-induced forgetting effect are relatively rare. We replicated studies by Wiechert et al. (2023) and Wessel et al. (2023), which suggested that an online TNT task, compared to the in-lab task, might not result in a significant SIF effect. Both studies were successfully replicated, and a significant SIF-SP effect was found only in the laboratory condition, suggesting that the TNT task might not be as generalizable as previously thought. Furthermore, exploratory analyses

were conducted on the possible influences of emotional regulation, as an easily testable part of cognitive control, and psychopathy, as a major aspect of forensic psychology, on the SIF phenomenon. To our knowledge, this is the first time these specific aspects have been tested with the TNT task. No significant results were found for the two analyses, but the exact relationship remains unclear due to its novelty. Future research should try to further explore the contributions of cognitive control and psychopathy on suppression-induced forgetting, as well as replicate our findings regarding the testing context to ensure the TNT tasks' external validity.

References

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural Systems Underlying the Suppression of Unwanted Memories. *Science*, *303*(5655), 232-235. https://doi.org/10.1126/science.1089504

Anderson, M.C., Huddleston, E. (2012). Towards a Cognitive and Neurobiological Model of Motivated Forgetting. In *True and False Recovered Memories: Toward a Reconciliation of the Debate* (Vol. 58). https://doi.org/10.1007/978-1-4614-1195-6_3

Battista, F., Lanciano, T., Curci, A., Mirandola, C., & Otgaar, H. (2023). I lie because I am good at: Psychopathic traits do not influence the effects of fabrication on memory. *Journal of Cognitive Psychology*, *35*(6–7), 635–649. https://doi.org/10.1080/20445911.2023.2245603

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron, 76*(2), 450-460. https://doi.org/10.1016/j.neuron.2012.07.025

Brewin, C. R. (2011). The nature and significance of memory disturbance in posttraumatic stress disorder. *Annual review of clinical psychology*, *7*, 203-227. https://doi.org/10.1146/annurev-clinpsy-032210-104544

Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images in psychological disorders: characteristics, neural mechanisms, and treatment implications. *Psychological review*, *117*(1), 210–232. https://doi.org/10.1037/a0018113

Castiglione, A., & Aron, A. R. (2021). Unwanted Memory Intrusions Recruit Broad Motor Suppression. *Journal of Cognitive Neuroscience*, *33*(1), 119–128. https://doi.org/10.1162/jocn_a_01642

Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to Forget: Inhibitory-Control Deficits Compromise Memory Suppression in Posttraumatic Stress Disorder. *Psychological Science*, *26*(5), 604-616. https://doi.org/10.1177/0956797615569889

Chen, S., Mao, X., & Wu, Y. (2022). Can't stop thinking: The role of cognitive control in suppression-induced forgetting. *Neuropsychologia, 172*, 108274. https://doi.org/10.1016/j.neuropsychologia.2022.108274

Clark, C. (2021). Forgetting as a consequence of retrieval suppression: a meta-analytic review of the think/no-think paradigm (Doctoral Dissertation, Memorial University of Newfoundland). http://central.bac-lac.gc.ca/.redirect?app=damspub&id=d2955a3d-bc4d-49a9-843d-551bb38677cb

Costanzi, M., Cianfanelli, B., Santirocchi, A., Lasaponara, S., Spataro, P., Rossi-Arnaud, C., & Cestari, V. (2021). Forgetting Unwanted Memories: Active Forgetting and Implications for the Development of Psychological Disorders. *Journal of personalized medicine*, *11*(4), 241. https://doi.org/10.3390/jpm11040241

DeLisi, M. (2009). Psychopathy is the Unified Theory of Crime. *Youth Violence and Juvenile Justice*, *7*(3), 256-273. https://doi.org/10.1177/1541204009333834

Falkenbach, D. M., Reinhard, E. E., & Zappala, M. (2021). Identifying Psychopathy Subtypes Using a Broader Model of Personality: An Investigation of the Five Factor Model Using Model-Based Cluster Analysis. *Journal of interpersonal violence*, *36*(15-16), 7161–7184. https://doi.org/10.1177/0886260519831388

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (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

Garofalo, C., Neumann, C. S., Kosson, D. S., & Velotti, P. (2020). Psychopathy and emotion dysregulation: More than meets the eye. *Psychiatry Research*, *290*, 113160. https://doi.org/10.1016/j.psychres.2020.113160

Gillespie, S. M., Lee, J., Williams, R., & Jones, A. (2022). Psychopathy and response inhibition: A meta-analysis of go/no-go and stop signal task performance. *Neuroscience & Biobehavioral Reviews*, *142*. https://doi.org/10.1016/j.neubiorev.2022.104868

Gouveia, V.V., Moura, H.M., Oliveira, I.C., Ribeiro, M.G., Rezende, A.T., & Brito, T.R. (2018). Emotional Regulation Questionnaire (ERQ): Evidence of Construct Validity and Internal Consistency. *Psico-USF*, *23*(3), 461–471. https://doi.org/10.1590/1413-82712018230306

Gross, J. J., & John, O. P. (2003). *Emotion Regulation Questionnaire (ERQ)* [Database record]. APA PsycTests. https://doi.org/10.1037/t06463-000

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology*, *4*, 217–246. https://doi.org/10.1146/annurev.clinpsy.3.022806.091452

Hare, R. D., Forth, A. E., & Hart, S. D. (1989). The psychopath as prototype for pathological lying and deception. In *Credibility assessment* (pp. 25-49). https://doi.org/10.1007/978-94-015-7856-1_2

Harvey, A. G., Watkins, E., & Mansell, W. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press, USA. https://doi.org/10.1093/med:psych/9780198528883.001.0001

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, *12*(3), 484–489. https://doi.org/10.3758/BF03193792

Hooper, L., Stockton, P., Krupnick, J., & Green, B., (2011). Development, use, and psychometric properties of the Trauma History Questionnaire. *Journal of Loss and Trauma, 16*, 258-283. https://doi.org/10.1080/15325024.2011.572035

Kok, R. (2017). Emotion regulation. In *Encyclopedia of Personality and Individual Differences* (pp. 1-10). https://doi.org/10.1007/978-3-319-28099-8_811-1

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta psychologica*, *127*(3), 623-635. https://doi.org/10.1016/j.actpsy.2007.12.004

Levy, B., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *The Journal of Neuroscience*, *32*(47), 16785–16794. https://doi.org/10.1523/jneurosci.2640-12.2012

Mackie, M. A., Van Dam, N. T., & Fan, J. (2013). Cognitive control and attentional functions. *Brain and cognition*, *82*(3), 301–312. https://doi.org/10.1016/j.bandc.2013.05.004

Martin, R. E., & Ochsner, K. N. (2016). The Neuroscience of Emotion Regulation Development: Implications for Education. *Current opinion in behavioral sciences*, *10*, 142–148. https://doi.org/10.1016/j.cobeha.2016.06.006

McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of psychiatric research*, *83*, 37-46. https://doi.org/10.1016/j.jpsychires.2016.08.001

Miller, J. D., Gaughan, E. T., & Pryor, L. R. (2008). The Levenson Self-Report Psychopathy Scale: An examination of the personality traits and disorders associated with the LSRP factors. *Assessment*, *15*(4), 450-463. https://doi.org/10.1177/1073191108316888

Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical psychology review*, *20*(1), 113–136. https://doi.org/10.1016/s0272-7358(98)00096-8

Moul, C., & Nickerson, A. (2017). Psychopathic personality traits as a protective factor against the development of intrusive memories. *Journal of Experimental Psychopathology*, *8*(1), 2–13. https://doi-org/10.5127/jep.055516

Murphy, C., & Vess, J. (2003). Subtypes of psychopathy: proposed differences between narcissistic, borderline, sadistic, and antisocial psychopaths. *The Psychiatric quarterly*, *74*(1), 11–29. https://doi.org/10.1023/a:1021137521142

Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology: An Interdisciplinary Journal*, *49*(4), 1063–1107. https://doi.org/10.1111/j.1745-9125.2011.00252.x

Pruessner, L., Barnow, S., Holt, D. V., Joormann, J., & Schulze, K. (2020). A cognitive control framework for understanding emotion regulation flexibility. *Emotion (Washington, D.C.)*, *20*(1), 21–29. https://doi.org/10.1037/emo0000658

Psederska, E., Yankov, G. P., Bozgunov, K., Popov, V., Vasilev, G., & Vassileva, J. (2020). Validation of the Levenson Self-Report Psychopathy Scale in Bulgarian Substance-Dependent Individuals. *Frontiers in psychology*, *11*, 1110. https://doi.org/10.3389/fpsyg.2020.01110

Qualtrics [Computer Software], Qualtrics, Provo, Utah, https://www.qualtrics.com/

Ragbeer, S. N., & Burnette, M. L. (2013). Remember me? Psychopathic traits and emotional memory in an undergraduate sample. *Personality and Individual Differences*, *54*(3), 372–377. https://doi.org/10.1016/j.paid.2012.10.001

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of personality and social psychology*, *79*(3), 410–424. https://doi.org/10.1037//0022-3514.79.3.410

Sauter, M., Stefani, M. & Mack, W. (2022). Equal Quality for Online and Lab Data: A Direct Comparison from Two Dual-Task Paradigms. *Open Psychology*, *4*(1), 47-59. https://doi.org/10.1515/psych-2022-0003

Semmelmann, K., Weigelt, S. (2017). Online psychophysics: reaction time effects in cognitive experiments. *Behavior Research Methods*, *49*(4), 1241–1260. https://doi.org/10.3758/s13428-016-0783-4

Stramaccia, D. F., Meyer, A.-K., Rischer, K. M., Fawcett, J. M., & Benoit, R. G. (2021). Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *Journal of Experimental Psychology: General*, *150*(5), 828–850. https://doi.org/10.1037/xge0000971

van der Meer, C. A. I., Te Brake, H., van der Aa, N., Dashtgard, P., Bakker, A., & Olff, M. (2018). Assessing Psychological Resilience: Development and Psychometric Properties of the English and Dutch Version of the Resilience Evaluation Scale (RES). *Frontiers in psychiatry*, *9*, 169. https://doi.org/10.3389/fpsyt.2018.00169

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale - Revised. In J. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399-411). Guilford. https://link.springer.com/chapter/10.1007/978-0-387-70990-1_10

Wessel, I., Albers, C. J., Zandstra, A. R. E., & Heininga, V. E. (2020). A multiverse analysis of early attempts to replicate memory suppression with the Think/No-think Task. *Memory (Hove, England)*, *28*(7), 870–887. https://doi.org/10.1080/09658211.2020.1797095

Wessel, I., Lehmann, R., & Wiechert, S. (2023). A Replication of Wiechert et al.'s (2023) Online Think/No-Think Study in the Lab. PsyArXiv. https://doi.org/10.31234/osf.io/m5kjy

Wiechert, S., Loewy, L., Wessel, I., Fawcett, J. M., Ben-Shakhar, G., Pertzov, Y., & Verschuere, B. (2023). Suppression-induced forgetting: A pre-registered replication of the think/no-think paradigm. *Memory*, *31*(7), 989-1002. https://doi.org/10.1080/09658211.2023.2208791

Zwaan, R. A., Pecher, D., Paolacci, G., Bouwmeester, S., Verkoeijen, P. P. J. L., Dijkstra, K., & Zeelenberg, R. (2018). Participant Nonnaiveté and the reproducibility of cognitive psychology. *Psychonomic Bulletin & Review*, *25*(5), 1968–1972. https://doi.org/10.3758/s13423-017-1348-y

Appendix A

Emotion Regulation Questionnaire

This Appendix shows the questions of the Emotion Regulation Questionnaire and how it was presented to participants in Qualtrics.

We would like to ask you some questions about your emotional life, in particular, how you control (that is, regulate and manage) your emotions.The questions below involve two distinct aspects of your emotional life.
One is your emotional experience, or what you feel like inside. The other is your emotional expression, or how you show your emotions in the way you talk, gesture, or behave.
Although some of the following questions may seem similar to one another, they differ in important ways.

	1 (strongly disagree)	2	3	4 (neutral)	5	6	7 (strongly agree)
When I want to feel more positive emotion (such as joy or amusement), I change what I'm thinking about.	○	○	○	○	○	○	○
I keep my emotions to myself.	○	○	○	○	○	○	○
When I want to feel less negative emotion (such as sadness or anger), I change what I'm thinking about.	○	○	○	○	○	○	○
When I am feeling positive emotions, I am careful not to express them.	○	○	○	○	○	○	○
When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm.	○	○	○	○	○	○	○

	1 (strongly disagree)	2	3	4 (neutral)	5	6	7 (strongly agree)
I control my emotions by not expressing them.	○	○	○	○	○	○	○
When I want to feel more positive emotion, I change the way I'm thinking about the situation.	○	○	○	○	○	○	○
I control my emotions by changing the way I think about the situation I'm in.	○	○	○	○	○	○	○
When I am feeling negative emotions, I make sure not to express them.	○	○	○	○	○	○	○
When I want to feel less negative emotion, I change the way I'm thinking about the situation.	○	○	○	○	○	○	○

Appendix B

Levenson Self-Report Psychopathy Scale

This Appendix shows the questions of the Levenson Self-Report Psychopathy Scale and how it was presented to participants in Qualtrics.

The test consists of twenty six statements that could possibly apply to you.
You must rate each on how much you agree with it on a scale of (1) strongly disagree (2) disagree (3) neither agree nor disagree (4) agree (5) strongly agree.
Most people will will complete the test in five minutes or so.

	Disagree		Neutral		Agree
Success is based on survival of the fittest; I am not concerned about the losers.	○	○	○	○	○
I find myself in the same kinds of trouble, time after time.	○	○	○	○	○
For me, what's right is whatever I can get away with.	○	○	○	○	○
I am often bored.	○	○	○	○	○
In today's world, I feel justified in doing anything I can get away with to succeed.	○	○	○	○	○
I find that I am able to pursue one goal for a long time.	○	○	○	○	○
My main purpose in life is getting as many goodies as I can.	○	○	○	○	○
	Disagree		Neutral		Agree
I don't plan anything very far in advance.	○	○	○	○	○
Making a lot of money is my most important goal.	○	○	○	○	○
I quickly lose interest in tasks I start.	○	○	○	○	○
I let others worry about higher values; my main concern is with the bottom line.	○	○	○	○	○
Most of my problems are due to the fact that other people just don't understand me.	○	○	○	○	○
People who are stupid enough to get ripped off usually deserve it.	○	○	○	○	○
Before I do anything, I carefully consider the possible consequences.	○	○	○	○	○

	Disagree		Neutral		Agree
Looking out for myself is my top priority.	○	○	○	○	○
I have been in a lot of shouting matches with other people.	○	○	○	○	○
I tell other people what they want to hear so that they will do what I want them to do.	○	○	○	○	○
When I get frustrated, I often "let off steam" by blowing my top.	○	○	○	○	○
I would be upset if my success came at someone else's expense.	○	○	○	○	○
Love is overrated.	○	○	○	○	○
I often admire a really clever scam.	○	○	○	○	○

	Disagree		Neutral		Agree
I make a point of trying not to hurt others in pursuit of my goals.	○	○	○	○	○
I enjoy manipulating other people's feelings.	○	○	○	○	○
I feel bad if my words or actions cause someone else to feel emotional pain.	○	○	○	○	○
Even if I were trying very hard to sell something, I wouldn't lie about it.	○	○	○	○	○
Cheating is not justified because it is unfair to others.	○	○	○	○	○

Appendix C

Distributions of the SIF score variable (online and in-lab)

This Appendix shows histograms and QQ-Plots of the SIF score variable in both the online and laboratory conditions. This serves as a visualization of the normal distributions noted in the *Results* section.

Figure C1

Histogram of the SIF score In-Lab

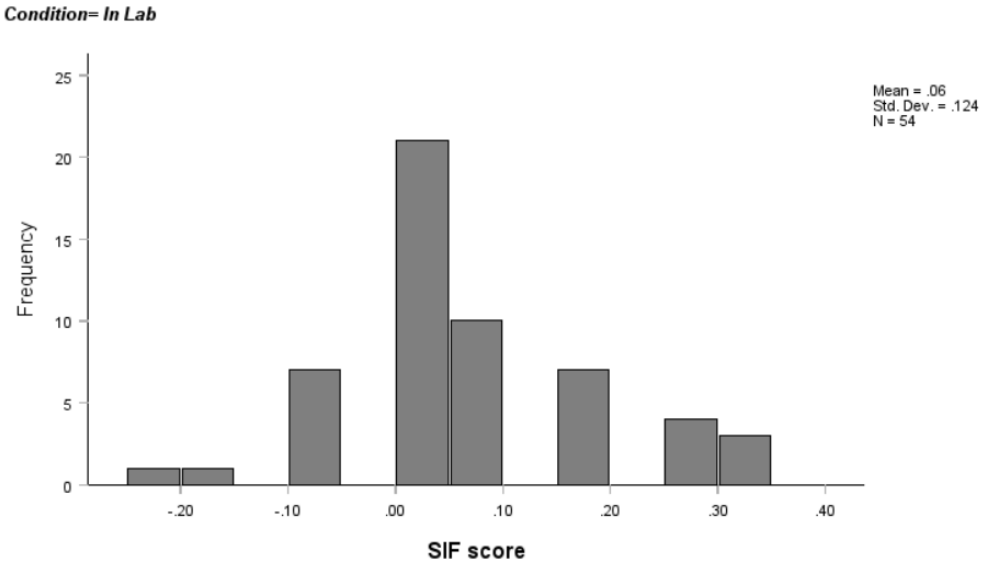


Figure C2

QQ-Plot of the SIF score In-Lab

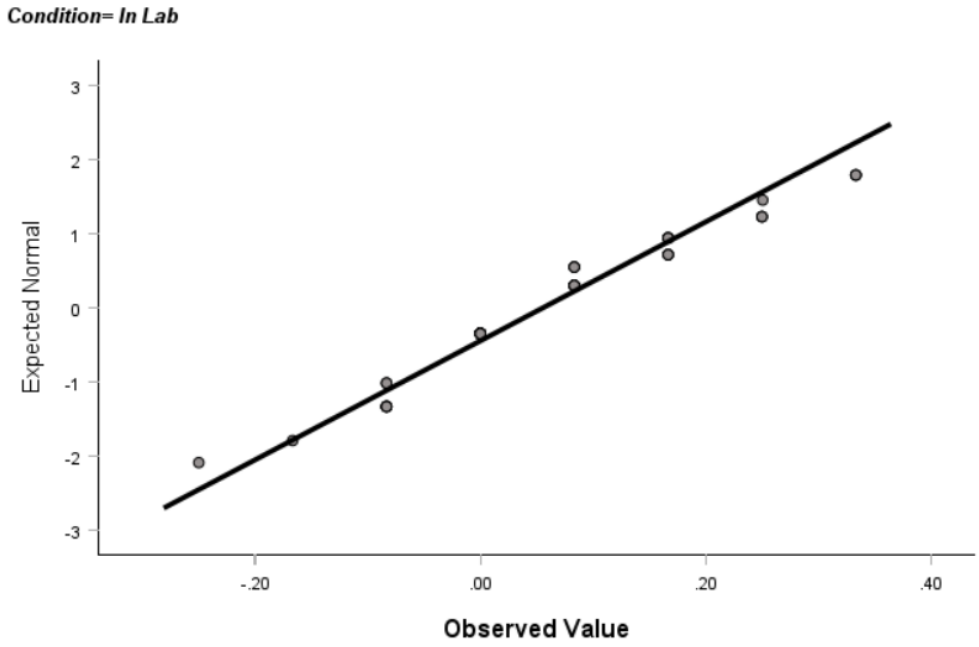


Figure C3

Histogram of the SIF score Online

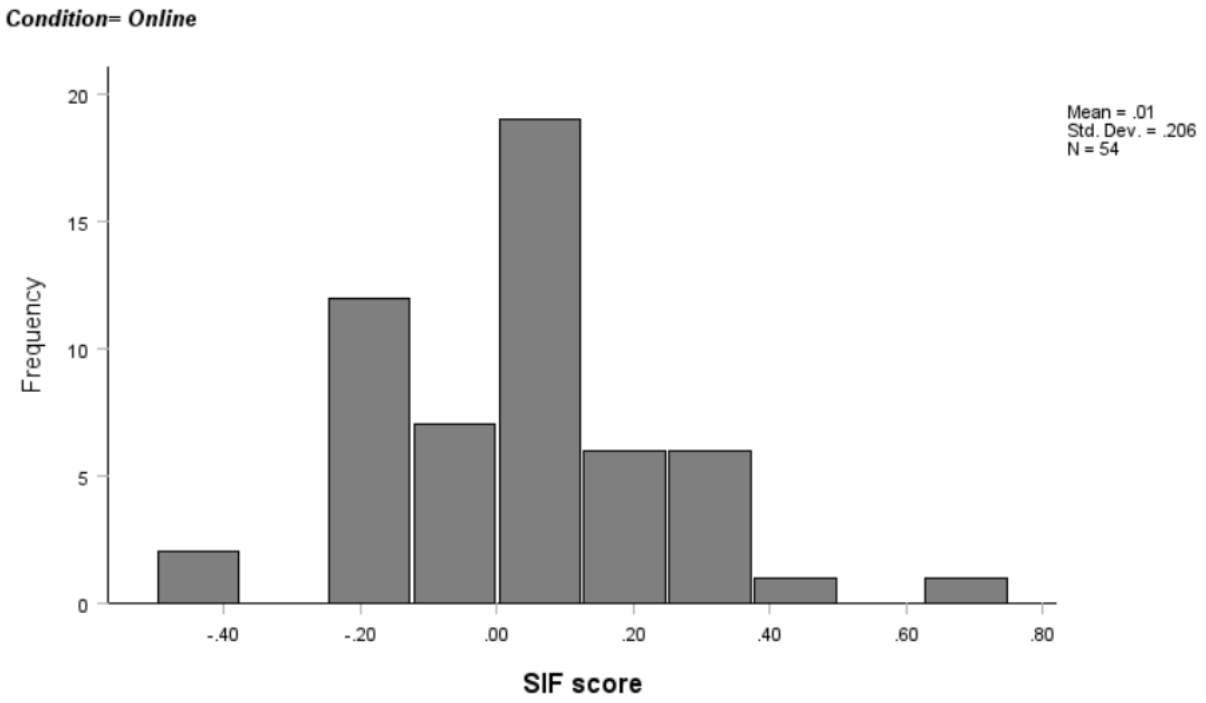
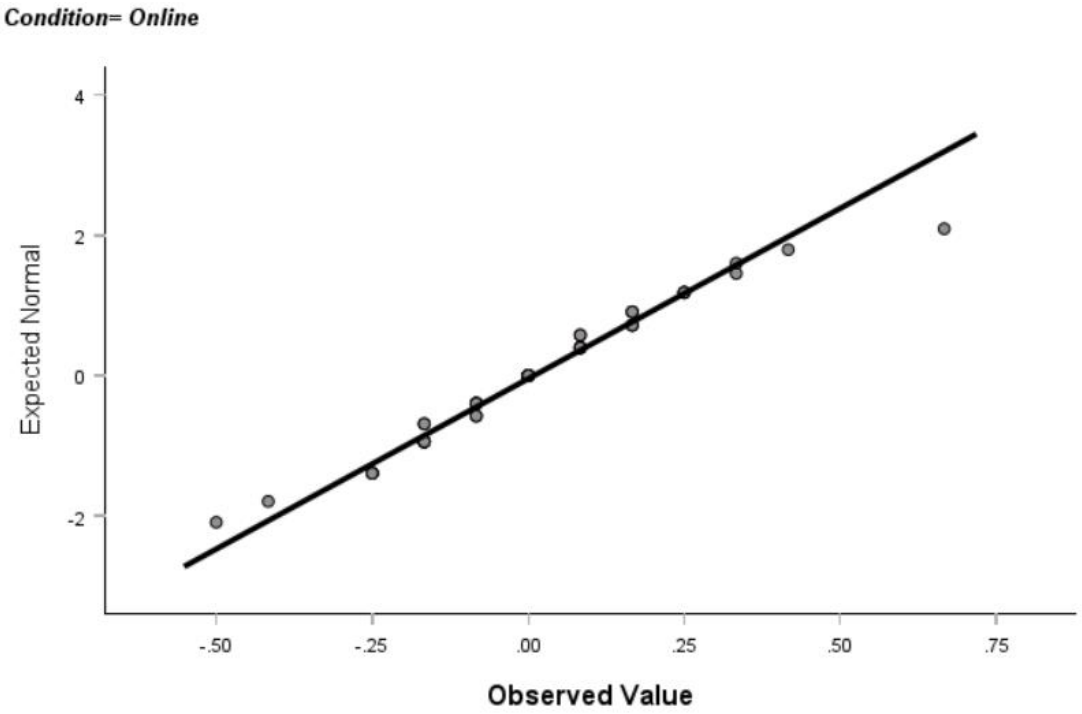


Figure C4

QQ-Plot of the SIF score Online



Appendix D

Distributions of the ERQ and LSRP scores

This Appendix shows the distributions of the total ERQ and total LSRP scores using histograms and QQ-Plots. This serves as a visualization of the questionnaire scores being normally distributed.

Figure D1

Histogram of the Total ERQ score In-Lab

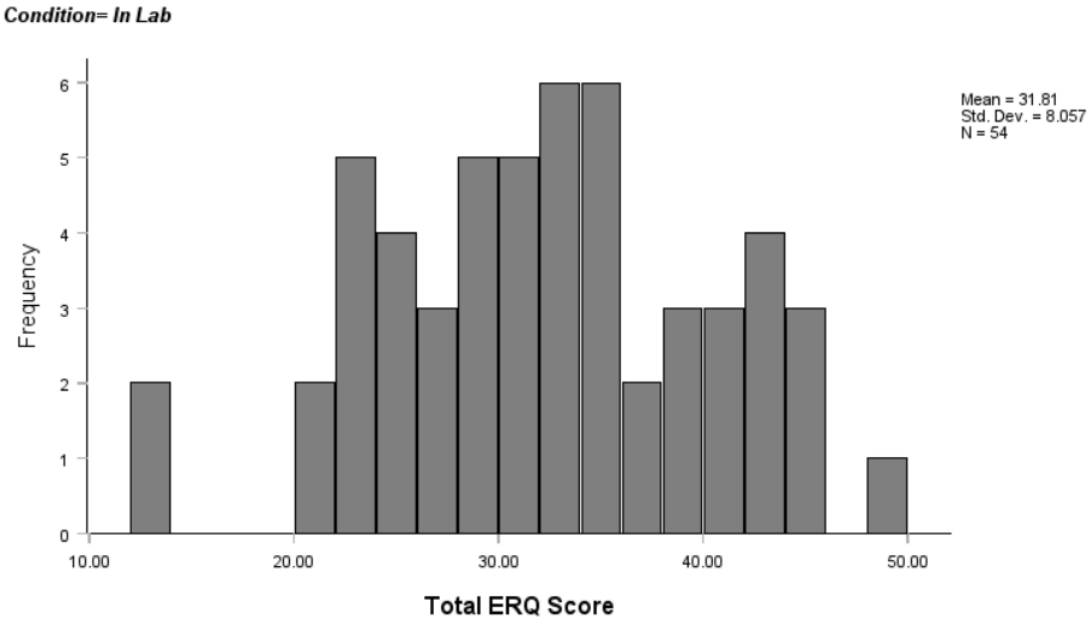


Figure D2

QQ-Plot of the Total ERQ score In-Lab

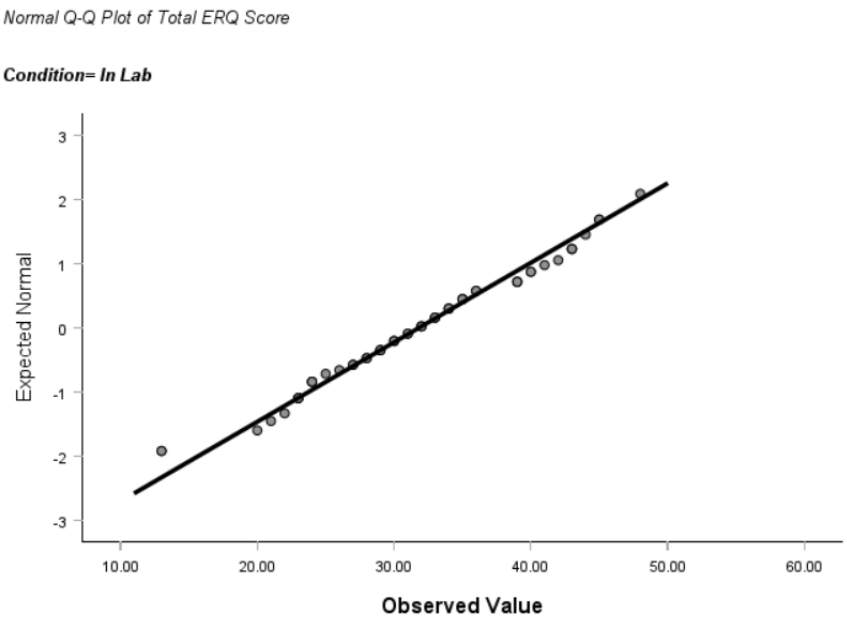


Figure D3

Histogram of the Total ERQ score Online

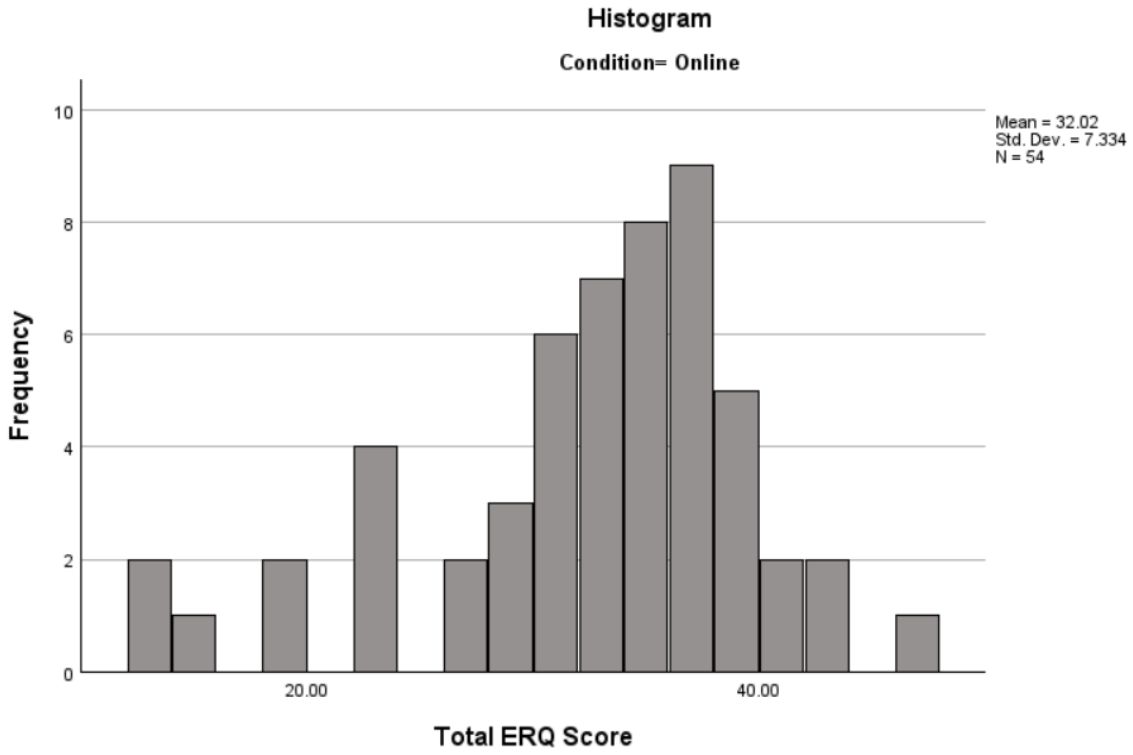


Figure D4

QQ-Plot of the Total ERQ score Online

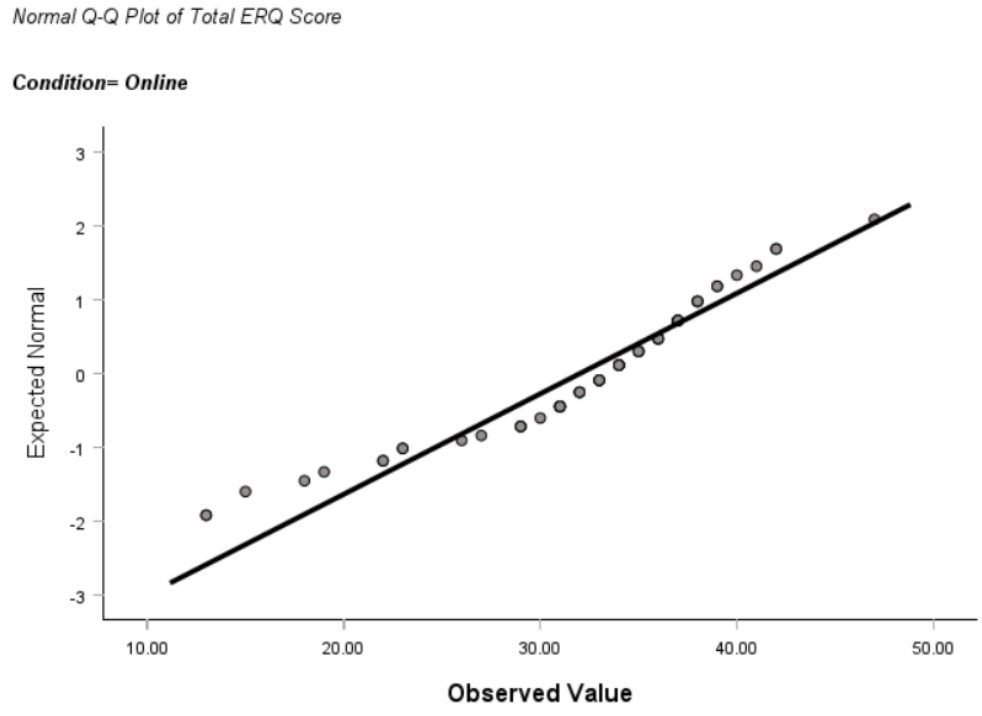


Figure D5

Histogram of the Total LSRP score In-Lab

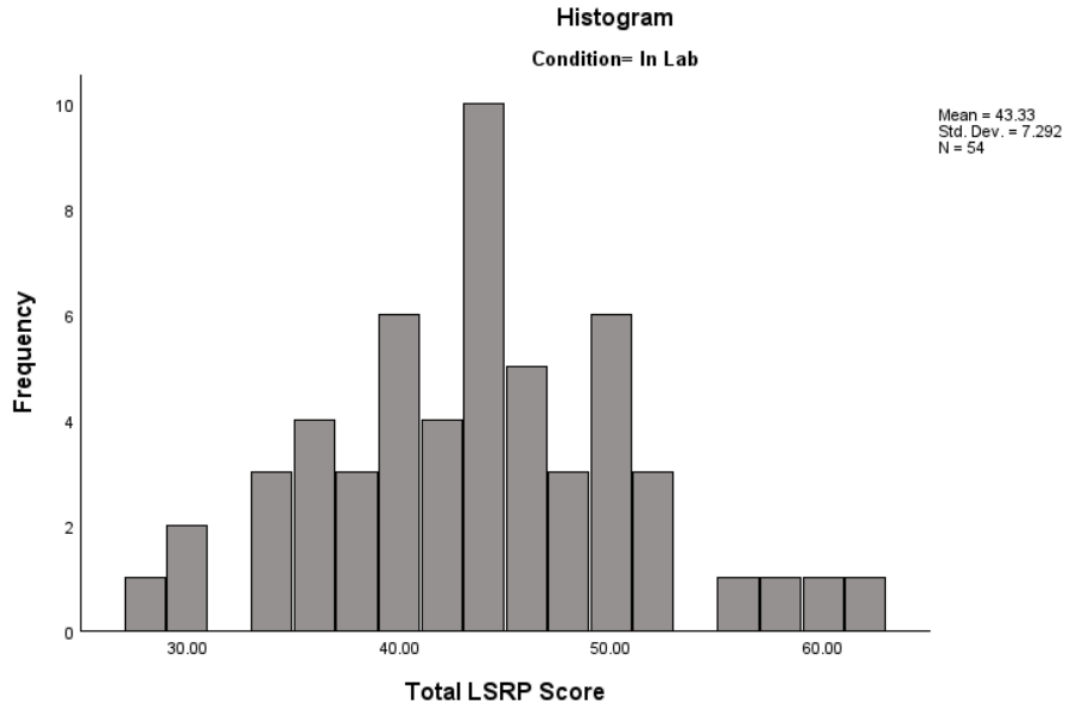


Figure D6

QQ-Plot of the Total LSRP score In-Lab

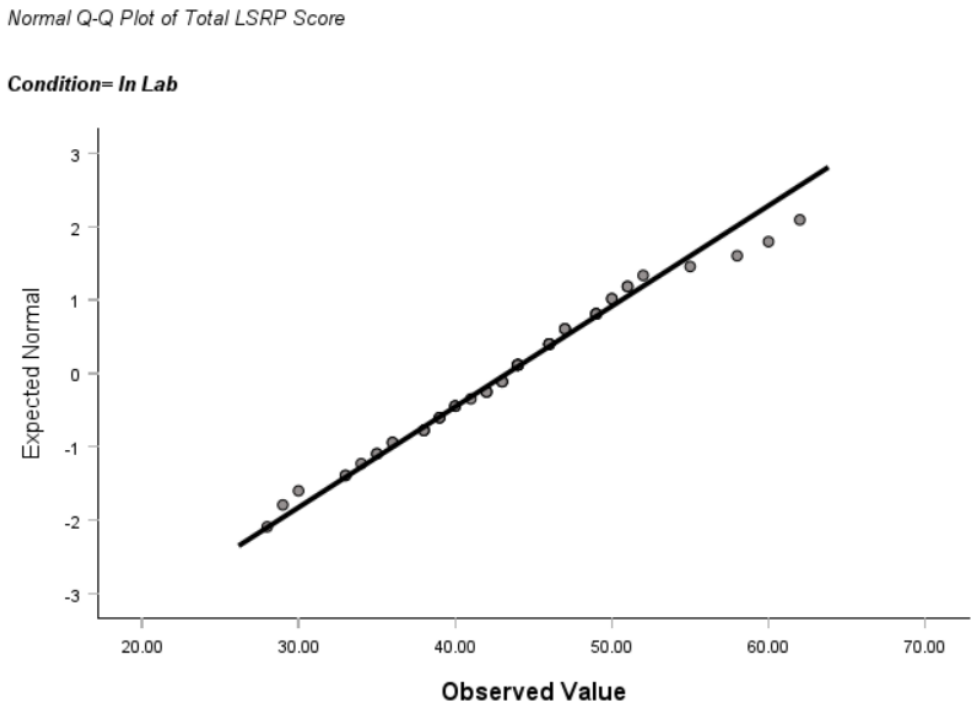


Figure D7

Histogram of the Total LSRP score Online

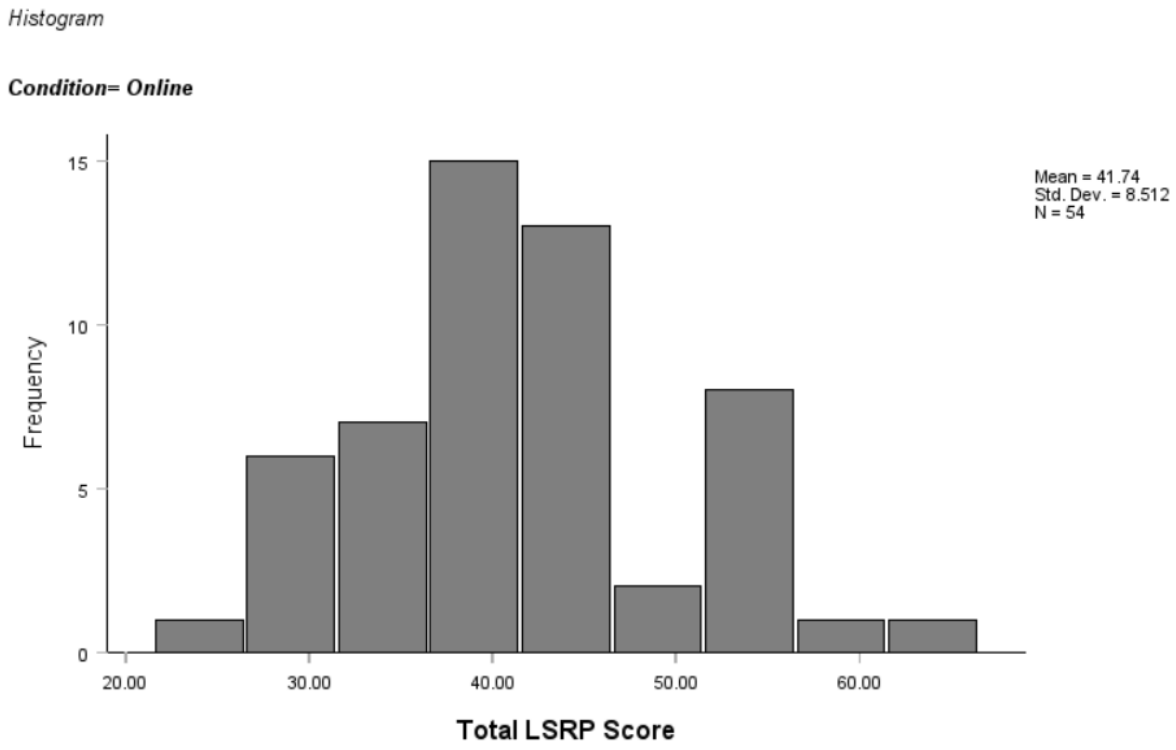
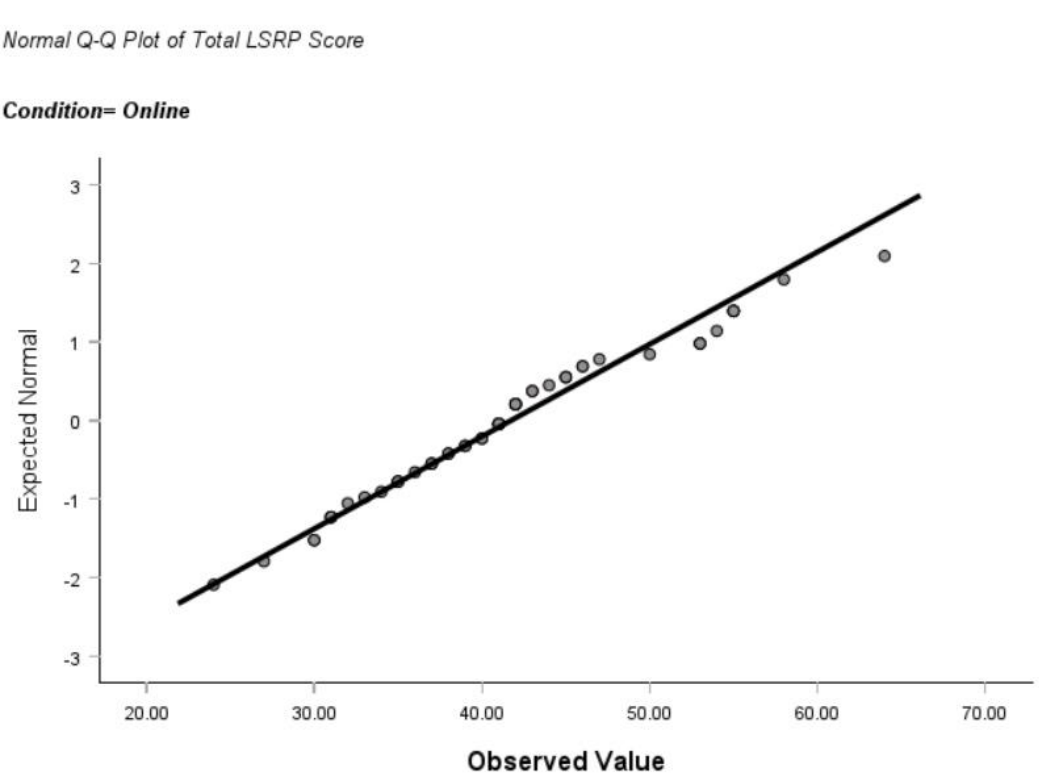


Figure D8

QQ-Plot of the Total LSRP score Online



Appendix E

Scatterplots of ERQ and LSRP scores correlations with the SIF score

This Appendix shows scatterplots of the ERQ and SIF score, as well as the LSRP and SIF score. As reported in *Results*, there are only very minor differences in correlation between the online and in-lab condition. Therefore, only show graphs representing the total correlation between the questionnaires and the SIF score are shown, as we believe these to be adequately representative of the general lack of relationship.

Figure E1

Scatterplot for the correlation between ERQ and SIF scores

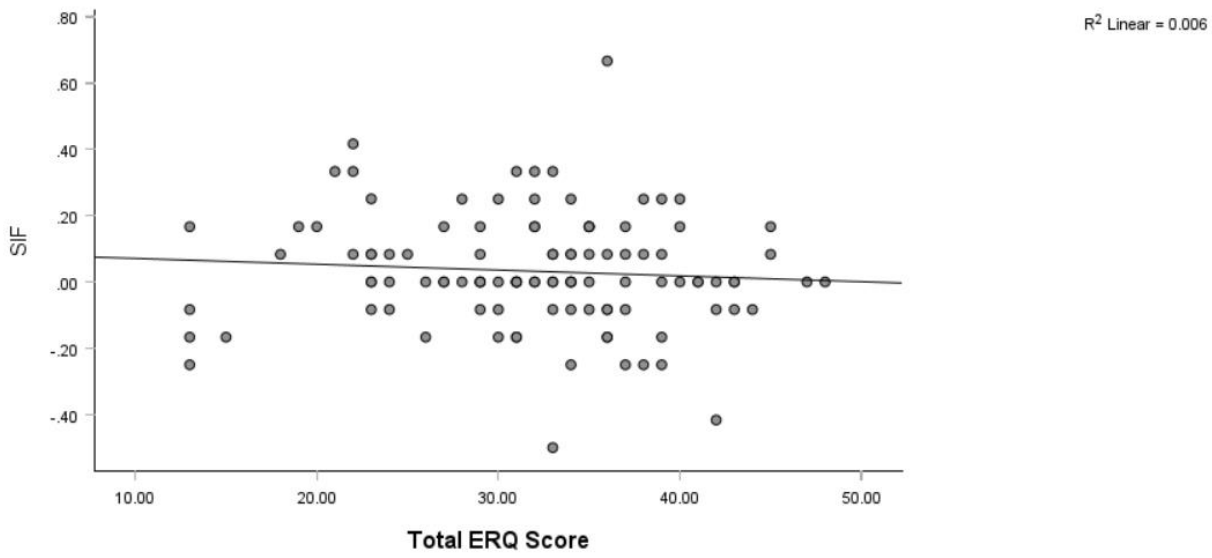
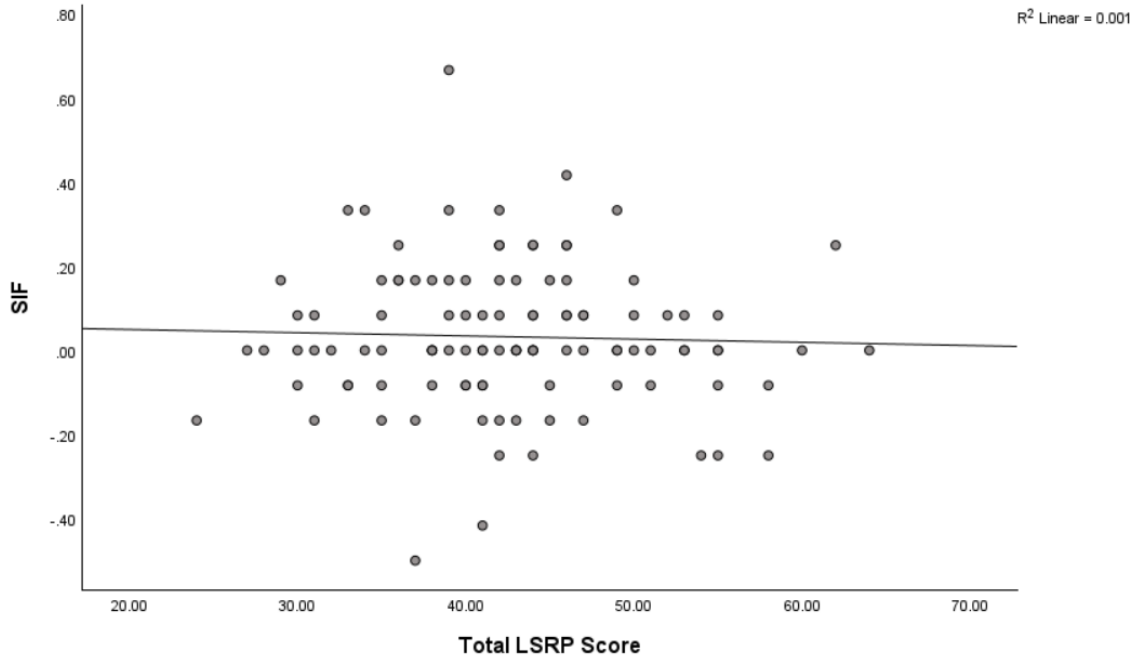


Figure E2

Scatterplot for the correlation between LSRP and SIF scores



Appendix F

Boxplot of Counterbalancing Conditions

This Appendix shows boxplots of the three Counterbalancing Conditions in regard to their average SIF mean scores.

Figure F1

Boxplot representing the different Counterbalancing Conditions' average SIF mean score