

**Universität
Basel**

Fakultät für
Psychologie



Investigating unwanted intrusive thoughts in laboratory settings and a natural environment

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der Fakultät für Psychologie der Universität Basel von

Martin Kollárik

aus Bratislava, Slowakische Republik

Basel, 2021

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch



Universität
Basel

Fakultät für
Psychologie



Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. Roselind Lieb

Prof. Dr. Andrew Gloster

Datum des Doktoratsexamen:

DekanIn der Fakultät für Psychologie



Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

- Kollárik, M., van den Hout, M., Heinzl, C. V., Hofer, P. D., Lieb, R., & Wahl, K. (2020). Effects of rumination on unwanted intrusive thoughts: A replication and extension. *Journal of Experimental Psychopathology*, 11(1).
- Kollárik, M., Hirsch, C. R., Heinzl, C. V., Lieb, R., & Wahl, K. (2021). *Validity of the Counter-App Method in the Assessment of Intrusions* [Manuscript submitted for publication]. Division of Clinical Psychology and Epidemiology, University of Basel.
- Kollárik, M., Heinzl, C. V., Miché, M., Lieb, R., & Wahl, K. (2021). *Exam-related unwanted intrusive thoughts and related neutralizing behaviors: Analogues to obsessions and compulsions* [Manuscript submitted for publication]. Division of Clinical Psychology and Epidemiology, University of Basel.

Basel, 26.05.2021

Martin Kollárik

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Roselind Lieb, for her continuous and invaluable support during my doctoral studies. Her immense knowledge and rich experience gave me the opportunity to grow both professionally as well as personally. I would also like to thank Karina Wahl, who supported me immensely during my studies and introduced me to the field of obsessive-compulsive disorder research. I was able to develop a vast range of skills that will undoubtedly guide my future career and life choices.

I would also like to thank Andrew Gloster, for acting as part of my dissertation committee, for his encouragement, support, and cheerful and enthusiastic attitude toward research.

I wish to thank my colleagues Marcel Miché, Patrizia Hofer, Marcia Rinner, Jeanette Villanueva, Sarah Kuhn, Johannes Massell, Andrea Meyer, Thea Zander-Schellenberg, Eric Mayor, Catherine Coste at the Division of Clinical Psychology and Epidemiology and Division of Clinical Psychology and Intervention, for their constant support, inspiring discussions, and great coffee breaks. My special thanks to Carlotta Heinzl, for her invaluable friendship, constructive feedbacks, and rich discussions. I further wish to thank my coauthors and all other contributors, who supported me during my studies.

My sincere gratitude also goes out to my family and friends, for their encouragement and support during my doctoral studies. My thanks to my mom, stepfather, siblings, grandparents, aunt, uncle, and cousin for supporting me in my life decisions and allowing me to follow my dreams. Most importantly, I would like to thank my wife, for her constant love, support, and patience. Without her, this thesis would not exist.

Table of Contents

Acknowledgements.....	IV
Abbreviations.....	VI
Abstract.....	1
Introduction.....	2
Method.....	7
Participants.....	7
Research design.....	7
Measures and tasks.....	9
Statistical analysis.....	11
Results.....	13
Discussion.....	15
Strengths and limitations.....	18
Implications.....	19
Conclusion.....	20
References.....	21
Appendices A-D.....	27

Abbreviations

EMA	Ecological momentary assessment
OC	Obsessive-compulsive
OCD	Obsessive-compulsive disorder
Severity of ER-UITs-N	Severity of exam-related unwanted intrusive thoughts and related neutralizing behaviors
UITs	Unwanted intrusive thoughts

Abstract

Unwanted intrusive thoughts (UITs) are essential to research of obsessive-compulsive disorder. This thesis aimed (1) to replicate results of a previous study in which rumination about a UIT maintained the urge to neutralize the UIT, (2) to evaluate the validity of a counter-app method in assessment of UIT frequency, and (3) to examine in what aspects exam-related UITs and related neutralizing behaviors (ER-UITs-N) are analogous to obsessions and compulsions. In study one, we temporarily activated a UIT in $N = 105$ students and randomly assigned them to rumination about the UIT, rumination about negative mood, or distraction. We assessed distress, urge to neutralize, depressed mood, and UIT frequency. In the next study, we activated a UIT in $N = 142$ students and assessed the UIT frequency with the counter-app method and three other measures (convergent validity criteria). We also assessed discriminant and predictive validity criteria. Finally, we assessed severity of ER-UITs-N, obsessive-compulsive symptoms, anxiety, distress, urge to neutralize, depressed mood, and stress in $N = 29$ students using the ecological momentary assessment. Rumination about the UIT maintained the urge to neutralize and distress compared to rumination about negative mood and distraction. Correlations between the counter-app frequency and the validity criteria supported convergent and predictive, but not discriminant, validity of the counter-app method. The severity of ER-UITs-N was positively associated with obsessive-compulsive symptoms, anxiety, distress, urge to neutralize, and stress, but not with depressed mood. Results suggest that rumination about UITs might contribute to their persistence. The counter-app method validly assesses UIT frequency but is not appropriate when the differentiation of frequency from other constructs (e.g., duration) is needed. Finally, the ER-UITs-N seem analogous to obsessions and compulsions in some aspects (e.g., association with distress), but not all (e.g., no relation to depressed mood).

Introduction

Brief and sudden unwanted intrusive thoughts (UITs), such as thoughts related to harming others, are normal and universal, experienced by the majority of individuals (Rachman & de Silva, 1978; Radomsky et al., 2014). In 1978, Rachman and de Silva demonstrated that UITs share similarities with obsessions that characterize obsessive-compulsive disorder (OCD). Specifically, the authors concluded that UITs and obsessions are comparable in content but differ in severity: obsessions occur more frequently, are more distressing, and are harder to dismiss than UITs. Expanding upon this finding, Salkovskis (1985) proposed that UITs can escalate into clinical obsessions when they elicit automatic thoughts related to responsibility or blame. These automatic thoughts then cause mood disturbances and to cope with the mood, individuals engage in neutralizing responses (termed compulsions in the context of OCD). Since then, a vast number of studies has investigated UITs in nonclinical and clinical samples, providing evidence that supports the role of UITs in etiology of obsessions (for discussion, see Abramowitz & Jacoby, 2014).

Despite the large number of findings on UITs, there are still areas of interest that remain unexplored, such as the effect of rumination about UITs on obsessive-compulsive (OC) symptoms, valid assessment of UIT frequency in experimental studies, and type of UITs typically occurring in student samples. Authors have argued that UITs can trigger rumination (Harvey et al., 2004), that is, recurrent dwelling on the thoughts while focusing on feelings of distress and on the possible causes and consequences of these feelings (Nolen-Hoeksema, 1991; Nolen-Hoeksema et al., 2008). This was supported by several non-experimental studies that demonstrated associations between rumination and OC symptoms in individuals with no diagnosis of mental disorder (Grisham & Williams, 2009; Wahl et al., 2011) and in individuals diagnosed with OCD (Dar & Iqbal, 2015; Heinzl et al., 2021). In addition, an

experimental study investigated the effects of rumination about a UIT in undergraduate students and individuals from community (Wahl et al., 2019). In this study, rumination about the UIT maintained the urge to neutralize the UIT in comparison to rumination about depressed mood and distraction. Rumination about the UIT did not affect the associated distress, depressed mood, and frequency of UITs. This suggests that rumination about a UIT prevents the natural decay of urge to neutralize the UIT that was observed in other studies with student samples (Rachman et al., 1996; van den Hout et al., 2001). Replication studies are needed to confirm the causal effect of rumination about a UIT on the urge to neutralize the thought. Replicating the result would strengthen the argument that rumination plays a role in OCD.

Another area of interest is the real-time assessment of UIT frequency in experimental studies. A common method is the use of event marking during which participants are instructed to record the occurrences of UITs with a counter (Abramowitz et al., 2001). Although many researchers have employed event marking to assess UITs, there is little evidence about the method's psychometric properties. In a recent study (Wahl, Lieb, et al., 2020), the authors investigated the validity and reliability of a counter-app method (a contemporary event marking) in assessment of UIT frequency. The findings supported convergent validity and test-retest reliability but not discriminant and predictive validity of the counter-app method. The study used retrospective measures to evaluate the psychometric properties of the counter-app method. Since retrospective measures can be prone to memory biases (Sato & Kawahara, 2011), other real-time measures of UITs should be used to validate the counter-app method. For example, thought-sampling and thinking-aloud methods are alternative real-time measures of thoughts (Smallwood & Schooler, 2015; Zanov & Davison, 2010). Studies provide evidence that both methods have convergent validity (Davison et al., 1995; Faber et al., 2018; Steindorf & Rummel, 2020; Zanov & Davison,

2010). Frequency is one of the main aspects that differentiates UITs from clinical obsessions (Bouvard et al., 2017; Rachman & de Silva, 1978) and thus, having a valid assessment tool of UIT frequency is crucial for future research of OCD.

Finally, research on OC symptoms could benefit from considering exam-related UITs and related neutralizing behaviors as analogues to obsessions and compulsions. Previous studies indicate that students typically experience exam-related UITs and engage in related neutralizing behaviors during exam periods (Keinan, 2002; Lepore, 1997; Wahl, Hofer, et al., 2020). By definition, exam-related UITs share characteristics with obsessions (American Psychiatric Association [APA], 2013; McKay et al., 2004; Wahl, Hofer, et al., 2020): For example, both are idiosyncratic, unwanted, intrusive, and recurrent. Exam-related neutralizing behaviors also share characteristics with compulsions by definition (APA, 2013; Wahl, Hofer, et al., 2020): Both act as responses to their respective intrusion in order to handle the experienced anxiety. There are mainly two reasons why exam-related UITs and related neutralizing behaviors might be of interest to research. First, researchers typically investigate UITs in nonclinical samples that occur with very low frequency (Belloch et al., 2004; Julien et al., 2007; Purdon & Clark, 1994b). Second, in experimental studies, a common practice is to activate UITs and neutralizing behaviors that might not be idiosyncratic (person specific) and might not occur outside of laboratory (De Putter et al., 2017; Gagné et al., 2018); obsessions occur frequently (Morillo et al., 2007; Rachman & de Silva, 1978), are highly idiosyncratic (McKay et al., 2004), and are experienced in different situations (Clark & Inozu, 2014; Julien et al., 2007). Thus, a research practice that allows to study more frequently occurring, idiosyncratic UITs in their natural environment would further benefit research on OC symptoms. Exam-related UITs and related neutralizing behaviors are by definition idiosyncratic and might occur frequently in their natural setting (during exam periods). However, no study has investigated in

what aspects the exam-related UITs and related neutralizing behaviors are analogous to obsessions and compulsions so far.

The main objectives of this dissertation were (1) to investigate the effect of rumination about a UIT by replicating the Wahl et al. (2019) study, (2) to evaluate the validity of counter-app method in assessment of UIT frequency using other real-time measures as well as retrospective and prospective measures as validation criteria, and (3) to investigate in what aspects the exam-related UITs and related neutralizing behaviors can be considered analogous to clinical obsessions and compulsions. The analogy would be demonstrated through positive associations of exam-related UITs and related neutralizing behaviors with OC symptoms, core aspects of OCD (anxiety, distress, and urge to neutralize; APA, 2013; Salkovskis, 1985), and with other relevant factors suggested to play a role in OCD (depressed mood and stress; Rachman, 1997; Salkovskis, 1985). We expected the associations with OC symptoms and core aspects of OCD to be at least medium-sized, while the associations with depressed mood and stress to be small to medium in size (since they are not symptoms of OCD but factors that might promote OC symptoms). The main objectives were addressed in three manuscripts – consisting of four independent studies – that constitute this thesis. The manuscripts with their specific hypotheses were as follows:

Manuscript I: *Effects of rumination on unwanted intrusive thoughts: A replication and extension*

- Hypothesis 1: Rumination about a UIT would attenuate the decline of distress, urge to neutralize, depressed mood, and UIT frequency, relative to rumination about negative mood and distraction.

Manuscript II: *Validity of the Counter-App Method in the Assessment of Intrusions*

- Hypothesis 2: The counter-app frequency would be positively associated with the convergent validity criteria, indicating convergent validity for the counter-app method.
- Hypothesis 3: The counter-app frequency would demonstrate smaller associations with the discriminant validity criteria compared to the associations with the convergent validity criteria, indicating discriminant validity for the counter-app method.
- Hypothesis 4: The counter-app frequency would be positively associated with a future criterion, indicating predictive validity for the counter-app method.

Manuscript III: *Exam-related unwanted intrusive thoughts and related neutralizing behaviors: Analogues to obsessions and compulsions*

- Hypothesis 5: The exam-related UITs and related neutralizing behaviors would be positively associated with OC symptoms and core aspects of OCD (anxiety, distress, and urge to neutralize), with medium to large effect sizes.
- Hypothesis 6: The exam-related UITs and related neutralizing behaviors would be positively associated with depressed mood and stress, with lower than medium effect sizes.

I note that this thesis addresses the main hypotheses of the Manuscripts I-III. For details on additional hypotheses and assessed variables, refer to Appendices A to C.

Method

Participants

The manuscripts consist of four independent studies: three experimental studies and one ecological momentary assessment (EMA) study. Participants in each study were undergraduate psychology students recruited at the University of Basel in exchange for course credit; participants in the study of Manuscript III received additionally monetary compensation. We included 105 students in Manuscript I ($M_{\text{age}} = 22.35$ years, $SD = 4.94$; 84.76% female), 77 and 65 in Manuscript II (two studies constituted the manuscript; study 1: $M_{\text{age}} = 22.84$, $SD = 5.82$; 80.52% female; study 2: $M_{\text{age}} = 22.63$, $SD = 4.89$; 73.85% female), and 29 in Manuscript III ($M_{\text{age}} = 21.36$ years, $SD = 1.87$; 92.9% female). Institutional Review Board of the University of Basel approved the study of Manuscript I. Ethics Committee of the Department of Psychology, University of Basel approved studies of Manuscripts II and III.

Research design

Manuscript I. All participants were tested individually in a laboratory. The experiment started with the activation of UIT and was followed by the assessment of distress, urge to neutralize, and depressed mood. The study continued with the assessment of UIT frequency and immediately after that, the participants rated their distress, urge to neutralize, and depressed mood. Subsequently, the participants were randomly assigned to UIT rumination group ($n = 34$), mood rumination group ($n = 35$), or distraction ($n = 36$). The experimental manipulation was followed by ratings of distress, urge to neutralize, and depressed mood. The study continued with the second assessment of UIT frequency and with the final ratings of distress, urge to neutralize, and depressed mood.

Manuscript II. The first study started with the activation of UIT and continued with one of two methods, depending on which experimental group the participants were

randomly assigned to. In the counter-app--thought-sampling group ($n = 38$), participants performed the counter-app method in phase 1 and the thought-sampling method in phase 2. In the thought-sampling--counter-app group ($n = 39$), the order of the methods was reversed. The experimental manipulation was implemented to control for potential order effects. Afterwards, we assessed retrospectively estimated frequency, duration, percentage of time, intensity, and intrusiveness of the UITs, concluding phase 1. Next, participants watched a short movie about visiting different places in Europe to minimize potential carry-over effects. Phase 2 started with the reactivation of UIT (identical to the activation of UIT in phase 1), continued with the counter-app or thought-sampling method, and concluded with the ratings of retrospectively estimated frequency, duration, percentage of time, intensity, and intrusiveness of UITs. The second study was identical to study one, with exception of thinking-aloud method replacing the thought-sampling method. We randomly assigned participants either to the counter-app--thinking-aloud group ($n = 33$) or the thinking-aloud--counter-app group ($n = 32$). In addition, behavioral neutralizing was assessed at the end of the second study.

Manuscript III. The EMA survey took place during the week before a stressful exam period. The study started with a baseline assessment and continued with a 1-week EMA period. In the baseline, we assessed the types of exam-related UITs and related neutralizing behaviors. During the EMA period, participants were prompted three times a day (morning, afternoon, and evening; 21 prompts per participant) on a study smartphone to rate the distress, urge to neutralize, and severity of exam-related UITs and related neutralizing behaviors (severity of ER-UITs-N) as well as anxiety, depressed mood, OC symptoms, and stress.

Measures and tasks

Manuscript I. To activate a UIT, we used a common research practice in the field of OCD research (car accident sentence; for details, see Rachman et al., 1996; van den Hout et al., 2002; van den Hout et al., 2001). We assessed distress, urge to neutralize, and depressed mood with visual analogue scales ranging from 0 (*very low*) to 9 (*very high*). The UIT frequency was assessed with a counter app (FunCoolApps, 2016) installed on a smartphone (counter-app method). The participants were instructed to monitor their thoughts during a 5 min period during which they had to press the volume-up button whenever the UIT entered their mind (adapted from: Marcks & Woods, 2005; Wahl et al., 2019). Participants were randomly assigned to three different experimental groups: UIT rumination, mood rumination, or distraction. The experimental manipulation was implemented to induce three different thinking styles in participants: rumination about a UIT (modified from Morrow & Nolen-Hoeksema, 1990), rumination on negative mood (based on the depressive rumination task; Blagden & Craske, 1996; Huffziger & Kuehner, 2009; Lyubomirsky et al., 2003; Morrow & Nolen-Hoeksema, 1990), or distraction (based on the original distraction task; Huffziger & Kuehner, 2009; Morrow & Nolen-Hoeksema, 1990). The manipulation was identical to the one used in Wahl et al. (2019).

Manuscript II. The activation of UIT was identical to the one in Manuscript I. The UIT frequency was assessed with a retrospective measure (retrospectively estimated frequency; item adapted from Wahl, Lieb, et al., 2020) and with real-time measures: a counter-app method (counter-app frequency), thought-sampling method (thought-sampling frequency), and thinking-aloud method (thinking-aloud frequency). The counter-app method was identical to the one in Manuscript I. During the thought-sampling method, participants were probed 12 times during a 5 min period and with each probe they had to briefly report their thoughts at the time of the probe (adapted

from: Hirsch et al., 2013; Hirsch et al., 2015). In the thinking-aloud method, participants were instructed to verbalize their thoughts during a 5 min period (adapted from: Molina et al., 1998). The verbal reports were recorded, transcribed, and later analyzed. In addition, during each real-time method, participants were asked to focus on their breathing (adapted from: Hirsch et al., 2013; Hirsch et al., 2015; Wahl, Lieb, et al., 2020). We also assessed retrospectively estimated duration, percentage of time, intensity, and intrusiveness of UITs with one item each adapted from Wahl, Lieb, et al. (2020). Each retrospective measure referred to the last 5 min. We used a neutralizing scale consisting of 11 items to assess behavioral neutralizing. The items were modified from previous work on neutralizing (Freeston et al., 1991; Goodman et al., 1989; Kollárik et al., 2020; Purdon & Clark, 1993, 1994a; Rachman et al., 1996) and each item represents a different behavioral neutralizing strategy with answers: 1 (*strategy occurred*) or 0 (*strategy did not occur*).

Manuscript III. Baseline measure: We identified the two most frequent exam-related UITs and two most frequent exam-related neutralizing behaviors for each participant using the Stress-Related Thoughts and Behavior List (Wahl, Hofer, et al., 2020). EMA measures: We assessed the severity of ER-UITs-N with a modified Yale–Brown Obsessive Compulsive Scale (Goodman et al., 1989; German version: Hand & Büttner-Westphal, 1991). When answering, participants had to think about their two most frequent exam-related UITs and the two most frequent related neutralizing behaviors, as identified by the baseline measure. We assessed anxiety and depressed mood with one item each on a scale of 0 (*not at all*) to 100 (*extremely*), distress and urge to neutralize with one item each on a scale of 0 (*very low*) to 5 (*extremely high*), and stress on a scale of 0 (*not at all*) to 5 (*extremely*). OC symptoms were assessed with a modified Obsessive-Compulsive Inventory, Revised (Foa et al., 2002; German

version: Goenner et al., 2007). Each EMA measure referred to the last 30 min when assessing the constructs.

Statistical analysis

Manuscript I. To examine Hypothesis 1, we first calculated the outcome variables distress, urge to neutralize, and depressed mood as the mean difference between ratings that took place directly before and after experimental manipulation. The outcome variable UIT frequency was calculated as the mean difference between the first and second assessment of UIT frequency. We conducted planned contrasts comparing the UIT rumination group to the combined mood rumination and distraction groups on these outcome variables. We conducted an additional contrast to compare the mood rumination group to distraction group (for more details, see Results section below). The significance level was set at $p < .05$.

Manuscripts II. To examine Hypotheses 2, 3, and 4, we calculated Pearson's correlation coefficients between the counter-app frequency and validation criteria. The definitions of convergent, discriminant, and predictive validity that guided this manuscript are based on several studies from the area of psychometrics (DeVon et al., 2007; Drost, 2011; Heale & Twycross, 2015; Post, 2016) and are discussed in detail in Appendix B. Convergent validity of the counter-app method would be demonstrated through positive correlations of the counter-app frequency with convergent validity criteria (thought-sampling frequency, thinking-aloud frequency, and retrospectively estimated frequency of UITs; Hypothesis 2). The discriminant validity of the counter-app method would be demonstrated through correlations of the counter-app frequency with the discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of UITs) that are smaller than the correlations of the counter-app frequency with the convergent validity criteria (Hypothesis 3). The predictive validity of the counter-app method would be shown

through positive correlation between the counter-app frequency and subsequent behavioral neutralizing (Hypothesis 4). In both studies, the correlations did not differ between the experimental groups (e.g., counter-app--thought-sampling vs. thought-sampling--counter-app), indicating that order effects are unlikely. Thus, we conducted the correlation analysis on both groups combined, separately for each study. We report confidence intervals to examine the statistical significance of and differences between the correlations.

Manuscript III. To examine Hypotheses 5 and 6, we calculated separate multilevel models with the severity of ER-UITs-N as the outcome variable and OC symptoms, anxiety, distress, urge to neutralize, depressed mood, and stress as the predictor variables. For each multilevel model, we defined a two-level model with the measurement occasions ($n = 21$ per participant; Level 1) nested within individuals (Level 2). In addition, we included the variable time into each model to control for time trends. Consistent with Barr et al. (2013) and Schmidt-Catran and Fairbrother (2016), we defined the intercept and slope (for each predictor) as random effects. Consistent with Lorah (2018) and Selya et al. (2012), we calculated Cohen's f^2 (Cohen, 1988) as the indicator of effect size for the predictors. The size of f^2 can be interpreted as: 0.02 is a small effect, 0.15 is a medium effect, and 0.35 is a large effect (Cohen, 1988). We interpreted the in-between f^2 values as a range (e.g., a value of 0.05 would be a small to medium effect; values above 0.35 were large effects). The significance level was $p < .05$.

For more details on statistical analysis (e.g., handling of outliers, manipulation checks), see Appendices A to C.

Results

Manuscript I.

Hypothesis 1. Distress and urge to neutralize declined to a smaller degree in the UIT rumination group compared to combined mood rumination and distraction groups (distress: $t(102) = 2.45, p = .016$; urge to neutralize: $t(81.58) = 2.74, p = .007$). The rumination about the UIT did not differ from rumination about negative mood and distraction in its effect on depressed mood ($t(102) = 1.56, p = .121$) and UIT frequency ($t(102) = 0.45, p = .656$). For means and standard deviations, see Appendix A.

Visual inspection of the means suggests that rumination about negative mood was similar to rumination about the UIT in its effect on the outcome variables. Thus, we conducted an additional analysis and compared rumination about negative mood to distraction. In contrast to distraction, rumination about negative mood attenuated the decrease of distress ($t(102) = 2.15, p = .034$), urge to neutralize ($t(60.84) = 3.42, p = .001$), and depressed mood ($t(102) = 3.54, p = .001$).

Manuscript II.

Hypothesis 2, 3, and 4. Table 1 depicts correlations of the counter-app frequency with the convergent, discriminant, and predictive validity criteria. The counter-app frequency was positively correlated with the convergent validity criteria (thought-sampling frequency, thinking-aloud frequency, and retrospectively estimated frequency of UITs). The associations between the counter-app frequency and discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of UITs) were not smaller than those between the counter-app frequency and convergent validity criteria. The counter-app frequency was positively correlated with the predictive validity criterion (behavioral neutralizing).

Table 1. Convergent, discriminant, and predictive validity of the counter-app method
(Pearson correlations) separately for Study 1 and 2

Variable	Study 1		Study 2	
	Counter-app frequency		Counter-app frequency	
	<i>r</i>	95% CI	<i>r</i>	95% CI
Convergent validity				
Thought-sampling frequency	.38	[.171, .557]	-	-
Thinking-aloud frequency ^a	-	-	.25	[.002, .469]
Retrospectively estimated frequency	.32	[.103, .508]	.64	[.469, .765]
Discriminant validity				
Retrospectively estimated				
Duration	.18	[-.046, .388]	.43	[.208, .610]
Percentage of time	.22	[-.004, .423]	.60	[.417, .736]
Intensity	.21	[-.015, .414]	.52	[.316, .678]
Intrusiveness	.31	[.092, .499]	.56	[.366, .707]
Predictive validity				
Behavioral neutralizing	-	-	.27	[.028, .482]

Note. CI = confidence interval.

^a Two participants had missing values.

Manuscript III.

Hypothesis 5 and 6. Table 2 displays the main results. The severity of ER-UITs-N was positively associated with OC symptoms and core aspects of OCD (anxiety, distress, and urge to neutralize). The effect sizes were small to medium for the associations with OC symptoms and anxiety, and large for the associations with distress and urge to neutralize. The severity of ER-UITs-N was positively associated with stress but not with depressed mood. The effect size for the association with stress was small to medium.

Table 2. Associations of the severity of ER-UITs-N (outcome) with OC symptoms, core aspects of OCD, depressed mood, and stress (predictors)

Predictors (fixed effects)	Severity of ER-UITs-N		
	Coefficient (SE)	<i>p</i>	<i>f</i> ²
OC symptoms	0.46 (0.13)	.004	.06
Anxiety	0.07 (0.02)	.002	.14
Distress	2.73 (0.34)	< .001	.58
Urge to neutralize	2.68 (0.34)	< .001	.50
Depressed mood	0.04 (0.02)	.052	.03
Stress	1.17 (0.37)	.004	.12

Note. Each fixed effect represents a separate multilevel model with severity of ER-UITs-N as the outcome. ER-UITs-N = exam-related unwanted intrusive thoughts and related neutralizing behaviors; OC = obsessive-compulsive.

Discussion

The present thesis addressed three relevant areas related to research on UITs. First, we investigated the effect of rumination about a UIT on OC symptoms, second, we evaluated the validity of the counter-app method in assessment of UIT frequency, and finally, we examined the analogy of exam-related UITs and related neutralizing behaviors to obsessions and compulsions.

Consistent with Wahl et al. (2019), rumination about the UIT attenuated the decrease of urge to neutralize compared to rumination about negative mood and distraction. Additionally and in contrast to Wahl et al. (2019), rumination about the UIT also attenuated the decrease of distress in comparison to the other two thinking styles. This suggests that ruminating about one's UITs might prevent the decline of urge to neutralize and distress. Rumination about the UIT did not differ from rumination about negative mood and distraction in its effect on depressed mood and UIT frequency. Interestingly, rumination about negative mood also attenuated the decrease of distress, urge to neutralize, and depressed mood in contrast to distraction. This indicates that two different types of rumination had comparable effects on the outcome variables. In brief, the Hypothesis 1 was partially supported, and our findings suggest

that rumination, irrespective of UIT or negative mood focus, contributes to the persistence of distress and urge to neutralize caused by UITs.

The findings were also recently replicated and extended by our research group in patients diagnosed with OCD (Wahl, van den Hout, et al., in press). The authors showed that rumination about an obsessive thought and rumination about negative mood led to a reduced decline of distress, urge to neutralize, depressed mood, and obsession frequency, relative to distraction. Based on the findings of this and previous studies, it might be that both types of rumination co-occur in everyday life and contribute to persistence of dysphoric feelings (e.g., distress, urge to neutralize, and depressed mood) caused by the occurrence of UIT or obsession. The effect of rumination seems to exert to distress, urge to neutralize, and depressed mood but not to frequency of UITs in individuals with no mental disorder. In individuals diagnosed with OCD, the rumination seems to also influence the frequency of obsessions.

The second aim was to evaluate the validity of the counter-app method in assessment of UIT frequency. The counter-app frequency was positively associated with the convergent validity criteria (thought-sampling frequency, thinking-aloud frequency, and retrospectively estimated frequencies of UITs), demonstrating convergent validity for the counter-app method. The associations of the counter-app frequency with the discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of UITs) were not smaller than the associations with the convergent validity criteria. Findings replicate and extend Wahl, Lieb, et al. (2020), who also reported results supporting convergent, but not discriminant, validity of the counter-app method. With regard to predictive validity, the counter-app frequency was positively associated with the subsequent behavioral neutralizing, supporting predictive validity of the counter-app method. In Wahl, Lieb, et al. (2020), the findings did not support the predictive validity of the counter-app method.

In sum, results support Hypothesis 2 and 4, but not Hypothesis 3. Our results suggest that the counter-app method can assess the frequency of UITs in a valid way but is less applicable to research that aims to differentiate frequency from other characteristics of UITs such as duration.

Finally, we investigated the analogy of exam-related UITs and related neutralizing behaviors to obsessions and compulsions. To evaluate the analogy, we examined the associations of exam-related UITs and related neutralizing behaviors with OC symptoms, core aspects of OCD (anxiety, distress, and urge to neutralize), and other relevant factors thought to play a role in OCD (depressed mood and stress). In general, the associations support the analogy of exam-related UITs and related neutralizing behaviors to obsessions and compulsions in some aspects but not all. Specifically, the severity of ER-UITs-N was positively associated with OC symptoms, anxiety, distress, urge to neutralize, and stress, but was not related to depressed mood. The associations with OC symptoms and anxiety were small to medium in their effect size, that is, smaller than expected. As predicted, the associations with distress and urge to neutralize were medium to large and the association with stress was small to medium. Thus, the Hypotheses 5 and 6 were partially supported. The relatively small association between the severity of ER-UITs-N and OC symptoms could be attributed to the questionnaire (Obsessive-Compulsive Inventory, Revised) that we used to assess OC symptoms. A common type of exam-related UITs reported by the students were superstitious thoughts (for more details, see Appendix C) and Obsessive-Compulsive Inventory, Revised does not assess superstition. Superstitious thoughts are also common in OCD (Garcia-Soriano et al., 2011; Katerberg et al., 2010). Future studies might want to evaluate the analogy also with other measures of OC symptoms. For example, the Obsessional Intrusive Thoughts Inventory (Garcia-Soriano et al., 2011) assesses also superstition.

Strengths and limitations

Manuscript I and II have a number of strengths. First of all, all three studies are replication-extension studies. Bonett (2012) argued that replication research is rare in psychology and emphasized the need of replication-extension studies to improve the quality of published psychological research. Replication-extension designs allow more precise estimation of effect sizes and increase the generalizability of statistical results to other populations (Bonett, 2012). Next, in Manuscript II, we included real-time measures to validate the counter-app method in assessment of UITs. Relative to retrospective measures, real-time measures minimize retrospective biases (Trull & Ebner-Priemer, 2013; Trull & Ebner-Priemer, 2020). A strength of Manuscript III is that we used the EMA methodology to investigate the exam-related UITs and related neutralizing behaviors. EMA captures experiences in real- or near-real time, minimizes retrospective biases, and can assess various psychological constructs in their natural environment (Trull & Ebner-Priemer, 2020). Thus, the ecological validity of our measure exceeds that of more traditional measures which are conducted, for example, in laboratory.

In Manuscript I and II, there are several limitations to acknowledge. First, in both manuscripts, we used single-item measures to assess some of the dependent variables (e.g., distress in Manuscript I; retrospectively estimated frequency in Manuscript II). This raises questions about the psychometric properties of the measures. Next, since the findings were replicated by the same research group, we cannot rule out the possibility that the researchers' expectations might have affected the present findings. In addition, we did not account for participants' reactivity to the counter-app method. Reactivity occurs when the subject is affected by the methodology of the study (Lavrakas, 2008). Evidence indicates that participants might overestimate the frequency of UITs with the counter-app method (Wahl, Lieb, et al.,

2020). Similar to Manuscript I and II, a limitation of the third manuscript is that we did not control for participants' reactivity to the EMA methodology. In addition, we also used single-item measures to assess some of the relevant variables (e.g., anxiety, depressed mood) and the psychometric properties of these measures were not explored. Next, we did not include a sample consisting of individuals diagnosed with OCD and thus, did not directly compare exam-related UITs and related neutralizing behaviors to obsessions and compulsions. Finally, results of all three manuscripts might not generalize beyond our student samples to the population.

Implications

Manuscript I replicates and extends previous findings (Wahl et al., 2019) and strengthens the argument that rumination plays a role in OCD. This was also supported by the recent study that investigated rumination in patients diagnosed with OCD (Wahl, van den Hout, et al., in press). Taken together, the findings have an important implication for practice: Treatments of OCD, such as cognitive-behavioral therapy, might want to consider targeting and reducing rumination in patients.

Researchers have questioned the validity of the counter-app method in assessment of UIT frequency (Kollárik et al., 2020; Wahl et al., 2019). Manuscript II provides evidence that the counter-app method is a valid assessment tool of UITs. Based on our results, we recommend using the counter-app method in experimental studies to assess the real-time frequency of UITs in short time intervals (5 min). The method might be less appropriate for research that aims to differentiate frequency from other constructs, such as duration or intensity. To account for the possible overreporting shown in Wahl, Lieb, et al. (2020), we additionally suggest that the counter-app frequency should not be interpreted as absolute frequency of UITs but rather as estimate of the construct.

Finally, we recommend considering exam-related UITs and related neutralizing behaviors in the field of OCD research as analogues to clinical obsessions and compulsions. This will provide researchers with the opportunity to study relatively frequent, idiosyncratic UITs and related neutralizing behaviors in their natural setting. This might extend studies during which participants report UITs that occur with a low frequency as well as experimental studies that temporarily activate UITs that are not necessarily idiosyncratic. During the EMA period of the study, the idiosyncratic exam-related UITs occurred on average 7 times and the exam-related neutralizing behaviors on average 6 times (for more details, see Appendix C).

Conclusion

The present thesis contributes to research related to OCD and provides important implications for practice and research. In Manuscript I, we replicated and extended previous results on the effects of rumination about a UIT on OC symptoms. Results suggest that rumination might play a role in persistence of UITs. In Manuscript II, we showed that the counter-app method has convergent and predictive, but not discriminant, validity when assessing the frequency of UITs. In Manuscript III, we evaluated the analogy of exam-related UITs and related neutralizing behaviors to clinical obsessions and compulsions and recommend considering the exam-related UITs and related neutralizing behaviors if researchers intend to investigate idiosyncratic UITs that occur frequently in their natural environment.

References

- Abramowitz, J. S., & Jacoby, R. J. (2014). Obsessive-compulsive and related disorders. In D. H. Barlow (Ed.), *The Oxford handbook of clinical psychology* (pp. 313-338). Oxford University Press.
- Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: a meta-analysis of controlled studies. *Clinical Psychology Review*, 21(5), 683-703. [https://doi.org/10.1016/s0272-7358\(00\)00057-x](https://doi.org/10.1016/s0272-7358(00)00057-x)
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Belloch, A., Morillo, C., Lucero, M., Cabedo, E., & Carrio, C. (2004). Intrusive thoughts in non-clinical subjects: The role of frequency and unpleasantness on appraisal ratings and control strategies. *Clinical Psychology & Psychotherapy*, 11(2), 100-110. <https://doi.org/10.1002/cpp.397>
- Blagden, J. C., & Craske, M. G. (1996). Effects of active and passive rumination and distraction: A pilot replication with anxious mood. *Journal of Anxiety Disorders*, 10(4), 243-252. [https://doi.org/Doi.10.1016/0887-6185\(96\)00009-6](https://doi.org/Doi.10.1016/0887-6185(96)00009-6)
- Bonett, D. G. (2012). Replication-Extension Studies. *Current Directions in Psychological Science*, 21(6), 409-412. <https://doi.org/10.1177/0963721412459512>
- Bouvard, M., Fournet, N., Denis, A., Sixdenier, A., & Clark, D. (2017). Intrusive thoughts in patients with obsessive compulsive disorder and non-clinical participants: a comparison using the International Intrusive Thought Interview Schedule. *Cognitive Behaviour Therapy*, 46(4), 287-299. <https://doi.org/10.1080/16506073.2016.1262894>
- Clark, D. A., & Inozu, M. (2014). Unwanted intrusive thoughts: Cultural, contextual, covariational, and characterological determinants of diversity. *Journal of Obsessive-Compulsive and Related Disorders*, 3(2), 195-204. <https://doi.org/10.1016/j.jocrd.2014.02.002>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Routledge.
- Dar, K. A., & Iqbal, N. (2015). Worry and Rumination in Generalized Anxiety Disorder and Obsessive Compulsive Disorder. *Journal of Psychology*, 149(8), 866-880. <https://doi.org/10.1080/00223980.2014.986430>
- Davison, G. C., Navarre, S. G., & Vogel, R. S. (1995). The Articulated Thoughts in Simulated Situations Paradigm - a Think-Aloud Approach to Cognitive Assessment. *Current Directions in Psychological Science*, 4(1), 29-33. <https://doi.org/10.1111/1467-8721.ep10770963>
- De Putter, L. M., Van Yper, L., & Koster, E. H. (2017). Obsessions and compulsions in the lab: A meta-analysis of procedures to induce symptoms of obsessive-compulsive disorder. *Clinical Psychology Review*, 52, 137-147. <https://doi.org/10.1016/j.cpr.2017.01.001>
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, 39(2), 155-164. <https://doi.org/10.1111/j.1547-5069.2007.00161.x>
- Drost, E. (2011). Validity and Reliability in Social Science Research. . *Education research and perspectives*, 38, 105-123.

- Faber, M., Bixler, R., & D'Mello, S. K. (2018). An automated behavioral measure of mind wandering during computerized reading. *Behavior Research Methods*, 50(1), 134-150. <https://doi.org/10.3758/s13428-017-0857-y>
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485-496. <https://doi.org/10.1037//1040-3590.14.4.485>
- Freeston, M. H., Ladouceur, R., Thibodeau, N., & Gagnon, F. (1991). Cognitive intrusions in a non-clinical population. I. Response style, subjective experience, and appraisal. *Behaviour Research and Therapy*, 29(6), 585-597. <https://www.ncbi.nlm.nih.gov/pubmed/1759957>
- FunCoolApps. (2016). *Click Counter* (Version 1.4) [Mobile app]. Google Play Store. <https://play.google.com/store/apps/details?id=com.useless.counter>
- Gagné, J. P., Kelly-Turner, K., & Radomsky, A. S. (2018). From the laboratory to the clinic (and back again): How experiments have informed cognitive-behavior therapy for obsessive-compulsive disorder. *Journal of Experimental Psychopathology*, 9(4). <https://doi.org/10.1177/2043808718810030>
- Garcia-Soriano, G., Belloch, A., Morillo, C., & Clark, D. A. (2011). Symptom dimensions in obsessive-compulsive disorder: from normal cognitive intrusions to clinical obsessions. *Journal of Anxiety Disorders*, 25(4), 474-482. <https://doi.org/10.1016/j.janxdis.2010.11.012>
- Gönner, S., Leonhart, R., & Ecker, W. (2007). Das Zwangsinventar OCI-R - die Deutsche Version des Obsessive-Compulsive Inventory-Revised: Ein kurzes Selbstbeurteilungsinstrument zur mehrdimensionalen Messung von Zwangssymptomen [The German version of the Obsessive-Compulsive Inventory-Revised: A brief self-report measure for the multidimensional assessment of obsessive-compulsive symptoms]. *Psychotherapie Psychosomatik Medizinische Psychologie*, 57(9-10), 395-404. <https://doi.org/10.1055/s-2007-970894>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry*, 46(11), 1006-1011. <https://doi.org/10.1001/archpsyc.1989.01810110048007>
- Grisham, J. R., & Williams, A. D. (2009). Cognitive control of obsessional thoughts. *Behaviour Research and Therapy*, 47(5), 395-402. <https://doi.org/10.1016/j.brat.2009.01.014>
- Hand, I., & Büttner-Westphal, H. (1991). Die Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Ein halbstrukturiertes Interview zur Beurteilung des Schweregrades von Denk- und Handlungszwängen [The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS): A semistructured interview for assessing severity of compulsive cognitions and behavior]. *Verhaltenstherapie*, 1, 223-225. <https://doi.org/10.1159/000257972>
- Harvey, A., Watkins, E., Mansell, W., & Shafran, R. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press. <https://doi.org/10.1093/med:psych/978019528883.001.0001>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66-67. <https://doi.org/10.1136/eb-2015-102129>
- Heinzel, C. V., Kollárik, M., Miché, M., Clamor, A., Ertle, A., Lieb, R., & Wahl, K. (2021). *Is a ruminative thinking style related to obsessive-compulsive symptom severity*

- beyond its associations with depressive and anxiety symptom severity [Manuscript submitted for publication]? Division of Clinical Psychology and Epidemiology, University of Basel.
- Hirsch, C. R., Mathews, A., Lequertier, B., Perman, G., & Hayes, S. (2013). Characteristics of worry in Generalized Anxiety Disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 44(4), 388-395. <https://doi.org/10.1016/j.ibtep.2013.03.004>
- Hirsch, C. R., Perman, G., Hayes, S., Eagleson, C., & Mathews, A. (2015). Delineating the Role of Negative Verbal Thinking in Promoting Worry, Perceived Threat, and Anxiety. *Clinical Psychological Science*, 3(4), 637-647. <https://doi.org/10.1177/2167702615577349>
- Huffziger, S., & Kuehner, C. (2009). Rumination, distraction, and mindful self-focus in depressed patients. *Behaviour Research and Therapy*, 47(3), 224-230. <https://doi.org/10.1016/j.brat.2008.12.005>
- Julien, D., O'Connor, K. P., & Aardema, F. (2007). Intrusive thoughts, obsessions, and appraisals in obsessive-compulsive disorder: A critical review. *Clinical Psychology Review*, 27(3), 366-383. <https://doi.org/10.1016/j.cpr.2006.12.004>
- Katerberg, H., Delucchi, K. L., Stewart, S. E., Lochner, C., Denys, D. A. J. P., Stack, D. E., Andresen, J. M., Grant, J. E., Kim, S. W., Williams, K. A., den Boer, J. A., van Balkom, A. J. L. M., Smit, J. H., van Oppen, P., Polman, A., Jenike, M. A., Stein, D. J., Mathews, C. A., & Cath, D. C. (2010). Symptom dimensions in OCD: Item-level factor analysis and heritability estimates. *Behavior Genetics*, 40(4), 505-517. <https://doi.org/10.1007/s10519-010-9339-z>
- Keinan, G. (2002). The effects of stress and desire for control on superstitious behavior. *Personality and Social Psychology Bulletin*, 28(1), 102-108. <https://doi.org/10.1177/0146167202281009>
- Kollárik, M., van den Hout, M., Heinzl, C. V., Hofer, P. D., Lieb, R., & Wahl, K. (2020). Effects of rumination on unwanted intrusive thoughts: A replication and extension. *Journal of Experimental Psychopathology*, 11(1). <https://doi.org/10.1177/2043808720912583>
- Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. <https://dx.doi.org/10.4135/9781412963947.n448>
- Lepore, S. J. (1997). Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *Journal of Personality and Social Psychology*, 73(5), 1030-1037. <https://doi.org/10.1037/0022-3514.73.5.1030>
- Lorah, J. (2018). Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, 6. <https://doi.org/10.1186/s40536-018-0061-2>
- Lyubomirsky, S., Kasri, F., & Zehm, K. (2003). Dysphoric rumination impairs concentration on academic tasks. *Cognitive Therapy and Research*, 27(3), 309-330. <https://doi.org/10.1023/A:1023918517378>
- Marcks, B. A., & Woods, D. W. (2005). A comparison of thought suppression to an acceptance-based technique in the management of personal intrusive thoughts: a controlled evaluation. *Behaviour Research and Therapy*, 43(4), 433-445. <https://doi.org/10.1016/j.brat.2004.03.005>
- McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radomsky, A., Sookman, D., Taylor, S., & Wilhelm, S. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. *Clinical Psychology Review*, 24(3), 283-313. <https://doi.org/10.1016/j.cpr.2004.04.003>

- Molina, S., Borkovec, T. D., Peasley, C., & Person, D. (1998). Content analysis of worrisome streams of consciousness in anxious and dysphoric participants. *Cognitive Therapy and Research*, 22(2), 109-123. <https://doi.org/10.1023/A:1018772104808>
- Morillo, C., Belloch, A., & Garcia-Soriano, G. (2007). Clinical obsessions in obsessive-compulsive patients and obsession-relevant intrusive thoughts in non-clinical, depressed and anxious subjects: Where are the differences? *Behaviour Research and Therapy*, 45(6), 1319-1333. <https://doi.org/10.1016/j.brat.2006.11.005>
- Morrow, J., & Nolen-Hoeksema, S. (1990). Effects of Responses to Depression on the Remediation of Depressive Affect. *Journal of Personality and Social Psychology*, 58(3), 519-527. <https://doi.org/Doi 10.1037/0022-3514.58.3.519>
- Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, 100(4), 569-582. <https://www.ncbi.nlm.nih.gov/pubmed/1757671>
- Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science*, 3(5), 400-424. <https://doi.org/10.1111/j.1745-6924.2008.00088.x>
- Post, M. W. (2016). What to Do With "Moderate" Reliability and Validity Coefficients? *Archives of Physical Medicine and Rehabilitation*, 97(7), 1051-1052. <https://doi.org/10.1016/j.apmr.2016.04.001>
- Purdon, C., & Clark, D. A. (1993). Obsessive intrusive thoughts in nonclinical subjects. Part I. Content and relation with depressive, anxious and obsessional symptoms. *Behaviour Research and Therapy*, 31(8), 713-720. [https://doi.org/10.1016/0005-7967\(93\)90001-B](https://doi.org/10.1016/0005-7967(93)90001-B)
- Purdon, C., & Clark, D. A. (1994a). Obsessive intrusive thoughts in nonclinical subjects. Part II. Cognitive appraisal, emotional response and thought control strategies. *Behaviour Research and Therapy*, 32(4), 403-410. <https://www.ncbi.nlm.nih.gov/pubmed/8192639>
- Purdon, C., & Clark, D. A. (1994b). Perceived control and appraisal of obsessional intrusive thoughts: A replication and extension. *Behavioural and Cognitive Psychotherapy*, 22, 269-285.
- Rachman, S. (1997). A cognitive theory of obsessions. *Behaviour Research and Therapy*, 35(9), 793-802. [https://doi.org/10.1016/S0005-7967\(97\)00040-5](https://doi.org/10.1016/S0005-7967(97)00040-5)
- Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy*, 16(4), 233-248. [https://doi.org/10.1016/0005-7967\(78\)90022-0](https://doi.org/10.1016/0005-7967(78)90022-0)
- Rachman, S., Shafran, R., Mitchell, D., Trant, J., & Teachman, B. (1996). How to remain neutral: An experimental analysis of neutralization. *Behaviour Research and Therapy*, 34(11-12), 889-898. [https://doi.org/10.1016/S0005-7967\(96\)00051-4](https://doi.org/10.1016/S0005-7967(96)00051-4)
- Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., Clark, D. A., Coles, M. E., Doron, G., Fernández-Álvarez, H., Garcia-Soriano, G., Ghisi, M., Gomez, B., Inozu, M., Moulding, R., Shams, G., Sica, C., Simos, G., & Wong, W. (2014). Part 1—You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders*, 3(3), 269-279. <https://doi.org/10.1016/j.jocrd.2013.09.002>
- Salkovskis, P. M. (1985). Obsessional-compulsive problems - a cognitive-behavioral analysis. *Behaviour Research and Therapy*, 23(5), 571-583. [https://doi.org/10.1016/0005-7967\(85\)90105-6](https://doi.org/10.1016/0005-7967(85)90105-6)

- Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety Stress Coping*, 24(4), 359-367. <https://doi.org/10.1080/10615806.2010.543132>
- Schmidt-Catran, A. W., & Fairbrother, M. (2016). The random effects in multilevel models: Getting them wrong and getting them right. *European Sociological Review*, 32(1), 23-38. <https://doi.org/10.1093/esr/jcv090>
- Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2 , a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00111>
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66, 487-518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods*, 52(1), 162-176. <https://doi.org/10.3758/s13428-019-01214-4>
- Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, 9, 151-176. <https://doi.org/10.1146/annurev-clinpsy-050212-185510>
- Trull, T. J., & Ebner-Priemer, U. W. (2020). Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. *Journal of Abnormal Psychology*, 129(1), 56-63. <https://doi.org/10.1037/abn0000473>
- van den Hout, M., Kindt, M., Weiland, T., & Peters, M. (2002). Instructed neutralization, spontaneous neutralization and prevented neutralization after an obsession-like thought. *Journal of Behavior Therapy and Experimental Psychiatry*, 33(3-4), 177-189. [https://doi.org/10.1016/S0005-7916\(02\)00048-4](https://doi.org/10.1016/S0005-7916(02)00048-4)
- van den Hout, M., van Pol, M., & Peters, M. (2001). On becoming neutral: effects of experimental neutralizing reconsidered. *Behaviour Research and Therapy*, 39(12), 1439-1448. [https://doi.org/10.1016/S0005-7967\(00\)00109-1](https://doi.org/10.1016/S0005-7967(00)00109-1)
- Wahl, K., Ertle, A., Bohne, A., Zurovski, B., & Kordon, A. (2011). Relations between a ruminative thinking style and obsessive-compulsive symptoms in non-clinical samples. *Anxiety Stress and Coping*, 24(2), 217-225. <https://doi.org/10.1080/10615806.2010.482985>
- Wahl, K., Hofer, P. D., Meyer, A. H., & Lieb, R. (2020). Prior beliefs about the importance and control of thoughts are predictive but not specific to subsequent intrusive unwanted thoughts and neutralizing behaviors. *Cognitive Therapy and Research*, 44, 360-375. <https://doi.org/10.1007/s10608-019-10046-7>
- Wahl, K., Lieb, R., Kollarik, M., & Purdon, C. (2020). The appropriateness of using a counter app in experimental studies assessing unwanted intrusive thoughts. *Journal of Obsessive-Compulsive and Related Disorders*, 26. <https://doi.org/10.1016/j.jocrd.2020.100550>
- Wahl, K., van den Hout, M., Heinzl, C. V., Kollarik, M., Meyer, A., Benoy, C., Berberich, G., Domschke, K., Gloster, A., Gradwohl, G., Hofecker, M., Jähne, A., Koch, S., Külz, A. K., Moggi, F., Poppe, C., Riedel, A., Rufer, M., Stierle, C., ... Lieb, R. (in press). Rumination about obsessive symptoms and mood maintains obsessive-compulsive symptoms and depressed mood: An experimental study. *Journal of Abnormal Psychology*.
- Wahl, K., van den Hout, M., & Lieb, R. (2019). Rumination on unwanted intrusive thoughts affects the urge to neutralize in nonclinical individuals. *Journal of*

Obsessive-Compulsive and Related Disorders, 20, 4-12.
<https://doi.org/10.1016/j.jocrd.2018.02.002>

Zanov, M. V., & Davison, G. C. (2010). A Conceptual and Empirical Review of 25 Years of Cognitive Assessment Using the Articulated Thoughts in Simulated Situations (ATSS) Think-Aloud Paradigm. *Cognitive Therapy and Research*, 34(3), 282-291. <https://doi.org/10.1007/s10608-009-9271-9>

Appendix A (Manuscript I)

Effects of rumination on unwanted intrusive thoughts: A replication and extension

Martin Kollárik, Marcel van den Hout, Carlotta V. Heinzl, Patrizia D. Hofer, Roselind Lieb, Karina Wahl

published in *Journal of Experimental Psychopathology*

Effects of rumination on unwanted intrusive thoughts: A replication and extension

Journal of Experimental Psychopathology
January-March 2020: 1–16
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2043808720912583
journals.sagepub.com/home/jepp



Martin Kollárik 

University of Basel, Switzerland

Marcel van den Hout

Utrecht University, the Netherlands

Carlotta V. Heinzl 

University of Basel, Switzerland

Patrizia D. Hofer

University of Basel, Switzerland

Roselind Lieb

University of Basel, Switzerland

Karina Wahl 

University of Basel, Switzerland

Abstract

Studies indicate that rumination might play a role in obsessive–compulsive disorder. In a previous experimental study, rumination about an unwanted intrusive thought (UIT) maintained the urge to neutralize this thought. We sought to replicate and extend these findings with measures of behavioral and mental neutralizing. Additionally, we investigated possible mechanisms that might be involved in the effects of rumination on the UIT. We activated a UIT by asking students ($N = 105$) to write down a sentence stating that they wished a loved person would die in a car accident. Participants were randomly allocated to rumination about the UIT, rumination about negative mood, or distraction. As predicted, rumination about the UIT maintained the urge to neutralize the UIT, relative to rumination about negative mood and distraction. In addition, rumination about the UIT also maintained distress associated with the UIT compared to rumination about negative mood and distraction. The effects of rumination did not extend to behavioral or mental neutralizing. UIT frequency and vividness were unaffected by rumination. The present findings strengthen the confidence that rumination contributes to the maintenance of UITs.

Corresponding author:

Karina Wahl, Division of Clinical Psychology and Epidemiology, Department of Psychology, University of Basel, Missionsstrasse 60/62, 4055 Basel, Switzerland.

Email: karina.wahl@unibas.ch



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Keywords

Experimental study, extension, obsessive–compulsive disorder, replication, rumination, unwanted intrusive thoughts

Date received: 23 June 2019; accepted: 4 February 2020

Introduction

Individuals with obsessive–compulsive disorder (OCD) typically experience distressing obsessive thoughts (Rachman, 1997, 1998; Salkovskis, 1985), defined as “recurrent and persistent thoughts, urges, or images that are experienced, at some time during the disturbance, as intrusive and unwanted, and that in most individuals cause marked anxiety or distress” (American Psychiatric Association, 2013, p. 237). Once an obsession enters the mind, there are multiple behavioral and mental strategies an individual with OCD can use to respond (Freeston & Ladouceur, 1997). Analyzing the thought, its reasons, and causes has been identified as one of these strategies (Freeston & Ladouceur, 1997). This indicates that obsessive thoughts might be followed by rumination about the obsessive thoughts. For example, individuals diagnosed with OCD might ruminate about why they cannot get rid of their obsessive thoughts, what reasons caused the obsessive thoughts in the first place, and what might happen if the obsessive thoughts persist.

Rumination is defined as passive repetitive thinking about symptoms of distress and its causes, meanings, and consequences (Nolen-Hoeksema, 1991). Although originally researched in the area of depression (Nolen-Hoeksema, 1991; Nolen-Hoeksema et al., 2008), rumination has been associated with a number of mental disorders, such as anxiety (Harrington & Blankenship, 2002), alexithymia (Di Schiena et al., 2011), eating disorders (Naumann et al., 2015), and sleeping disorders (Carney et al., 2013). In addition, Nolen-Hoeksema et al. (2007) demonstrated that rumination might convey a risk of the onset of psychopathology. In their study, prior rumination increased the risk of subsequent onset of major depression, recurrent binge eating, and substance abuse over a 3-year period.

A number of studies with nonclinical individuals have demonstrated associations between rumination and obsessive–compulsive (OC) symptoms. For instance, Grisham and Williams (2009) and Wahl, Ertle et al. (2011) showed that the tendency to ruminate was positively correlated with OC symptoms in

student samples. Studies with clinical samples indicated that ruminative thoughts occur as frequently as obsessive thoughts in individuals diagnosed with OCD (Wahl, Schönfeld et al., 2011). Dar and Iqbal (2015) found positive correlations between rumination and OC symptoms in a mixed sample of individuals diagnosed with OCD or generalized anxiety disorder. Rumination was related to the unacceptable thoughts/neutralization domain of OCD when negative affect was controlled for in an unselected treatment-seeking sample (Raines et al., 2017). Overall, these studies indicate an association between rumination and OCD symptoms and raise the question of whether and in what way they might influence each other.

Wahl et al. (2019) investigated the immediate rumination effects on the distress associated with an unwanted intrusive thought (UIT), the urge to neutralize it, depressed mood, and the frequency of this thought in an experimental study. Rumination about the UIT led to an *attenuated* decrease of the urge to neutralize the UIT in comparison to rumination about negative mood and distraction. This means that rumination might contribute to the maintenance of intrusive thoughts by reducing the natural decrease of the urge to act upon them. The authors did not find an effect of rumination about the UIT on the distress, depressed mood, or frequency of the UIT.

Several mechanisms by which rumination might influence the persistence of intrusive thoughts have been discussed. Raines et al. (2017) suggested that rumination might promote the misinterpretation of naturally occurring intrusive thoughts by changing the appraisals of these thoughts. This idea was supported in the previous experimental study (Wahl et al., 2019). Rumination about the UIT resulted in a stronger belief that the thought might come true relative to rumination about negative mood and distraction. This suggests that ruminating about one’s UITs might strengthen dysfunctional appraisals of the UIT. In addition, rumination could result in the persistence of negative mood. Negative mood increases the frequency of intrusive thoughts (Clark, 2002; Reynolds & Salkovskis, 1991) resulting in a vicious circle of

negative mood and intrusive thoughts (Wahl, Schönfeld, et al., 2011). Moreover, Grisham and Williams (2009) suggested that rumination might fuel the frequency of intrusive thoughts by increasing their accessibility as a result of a quicker spread of activation in the semantic network. Previous research has demonstrated a clear association between rumination and frequency of intrusive cognitions such as intrusive memories (James et al., 2016; Luo et al., 2013; White & Wild, 2016). For example, White and Wild (2016) showed that individuals who were trained to adopt an abstract mode of processing—such as rumination—in response to a traumatic film reported more intrusive memories than individuals who were trained to adopt a concrete processing style. Abstract processing was defined as rumination focused on the reasons, meanings, and consequences of the traumatic event and concrete processing was characterized by focusing on the concrete details of the event. In another study, rumination clearly correlated positively with intrusion frequency (Zetsche et al., 2009). Taken together, these results lead us to expect rumination about a UIT to affect not just the urge to neutralize the UIT but also the associated distress, depressed mood, frequency, and negative appraisals of the thought.

Two further candidates that might be involved in the effects of rumination on UITs are trait rumination and thought–action fusion (TAF). Trait rumination refers to the tendency to ruminate in daily life (Just & Alloy, 1997; Nolen-Hoeksema, 1991, 2000). Studies indicate that higher levels of trait rumination are associated with greater negative emotional outcomes (Moberly & Watkins, 2006; Watkins, 2004). TAF is the belief that experiencing an unacceptable thought is morally equivalent to acting according to the thought or that mere thinking about a particular event makes it more likely to happen (Rachman, 1997, 1998; Rassin et al., 2001; Salkovskis, 1985; Shafran et al., 1996). The effects of rumination on UITs might be particularly pronounced for individuals who have a strong tendency to ruminate or who strongly endorse beliefs about TAF. In this way, trait rumination and TAF might moderate the effects of rumination on UITs.

Finally, rumination might change the vividness of a UIT. In one study, a majority (81%) of individuals with OCD reported having mental images (Speckens et al., 2007). These images were mainly visual and were experienced as distressing and vivid (Lipton et al., 2010). Intrusive images in OCD seem to occur

more frequently than in anxiety disorders, are less often associated with past memories, and are typically viewed from a person's own vantage point rather than from an observer's perspective (Lipton et al., 2010). Rumination has been identified as one of the main triggers of intrusive images (Birrer et al., 2007), and one can assume that rumination affects the vividness of UITs (Birrer et al., 2007; Pearson et al., 2013).

The present study seeks to further clarify the imminent effect of rumination on UITs by replicating and extending previous findings by Wahl et al. (2019). The first aim was to replicate their main finding (rumination on a UIT attenuates the decline in the urge to neutralize) using an identical experimental paradigm in an independent sample. The second aim was to extend these findings by including measures of *actual* behavioral and mental neutralizing. While Wahl et al. (2019) investigated the urge to neutralize, an even stronger indication of the effects of rumination on UITs would be changes in *actual* neutralization. Additionally, we investigated several possible mechanisms that might be related to the influences of rumination on UITs.

We hypothesized that, relative to distraction and rumination about negative mood, rumination about a UIT would attenuate the natural waning of the urge to neutralize the UIT (replication of previous main result, Hypothesis 1). We additionally hypothesized that distress, depressed mood, and UIT frequency would decrease to a smaller degree after rumination about a UIT than after rumination about negative mood and distraction (Hypothesis 2a, b, and c). To extend the previous findings, we predicted that *actual* behavioral and mental neutralizing would be more pronounced after rumination about a UIT than after rumination about negative mood and distraction (Hypothesis 3a and b).

Concerning the potential mechanisms, we predicted that trait rumination and TAF, respectively, would moderate the relation between rumination about a UIT and the urge to neutralize (Hypothesis 4a and b). Specifically, we expected that the higher the level of trait rumination, the stronger the effects of rumination about a UIT on the urge to neutralize would be. Similarly, we predicted that the higher the TAF beliefs, the stronger the rumination about a UIT would affect the urge to neutralize. Furthermore, we hypothesized that rumination about a UIT would increase the negative appraisals of the UIT in comparison to rumination about negative mood and distraction (Hypothesis 5). Finally, we examined

Table 1. Sociodemographic variables, affect, and clinical characteristics.

Variable	Experimental group		
	UIT rumination (<i>n</i> = 34)	Mood rumination (<i>n</i> = 35)	Distraction (<i>n</i> = 36)
Age	21.74 (5.1)	21.69 (3.22)	23.58 (5.97)
Gender (% female)	79.4	88.6	86.1
PANAS: positive	34.59 (5.58)	33.20 (5.48)	33.08 (4.97)
PANAS: negative	23.44 (6.14)	23.51 (6.52)	22.47 (6.38)
OCI-R total	16.00 (8.57)	19.03 (11.70)	16.08 (7.81)
BDI	7.65 (5.44)	8.23 (7.51)	8.92 (4.98)
RRS-brood	11.00 (4.10)	10.69 (2.99)	10.11 (3.12)
TAFS total	22.09 (11.61)	24.37 (12.42)	24.56 (12.15)

Note. All values except for gender are means with standard deviations in parentheses. BDI = Beck Depression Inventory, OCI-R = Obsessive–Compulsive Inventory, Revised, PANAS = Positive and Negative Affect Schedule, RRS-brood = Ruminative Responses Scale, brooding subscale, TAFS = Thought–Action Fusion Scale, UIT = unwanted intrusive thought.

whether rumination about a UIT would affect the UIT vividness in comparison to rumination about negative mood and distraction in an exploratory analysis.

Method

Participants

All participants ($N = 105$) were undergraduate psychology students recruited at the University of Basel ($M_{\text{age}} = 22.35$ years, $SD = 4.94$; 89 females, 16 males). For their participation, they received course credit. During the experimental manipulation, the participants were randomly allocated to rumination about a UIT (UIT rumination group; $n = 34$), rumination about negative mood (mood rumination group; $n = 35$), or distraction ($n = 36$). The experimental groups did not differ in sociodemographic variables, positive or negative affect, depressive symptoms, OC symptoms, degree of brooding, or degree of TAF, all $ps > .05$ (see Table 1 for means and standard deviations [SDs]). The study was reviewed and approved by the institutional review board of the University of Basel (approval number: IRB 009-16-1).

Measures

Standardized questionnaires. The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988; German version: Krohne et al., 1996) is a measure of positive (10 items) and negative (10 items) affect with good reliability and validity (Crawford & Henry, 2004; Krohne et al., 1996). In this study, we measured the affect experienced within the last 12 months.

Cronbach's α was high in the current sample (for positive affect, $\alpha = .81$; for negative affect, $\alpha = .84$).

The Obsessive–Compulsive Inventory–Revised (OCI-R; Foa et al., 2002; German version: Goenner et al., 2007) is an 18-item self-report measure of OC symptoms consisting of six subscales (washing, obsessions, hoarding, ordering, neutralizing, and checking). The scale possesses good psychometric properties (Goenner et al., 2007, 2008). In the current sample, the internal consistency of the total scale was high with Cronbach's $\alpha = .85$.

The Beck Depression Inventory (BDI; Beck et al., 1979; German version: Hautzinger et al., 1995) is a 21-item self-report measure of depressive symptoms. The BDI is a widely used instrument in research with good reliability and validity (Beck et al., 1988). In this sample, Cronbach's $\alpha = .84$.

The Ruminative Responses Scale (RRS; Nolen-Hoeksema & Morrow, 1991; German version: Kuehner et al., 2007) is a 22-item self-report questionnaire that measures trait rumination. For this study, we used only the 5-item brooding scale (RRS-brood), which measures unproductive self-focused responses to sad mood. We chose this subscale because it is not contaminated by items focusing on depression (Treyner et al., 2003). For our study, we used RRS-brood to operationalize trait rumination. Psychometric properties of the brooding subscale are satisfactory (Treyner et al., 2003). Cronbach's α was acceptable in this sample ($\alpha = .70$).

The Thought–Action Fusion Scale (TAFS; Shafran et al., 1996; German version: Hansmeier et al., 2014) is a self-report measure of TAF consisting of two subscales: TAF-moral (12 items) and TAF-likelihood (7 items). The former focuses on a morality

bias and the latter on a probability bias. The TAF-likelihood further differentiates between negative consequences to oneself (TAF-LS) and to others (TAF-LO). The psychometric properties of the scale are good (Hansmeier et al., 2014; Meyer & Brown, 2013; Rassin et al., 2001). In our study, Cronbach's α for the total TAFS score was .89.

The Revised Obsessive Intrusions Inventory Part 2 (ROII Part 2; Purdon & Clark, 1993, 1994) was used to measure the appraisals of the activated UIT. We used 8 of the 10 appraisal items to assess unpleasantness of the UIT, a sense of guilt associated with the UIT, worry that the UIT would come true, unacceptability of the UIT, perceived likelihood of the UIT coming true, the importance of controlling the UIT, perceived harm or danger associated with the UIT, and perceived responsibility for the UIT coming true. We employed this measure to get an indication of the similarity in appraisal ratings between the induced UIT in our study and appraisals in individuals with OC symptoms. All items were rated on a 5-point Likert-type scale ranging from 0 (*not at all*) to 4 (*extremely*).

Assessment of distress, urge to neutralize, depressed mood, and vividness of UITs. Participants were asked to rate distress ("How distressed are you right now?"), urge to neutralize ("To what degree do you experience an urge to neutralize the UITs, that is, to undo the intrusive thought or to do something to prevent something bad happening?"), depressed mood ("How depressed are you right now?"), and vividness of the UITs ("Please indicate how vivid your intrusive thoughts are, that is, to what degree they appear as vivid images in your mind.") on visual analog scales (VASs) ranging from 0 (*very low/not at all vivid*) to 9 (*very high/extremely vivid*).

Assessment of UIT frequency. We assessed UIT frequency with a smartphone counter app. Participants were instructed to press the "+" volume button whenever the UIT occurred. The display was covered so that participants could not see the counter app.

Manipulation checks. To check whether the experimental manipulation worked, the participants were asked to rate their concentration ("What percentage of time were you able to concentrate on the sentences shown?"), degree of self-focus ("While the statements were presented, to what degree were you thinking about yourself?"), and degree of UIT focus ("While the statements were presented, to what degree were

you thinking about causes, meaning, and consequences of your intrusive thoughts or images?") on VASs ranging from 0% to 100%. If the manipulation was effective, participants in the distraction condition should score lower on both self-focus and UIT-focus variables in comparison to those in the rumination groups. In addition, participants in the UIT rumination group should be less self-focused and more UIT focused than those in the mood rumination group.

Assessment of behavioral and mental neutralizing. To assess behavioral and mental neutralizing strategies, we modified items of previous work in that area (Freeston et al., 1991; Goodman et al., 1989; Purdon & Clark, 1993, 1994; Rachman et al., 1996), supplemented by items about leaving the room and inwardly calming oneself down. Items were ordered so that they assessed first behavioral (11 items) and then mental (9 items) forms of neutralizing. Examples of behavioral neutralizing are ripping the paper with the sentence or changing the name. Examples of mental forms of neutralizing are thinking about something positive or saying a prayer. First, the experimenter observed whether the participant had performed any behavioral neutralizing strategies. Subsequently, the experimenter asked the participant whether he or she had used any mental neutralizing strategies. The presence of behavioral or mental neutralizing was rated by the experimenter as either 0 (absent) or 1 (present). Given that the participants mostly performed only one behavior to neutralize the UIT, we analyzed behavioral neutralizing as a dichotomous variable (performed vs. did not perform). The final score for mental neutralizing was the number of different strategies used to neutralize the UIT per participant (sum score), since all participants but one used at least one mental neutralizing strategy. For a detailed description of the neutralizing assessment, see the Appendix.

UIT activation. The UIT activation was identical to that of the previous study (Wahl et al., 2019) and was based on a previous paradigm used to study characteristics of neutralizing in nonclinical samples (Rachman et al., 1996; van den Hout et al., 2001, 2002). First, we provided the participants with a pen and a sheet of paper and asked them to think of a loved person and to get a vivid image of that person in their mind. Once they had a clear picture in their mind, they were instructed to write down and subsequently to read aloud the following sentence: "I wish that [loved person] would die today in a horrible car accident."

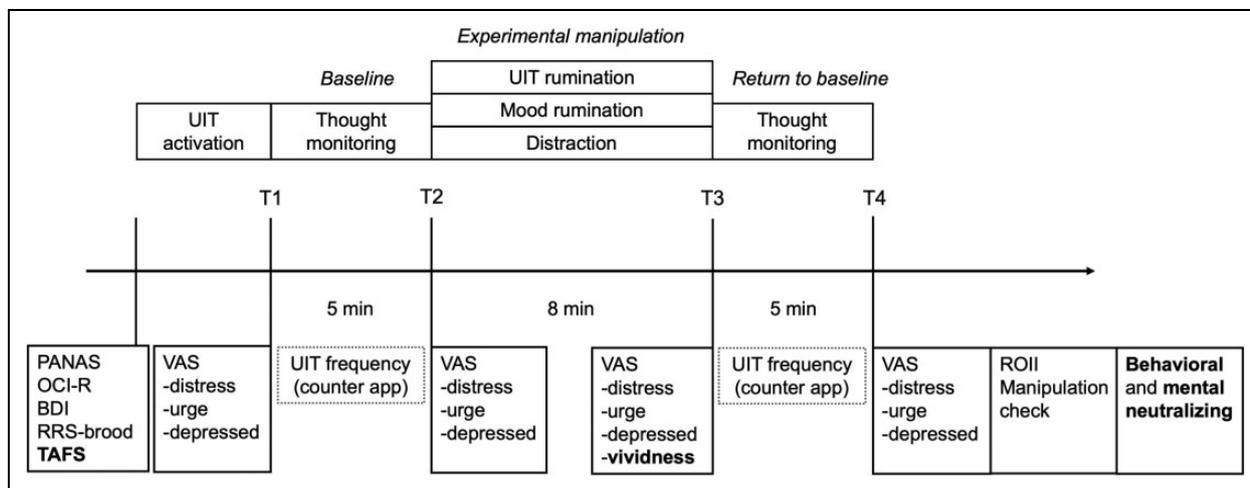


Figure 1. Experimental procedure. The additions that extend Wahl et al. (2019) are marked in bold. BDI = Beck Depression Inventory; OCI-R = Obsessive–Compulsive Inventory, Revised; PANAS = Positive and Negative Affect Schedule; ROI = Revised Obsessive Intrusions Inventory; RRS-brood = Ruminative Responses Scale-brooding subscale; TAFS = Thought–Action Fusion Scale; T1 = before baseline; T2 = after baseline; T3 = after experimental manipulation; T4 = after return to baseline; UIT = unwanted intrusive thought; VAS = visual analog scale. Adapted from K. Wahl et al. (2019). Copyright 2018 by Elsevier. Adapted with permission.

Three participants refused to write down the sentence and two did not read it aloud.

Thought monitoring. During the thought-monitoring phases, participants were instructed to observe their own thoughts for 5 min and to assess the UIT frequency using the counter app, in the same way as in Wahl et al. (2019). The instructions were as follows:

During the next 5 min, you may think about anything you like. You might think of your target unwanted intrusive thought, but you do not have to. However, if at any time you think of your target unwanted intrusive thought, please record the occurrence of each thought by pressing the “+” key on the smartphone once for each occurrence. It is important that you continue in the same way for the full 5 min. (adapted from Marcks & Woods, 2005)

Experimental manipulation. Following Wahl et al. (2019), we randomly assigned participants to the three groups: UIT rumination, mood rumination, or distraction. During the experimental manipulation, they were instructed to focus their attention on statements presented to them on a computer screen for 8 min. Each experimental condition used 28 statements to induce a designated thinking style or distraction. The participants could use the mouse to navigate through the statements.

To induce rumination about the UIT, we used a modified version of the rumination task by Morrow

and Nolen-Hoeksema (1990), which was identical to the one used by Wahl et al. (2019). The participants assigned to this condition were asked to focus on the reasons for, implications of, and possible consequences of their activated UIT (e.g., “Think about: the possible consequences of having intrusive thoughts or images”). In the rumination about negative mood condition, the participants had to focus on reasons for, meanings of, and possible consequences of their negative mood (e.g., “Think about: the way you feel inside”). This induction was based on the depressive rumination paradigm (Blagden & Craske, 1996; Huffziger & Kuehner, 2009; Lyubomirsky et al., 2003; Morrow & Nolen-Hoeksema, 1990). In the distraction condition, the participants were asked to distract themselves by thinking about everyday objects and situations. (e.g., “Think about: raindrops sliding down a window pane”). This condition was based on the original distraction task (Huffziger & Kuehner, 2009; Morrow & Nolen-Hoeksema, 1990).

Procedure

The experimental procedure of the study is depicted in Figure 1. To administer the study, we used the online survey tool Unipark (Questback GmbH, 2013). All participants gave their written informed consent and were tested individually in a quiet room, seated in front of a computer screen. At the beginning, participants were randomly allocated to one of the

three experimental groups and completed a set of standardized questionnaires (PANAS, OCI-R, BDI, RRS-brood, TAFS), followed by the activation of the UIT. The activation of the UIT was immediately followed by the first ratings on the VAS (distress, urge to neutralize, and depressed mood, at T1). Participants were subsequently asked to monitor and register their thoughts for 5 min (baseline), followed by the second ratings (distress, urge to neutralize, and depressed mood) on the VAS (T2). Participants were then asked to follow one of the three instructions of the experimental manipulation for 8 min, succeeded by a third rating of the VAS scales (distress, urge to neutralize, depressed mood, and UIT vividness, at T3). The study continued with the second thought-monitoring phase (return to baseline), followed by the fourth VAS ratings at T4 (distress, urge to neutralize, and depressed mood) and the completion of the ROII items and manipulation checks. Finally, behavioral and mental neutralizing were assessed. The procedure was identical to the procedure used in the study by Wahl et al. (2019) with the addition of three components: TAF was assessed as part of the questionnaire set at the beginning of the study, UIT vividness was assessed as part of the VAS ratings at T3, and behavioral and mental neutralizing were assessed at the end of the study.

Statistical analysis

Data were analyzed with IBM SPSS Statistics 23. First, we used box plots and z scores to identify outliers and detected four in UIT frequency. To reduce the bias, we applied Winsorizing by replacing these outliers with the next highest score that was not an outlier (Field, 2013). Next, we examined the assumption of normality, allowing small violations because our sample size included more than 30 participants per group (Field, 2013). To test whether the experimental groups differed in sociodemographic variables, clinical characteristics, or degree of concentration (first manipulation check), we used one-way analyses of variance (ANOVAs). We performed a χ^2 test to compare the groups on gender. To investigate whether the experimental groups differed in the expected directions during the experimental manipulation (second manipulation check), we conducted a 3×2 mixed-model ANOVA with group (UIT rumination, mood rumination, and distraction) as between-subjects factor and content of thinking (self-focus vs. UIT focus) as a within-subject factor.

We additionally carried out this analysis just for the rumination groups with a 2×2 mixed-model design, excluding distraction.

To examine Hypotheses 1 and 2 (a, b, and c), the outcome variables (distress, urge to neutralize, and depressed mood) were calculated as the mean difference between T2 and T3. We focused on these two time points as the crucial interval since they were immediately before and after the experimental manipulation and could also be directly compared with the Wahl et al.'s (2019) study. The outcome variable UIT frequency was calculated as the mean difference between the two thought-monitoring phases (baseline and return to baseline). Planned contrasts were conducted comparing the UIT rumination group with the combined mood rumination and distraction groups. Where necessary, we adjusted for heterogeneity of variances. To analyze the effect of rumination about the UIT on behavioral and mental neutralizing (Hypothesis 3a and b), we conducted a logistic regression and a Welch test (due to slight heterogeneity of variances on this variable), respectively.

To investigate the moderating effects of trait rumination and TAF (Hypothesis 4a and b), we conducted moderation analyses using PROCESS (Hayes, 2017). For the interactions, we report percentile bootstrap 95% confidence intervals (CIs), because this method is more robust than the standard CI (Field, 2013). Significant moderation effects were followed by a simple slopes analysis (Aiken & West, 1991; Rogosa, 1981). This analysis looks at the relation between the predictor and outcome at 1 SD above and below the mean value of the moderator. Specifically, we looked at the effect of rumination about the UIT on the urge to neutralize from T2 to T3 at lower (1 SD below mean), average (mean), and higher (1 SD above mean) levels of trait rumination and TAF, respectively. Urge to neutralize from T2 to T3 was calculated as the mean difference between these two time points. To analyze the effect of rumination about the UIT on UIT appraisals (Hypothesis 5), we carried out a multivariate ANOVA, which tests the difference between groups across multiple outcomes simultaneously and therefore controls better for multiple testing (Field, 2013). Last, changes in UIT vividness were analyzed with an ANOVA, followed by Tukey's post hoc test.

As effect sizes, we report Pearson's correlation coefficient (r ; planned contrasts), partial eta-squared (η_p^2 ; one-way independent and mixed ANOVAs), odds ratios (ORs ; logistic regression), and Cohen's (d ;

Table 2. Manipulation checks by experimental group.

Variable	Experimental group		
	UIT rumination ($n = 34$)	Mood rumination ($n = 35$)	Distraction ($n = 36$)
Concentration	82.06 (15.81)	83.14 (13.77)	84.86 (11.96)
Content of thinking			
Self-focus	73.32 (26.47)	84.57 (24.02)	30.14 (28.29)
UIT focus	62.82 (32.55)	49.20 (34.63)	19.31 (22.80)

Note. Values are means with standard deviations in parentheses. UIT = unwanted intrusive thought.

Tukey's post hoc test). Level of significance was set at .05.

Results

Manipulation Check

First, we analyzed whether participants were able to concentrate on the statements provided during the experimental manipulation. Results showed that on average, all participants were able to focus to a high degree on the statements, with no significant group differences, $F(2, 102) = 0.36, p = .696$ (see Table 2 for all manipulation check means and SDs). Next, we investigated differences between groups on the content of thinking. We expected that distraction would lead to lower scores on self-focus and UIT focus in comparison to rumination about the UIT and rumination about negative mood. In addition, we predicted that the rumination groups would differ from each other in the content being ruminated on (self-focus vs. UIT focus), with rumination about the UIT being more UIT focused and rumination about negative mood being more self-focused. A significant interaction between the experimental group and the content of rumination with the expected patterns suggests that three distinct thinking styles were successfully induced, $F(2, 102) = 6.41, p = .002, \eta_p^2 = .11$. The interaction remained significant when excluding distraction from the analysis, $F(1, 67) = 8.47, p = .005, \eta_p^2 = .11$, meaning that the two types of rumination differed in the expected direction (self-focus vs. UIT focus). In short, the experimental manipulation was successfully implemented.

Replication

Hypothesis 1: Effect of experimental manipulation on urge to neutralize.

Table 3 presents means and SDs for urge to neutralize for each time point. The analysis focuses on the

comparison between T2 and T3. Urge to neutralize showed an attenuated decline in those participants who had previously ruminated about their UIT compared to participants who had ruminated about negative mood and those who were distracted, $t(81.58) = 2.74, p = .007, r = .29$. For illustration and comparison with the study by Wahl et al. (2019), means and standard errors of the urge to neutralize are depicted graphically in Figure 2.

Hypothesis 2a, b, and c: Effects of experimental manipulation on distress, depressed mood, and UIT frequency.

Table 3 presents means and SDs for distress, depressed mood, and UIT frequency. Similar to urge to neutralize, distress showed an attenuated decline from T2 to T3 in those participants who were in the UIT rumination group compared to participants who were in the mood rumination and distraction groups, $t(102) = 2.45, p = .016, r = .24$. The effect of rumination about the UIT on depressed mood was not significant, $t(102) = 1.56, p = .121, r = .15$. Figure 3 shows the effects for distress and depressed mood. With regard to UIT frequency, there was no significant difference between the experimental groups, $t(102) = 0.45, p = .656, r = .04$.

Additional analysis: Rumination about negative mood versus distraction. Visual inspection of the data (Figures 2 and 3) and the means in Table 3 suggest that the two rumination groups followed a similar pattern, that is, a reduced decline in comparison to distraction. Thus, we decided to conduct an additional contrast to test the differences between rumination about negative mood and distraction for statistical significance. Relative to distraction, rumination about negative mood maintained the urge to neutralize, $t(60.84) = 3.42, p = .001, r = .40$; distress,

Table 3. Distress, urge to neutralize, depressed mood, UIT frequency, mental neutralizing, UIT appraisals, and UIT vividness by experimental group with time points.

Variable	Experimental group		
	UIT rumination (<i>n</i> = 34)	Mood rumination (<i>n</i> = 35)	Distraction (<i>n</i> = 36)
Distress			
T1	5.82 (2.14)	5.89 (1.57)	6.17 (1.98)
T2	3.62 (2.09)	4.40 (1.90)	4.31 (2.54)
T3	3.59 (2.18)	3.94 (2.14)	3.00 (2.08)
T4	2.44 (1.97)	3.09 (2.16)	2.44 (1.78)
Urge to neutralize			
T1	6.24 (2.66)	6.26 (2.58)	6.47 (2.62)
T2	4.50 (2.70)	4.29 (2.41)	4.94 (2.86)
T3	3.68 (2.91)	3.31 (2.39)	2.33 (2.27)
T4	2.74 (2.60)	2.66 (2.20)	2.56 (2.37)
Depressed mood			
T1	4.68 (2.53)	4.20 (2.40)	4.67 (2.41)
T2	3.12 (2.29)	3.06 (1.96)	3.56 (2.32)
T3	3.15 (2.34)	3.26 (2.25)	2.22 (2.02)
T4	2.24 (2.06)	2.74 (1.93)	2.08 (1.96)
UIT frequency			
Baseline	17.82 (16.26)	15.74 (12.29)	19.67 (15.49)
Return to baseline	12.41 (18.42)	11.20 (18.22)	11.22 (10.02)
Mental neutralizing			
UIT appraisals (ROII)	4.88 (1.45)	5.23 (1.52)	4.61 (1.89)
Unpleasantness	3.03 (1.14)	2.74 (1.17)	2.67 (1.12)
Guilt	2.06 (1.23)	1.91 (1.36)	1.89 (1.28)
Worry	1.50 (1.02)	1.54 (1.34)	1.28 (1.11)
Unacceptability	2.59 (1.21)	2.00 (1.26)	2.58 (1.20)
Likelihood	0.71 (0.72)	0.77 (0.65)	0.61 (0.60)
Control	2.26 (1.08)	2.14 (1.22)	1.72 (1.32)
Harm/danger	1.00 (1.10)	0.91 (1.10)	0.92 (0.94)
Responsibility	1.29 (1.14)	1.09 (1.27)	1.19 (1.09)
UIT vividness at T3			
	4.03 (2.62)	3.29 (3.03)	2.50 (2.89)

Note. Values are means with standard deviations in parentheses. ROII = Revised Obsessive Intrusions Inventory; T1 = before baseline; T2 = after baseline; T3 = after experimental manipulation; T4 = after return to baseline; UIT = unwanted intrusive thought.

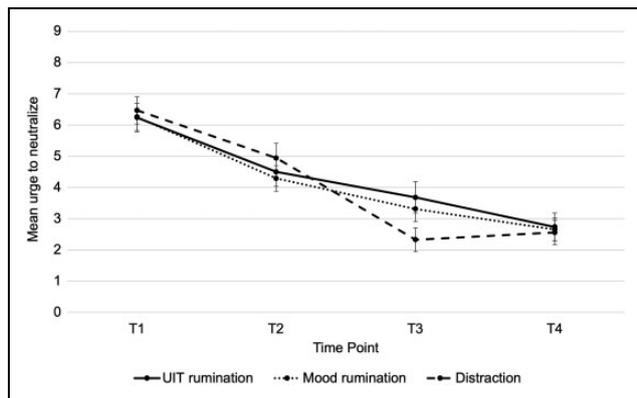


Figure 2. Means of urge to neutralize by group (UIT rumination, mood rumination, and distraction). Error bars represent standard errors. UIT = unwanted intrusive thought.

$t(102) = 2.15, p = .034, r = .21$; and depressed mood, $t(102) = 3.54, p = .001, r = .33$, from T2 to T3.

Extension

Hypothesis 3a and b: Effect of experimental manipulation on behavioral and mental neutralizing.

There was no significant effect of the experimental manipulation on behavioral neutralizing,¹ $b = 0.65, SE = 0.50, p = .196, OR = 1.92, 95\% CI [0.715, 5.157]$ (UIT rumination vs. distraction), $b = -0.19, SE = 0.54, p = .731, OR = 0.83, 95\% CI [0.288, 2.392]$ (UIT rumination vs. mood rumination), or mental neutralizing, $F(2, 67.63) = 1.20, p = .309$,

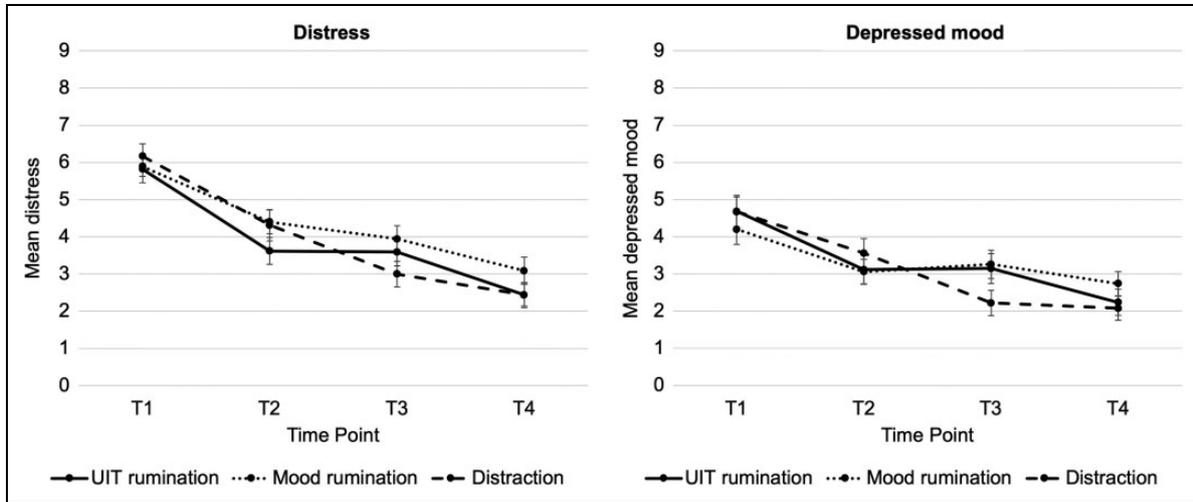


Figure 3. Means of distress and depressed mood by group (UIT rumination, mood rumination, and distraction). Error bars represent standard errors. UIT = unwanted intrusive thought.

$\eta_p^2 = .03$. Means and SDs of mental neutralizing are presented in Table 3.

Mechanisms

Hypothesis 4a and b: Moderating effects of trait rumination and TAF.

When comparing rumination about the UIT with distraction, trait rumination did not moderate the effect of the experimental manipulation on the urge to neutralize, $b = -0.28$, percentile 95% CI [-0.637, 0.027], $t = -2.25$, $p = .027$, since the percentile bootstrap CI included zero. TAF moderated the relation between the experimental group and the urge to neutralize from T2 to T3, $b = -0.09$, percentile 95% CI [-0.155, -0.012], $t = -2.33$, $p = .022$, when comparing rumination about the UIT with distraction. When comparing rumination about negative mood with distraction, the moderation was not significant, $b = -0.07$, percentile 95% CI [-0.145, 0.010], $t = -1.98$, $p = .051$. Further analysis showed that the effect of TAF occurred only in the distraction group and not in the two rumination groups (see Table 4). Specifically, when TAF score increased, so did the mean difference in urge to neutralize from T2 to T3; that is, urge to neutralize decreased to a greater degree.

Hypothesis 5: Effect of experimental manipulation on UIT appraisals.

We did not find evidence for an effect of rumination about the UIT on the UIT appraisals, $V = 0.15$,

Table 4. Conditional effects of TAF on urge to neutralize from T2 to T3 by experimental group.

Group	TAF	Mean difference in urge to neutralize from T2 to T3	t	p
UIT rumination	M -1 SD	0.94	-0.40	.691
	M	0.81		
	M +1 SD	0.68		
Mood rumination	M -1 SD	0.90	0.21	.830
	M	0.97		
	M +1 SD	1.03		
Distraction	M -1 SD	1.63	3.00	.003
	M	2.55		
	M +1 SD	3.46		

Note. The moderation occurred only in the distraction group, which is highlighted by the significant t statistic. As TAF score increased, so did the effect of distraction on the urge to neutralize. TAF = Thought-Action Fusion.

$F(16, 192) = 0.96$, $p = .504$ (for means and SDs, see Table 3).

Exploratory analysis: Effect of rumination on UIT vividness. The effect of the experimental group on the UIT vividness was not significant, $F(2, 102) = 2.51$, $p = .086$, $\eta_p^2 = .05$ (for means and SDs, see Table 3).

Discussion

The major goal of the study was to investigate whether rumination about a UIT reduces the natural decline of the urge to neutralize these thoughts relative to rumination about negative mood and

distraction, thereby replicating findings by Wahl et al. (2019). Consistent with Hypothesis 1, rumination about the UIT attenuated the general decline of the urge to neutralize the UIT relative to rumination about negative mood and distraction. Interestingly, we observed a similar effect on the urge to neutralize for individuals who had ruminated about negative mood, relative to those in the distraction group. In other words, individuals who ruminated—irrespective of the content of rumination—experienced a reduced decline in the urge to neutralize compared to individuals who were distracted. Findings partially replicate results by Wahl et al. (2019) who used a novel experimental paradigm for the first time. This replication in an independent sample further strengthens our confidence in the assumption that rumination about a UIT is involved in the maintenance of the urge to act upon these thoughts.

In contrast to the previous study (Wahl et al., 2019), rumination about the UIT also attenuated the decline of distress associated with the UIT, relative to rumination about the negative mood and distraction (Hypothesis 2a). These findings suggest that rumination about the UIT is involved not only in the maintenance of the urge to neutralize but possibly also in the persistence of the distress experienced with the UIT. It is possible that the nonsignificant effect on distress found earlier represented a power problem, as Wahl et al. (2019) assessed a smaller sample. In addition, rumination about negative mood maintained both the distress associated with the UIT and depressed mood, relative to distraction. The present findings suggest that the two types of rumination exerted similar effects on the urge to neutralize and distress. The reasons for this discrepancy in the effects of rumination about negative mood between the previous study (Wahl et al., 2019) and the current study are unclear. Manipulation checks in both studies indicate that two distinct types of rumination were successfully induced to a similar degree, using identical methodology. The differences between rumination on UIT and rumination on negative mood seem to be particularly relevant for individuals diagnosed with OCD (Wahl, Schönfeld et al., 2011), and future studies should retain this distinction and investigate whether they differ in their effects on OC symptoms.

Neither the current nor the previous study (Wahl et al., 2019) found an effect of the experimental manipulation on the frequency of UITs (Hypothesis 2c). Whether this can be interpreted as a robust finding, meaning that the frequency of UITs is unaffected

by previous rumination, or whether the findings are attributable to methodological difficulties such as the reliable and valid assessment of such transient phenomena as UITs remain to be seen in future studies, which could develop valid assessments of UITs that also differentiate between frequency and duration of UITs.

The second aim was to extend the previous findings by investigating whether also *actual* behavioral and mental neutralizing are affected by rumination. Although the effect was going in the predicted direction (i.e., participants in the UIT rumination group performing behavioral neutralizing more often than participants in the distraction group), the experimental groups did not significantly differ in their effect on behavioral neutralizing (Hypothesis 3a). This means that participants in each group engaged in actual neutralizing behavior such as ripping the paper or crossing out the name of the loved person to a similar degree. Likewise, rumination about the UIT did not result in a higher frequency of mental neutralizing strategies compared to rumination about the negative mood and distraction (Hypothesis 3b). Each group engaged in a high total number of mental neutralizing strategies (UIT rumination: $M = 4.88$, $SD = 1.45$; mood rumination: $M = 5.23$, $SD = 1.52$; distraction: $M = 4.61$, $SD = 1.89$), such as saying silently to oneself that the sentence “does not count” since one was told to write it down to fulfill the requirements of the study. Thus, the lack of group differences might be explained by a ceiling effect, attributable to the strong activation of a UIT. In sum, results do not support Hypothesis 3a and b. Future studies should address this question with more power and a more refined assessment of actual behavioral and mental neutralizing.

Several potential mechanisms involved in the associations between rumination and UITs were addressed in this study in a conjunct analysis to better understand how exactly rumination impacts the maintenance of UITs. We did not find evidence of trait rumination or TAF moderating the effect of rumination about the UIT on the urge to neutralize (Hypothesis 4a and b). Thus, it appears that rumination about a UIT affects the urge to neutralize regardless of the tendency to ruminate in daily life, or the predisposition to misinterpret the occurrence and meaning of UITs. Additionally, we did not find that rumination directly affected the appraisals of the UIT (e.g., rumination did not increase the perceived likelihood of the thought coming true, relative to distraction), which is consistent with the finding on TAF. Future studies

might focus on the assessment of key misinterpretations of UITs and their potential changes as a result of rumination.

Finally, we examined whether rumination about the UIT influenced the vividness of the UIT. Findings suggest that vividness did not differ between the groups. Future studies might want to include more aspects of imagery related to UITs to draw conclusions about their possible involvement in the effects of rumination on UITs.

There are several limitations to this study that should be addressed. First, the use of one-item measures as the main dependent variables poses questions about their psychometric properties. Second, the current and the previous study (Wahl et al., 2019) did not find an effect of rumination about the UIT on the frequency of UITs. These findings are surprising, given that in the current study rumination about the UIT had a broader impact on variables (urge to neutralize and distress). This raises the question of whether the use of a smartphone to assess frequency of UITs is a valid and reliable measure. Future studies might consider other measures of intrusive thoughts, for instance, the think-aloud approach (Zanov & Davison, 2010) or thought sampling (Hirsch et al., 2015). Finally, replications of key findings should ideally be conducted by an independent research group. Although the study was double blind, it cannot be completely ruled out that the previous experiences with the experimental paradigm or the researchers' expectations might have influenced the current results.

To conclude, the study replicated the main result from a previous study (Wahl et al., 2019) that rumination about a UIT attenuates the decrease in urge to neutralize compared to rumination about negative mood and distraction. Given the additional findings that rumination about the UIT affected distress, this strengthens the confidence in the argument that rumination has an influence on the maintenance of UITs. Findings are also consistent with previous studies suggesting that rumination might play a role in OCD (Dar & Iqbal, 2015; Grisham & Williams, 2009; Raines et al., 2017; Wahl et al., 2019). Findings warrant replication in individuals diagnosed with OCD to investigate whether rumination also affects obsessive thoughts in the same way as it affects UITs.

Acknowledgments

The authors would like to thank Florence Allegro, Antonino Cusano, Anja Laura Käch, Rahel Velte, and Leonie Kanzler for collecting the data.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The Swiss National Science Foundation (SNSF) is funding the PhD position for Martin Kollárik. The SNSF was not involved in the research question, data collection, data analysis, interpretation of results, or publication.

ORCID iD

Martin Kollárik  <https://orcid.org/0000-0002-7719-8065>
 Carlotta V. Heinzl  <https://orcid.org/0000-0002-2619-913X>
 Karina Wahl  <https://orcid.org/0000-0002-9232-0911>

Note

1. We changed the assignment of three strategies. The experimenter included two strategies (“I don’t wish that on anyone” and “it is not my wish but a task in the study”) in the behavioral neutralizing subscale (item: “Other”) that we did not consider as observable and hence categorized as mental neutralizing. In the mental neutralizing subscale (item: “Other”), one participant reported having looked out of the window as a strategy. This was an observable act. Therefore, we recategorized it as behavioral neutralizing.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. SAGE.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the beck depression inventory—25 years of evaluation. *Clinical Psychology Review, 8*, 77–100. [http://doi.org/10.1016/0272-7358\(88\)90050-5](http://doi.org/10.1016/0272-7358(88)90050-5)
- Birrer, E., Michael, T., & Munsch, S. (2007). Intrusive images in PTSD and in traumatised and non-traumatised depressed patients: A cross-sectional clinical study. *Behaviour Research and Therapy, 45*, 2053–2065. <http://doi.org/10.1016/j.brat.2007.03.005>

- Blagden, J. C., & Craske, M. G. (1996). Effects of active and passive rumination and distraction: A pilot replication with anxious mood. *Journal of Anxiety Disorders, 10*, 243–252. [http://doi.org/10.1016/0887-6185\(96\)00009-6](http://doi.org/10.1016/0887-6185(96)00009-6)
- Carney, C. E., Harris, A. L., Falco, A., & Edinger, J. D. (2013). The relation between insomnia symptoms, mood, and rumination about insomnia symptoms. *Journal of Clinical Sleep Medicine, 9*, 567–575. <http://doi.org/10.5664/jcsm.2752>
- Clark, D. A. (2002). Cognitive perspective on obsessive compulsive disorder and depression: Distinct and related features. In R. O. Frost & G. Steketee (Eds.), *Cognitive approaches to obsessions and compulsions: Theory, assessment, and treatment* (pp. 233–250). Pergamon.
- Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265. <http://doi.org/10.1348/0144665031752934>
- Dar, K. A., & Iqbal, N. (2015). Worry and rumination in generalized anxiety disorder and obsessive compulsive disorder. *Journal of Psychology, 149*, 866–880. <http://doi.org/10.1080/00223980.2014.986430>
- Di Schiena, R., Luminet, O., & Philippot, P. (2011). Adaptive and maladaptive rumination in alexithymia and their relation with depressive symptoms. *Personality and Individual Differences, 50*, 10–14. <http://doi.org/10.1016/j.paid.2010.07.037>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE.
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment, 14*, 485–496. <http://doi.org/10.1037/1040-3590.14.4.485>
- Freeston, M. H., & Ladouceur, R. (1997). What do patients do with their obsessive thoughts? *Behaviour Research and Therapy, 35*, 335–348. [http://doi.org/10.1016/S0005-7967\(96\)00094-0](http://doi.org/10.1016/S0005-7967(96)00094-0)
- Freeston, M. H., Ladouceur, R., Thibodeau, N., & Gagnon, F. (1991). Cognitive intrusions in a non-clinical population. I. Response style, subjective experience, and appraisal. *Behaviour Research and Therapy, 29*, 585–597.
- Goenner, S., Leonhart, R., & Ecker, W. (2007). Das Zwangsinventar OCI-R—die deutsche Version des Obsessive–Compulsive Inventory-Revised: Ein kurzes Selbstbeurteilungsinstrument zur mehrdimensionalen Messung von Zwangssymptomen [The German version of the Obsessive–Compulsive Inventory-Revised: A brief self-report measure for the multidimensional assessment of obsessive–compulsive symptoms]. *Psychotherapie Psychosomatik Medizinische Psychologie, 57*, 395–404.
- Goenner, S., Leonhart, R., & Ecker, W. (2008). The Obsessive-Compulsive Inventory-Revised (OCI-R): Validation of the German version in a sample of patients with OCD, anxiety disorders, and depressive disorders. *Journal of Anxiety Disorders, 22*, 734–749. <http://doi.org/10.1016/j.janxdis.2007.07.007>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Archives of General Psychiatry, 46*, 1006–1011.
- Grisham, J. R., & Williams, A. D. (2009). Cognitive control of obsessional thoughts. *Behaviour Research and Therapy, 47*, 395–402. <http://doi.org/10.1016/j.brat.2009.01.014>
- Hansmeier, J., Exner, C., Rief, W., & Glombiewski, J. (2014). Gedanken-Handlungs-Fusion und Zwangssymptome—eine Validierungsstudie der Thought Action Fusion-Skala (TAF-Skala) für den deutschen Sprachraum. *Zeitschrift für Klinische Psychologie und Psychotherapie, 43*, 204–214.
- Harrington, J. A., & Blankenship, V. (2002). Ruminative thoughts and their relation to depression and anxiety. *Journal of Applied Social Psychology, 32*, 465–485. <http://doi.org/10.1111/j.1559-1816.2002.tb00225.x>
- Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1995). *Beck-Depressions-Inventar (BDI). Testhandbuch* (2nd rev. ed.). Hans Huber.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.
- Hirsch, C. R., Perman, G., Hayes, S., Eagleson, C., & Mathews, A. (2015). Delineating the role of negative verbal thinking in promoting worry, perceived threat, and anxiety. *Clinical Psychological Science, 3*, 637–647. <http://doi.org/10.1177/2167702615577349>
- Huffziger, S., & Kuehner, C. (2009). Rumination, distraction, and mindful self-focus in depressed patients. *Behaviour Research and Therapy, 47*, 224–230. <http://doi.org/10.1016/j.brat.2008.12.005>
- James, E. L., Lau-Zhu, A., Clark, I. A., Visser, R. M., Hagens, M. A., & Holmes, E. A. (2016). The trauma film paradigm as an experimental psychopathology model of psychological trauma: Intrusive memories and

- beyond. *Clinical Psychology Review*, *47*, 106–142. <http://doi.org/10.1016/j.cpr.2016.04.010>
- Just, N., & Alloy, L. B. (1997). The response styles theory of depression: Tests and an extension of the theory. *Journal of Abnormal Psychology*, *106*, 221–229. <http://doi.org/10.1037/0021-843x.106.2.221>
- Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Investigations with a German version of the positive and negative affect schedule (PANAS). *Diagnostica*, *42*, 139–156.
- Kuehner, C., Huffziger, S., & Nolen-Hoeksema, S. (2007). *Response Styles Questionnaire: German version*. Hogrefe.
- Lipton, M. G., Brewin, C. R., Linke, S., & Halperin, J. (2010). Distinguishing features of intrusive images in obsessive-compulsive disorder. *Journal of Anxiety Disorders*, *24*, 816–822. <http://doi.org/10.1016/j.janxdis.2010.06.003>
- Luo, P. C., Jiang, Y. J., Dang, X. J., Huang, Y. S., Chen, X. J., & Zheng, X. F. (2013). Effects of different forms of verbal processing on the formation of intrusions. *Journal of Traumatic Stress*, *26*, 288–294. <http://doi.org/10.1002/jts.21800>
- Lyubomirsky, S., Kasri, F., & Zehm, K. (2003). Dysphoric rumination impairs concentration on academic tasks. *Cognitive Therapy and Research*, *27*, 309–330. <http://doi.org/10.1023/A:1023918517378>
- Marcks, B. A., & Woods, D. W. (2005). A comparison of thought suppression to an acceptance-based technique in the management of personal intrusive thoughts: A controlled evaluation. *Behaviour Research and Therapy*, *43*, 433–445. <http://doi.org/10.1016/j.brat.2004.03.005>
- Meyer, J. F., & Brown, T. A. (2013). Psychometric evaluation of the thought-action fusion scale in a large clinical sample. *Assessment*, *20*, 764–775. <http://doi.org/10.1177/1073191112436670>
- Moberly, N. J., & Watkins, E. R. (2006). Processing mode influences the relationship between trait rumination and emotional vulnerability. *Behavior Therapy*, *37*, 281–291. <http://doi.org/10.1016/j.beth.2006.02.003>
- Morrow, J., & Nolen-Hoeksema, S. (1990). Effects of responses to depression on the remediation of depressive affect. *Journal of Personality and Social Psychology*, *58*, 519–527. <http://doi.org/10.1037/0022-3514.58.3.519>
- Naumann, E., Tuschen-Caffier, B., Voderholzer, U., Caffier, D., & Svaldi, J. (2015). Rumination but not distraction increases eating-related symptoms in anorexia and bulimia nervosa. *Journal of Abnormal Psychology*, *124*, 412–420. <http://doi.org/10.1037/abn0000046>
- Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*, 569–582.
- Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*, 504–511. <http://doi.org/10.1037/0021-843x.109.3.504>
- Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology*, *61*, 115–121.
- Nolen-Hoeksema, S., Stice, E., Wade, E., & Bohon, C. (2007). Reciprocal relations between rumination and bulimic, substance abuse, and depressive symptoms in female adolescents. *Journal of Abnormal Psychology*, *116*, 198–207. <http://doi.org/10.1037/0021-843X.116.1.198>
- Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, *3*, 400–424. <http://doi.org/10.1111/j.1745-6924.2008.00088.x>
- Pearson, D. G., Deeprase, C., Wallace-Hadrill, S. M. A., Heyes, S. B., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical Psychology Review*, *33*, 1–23. <http://doi.org/10.1016/j.cpr.2012.09.001>
- Purdon, C., & Clark, D. A. (1993). Obsessive intrusive thoughts in nonclinical subjects. Part I. Content and relation with depressive, anxious and obsessional symptoms. *Behaviour Research and Therapy*, *31*, 713–720.
- Purdon, C., & Clark, D. A. (1994). Obsessive intrusive thoughts in nonclinical subjects. Part II. Cognitive appraisal, emotional response and thought control strategies. *Behaviour Research and Therapy*, *32*, 403–410.
- Questback GmbH. (2013). Unipark [Survey software]. <https://www.unipark.com/>
- Rachman, S. (1997). A cognitive theory of obsessions. *Behaviour Research and Therapy*, *35*, 793–802. [http://doi.org/10.1016/S0005-7967\(97\)00040-5](http://doi.org/10.1016/S0005-7967(97)00040-5)
- Rachman, S. (1998). A cognitive theory of obsessions: Elaborations. *Behaviour Research and Therapy*, *36*, 385–401. [http://doi.org/10.1016/S0005-7967\(97\)10041-9](http://doi.org/10.1016/S0005-7967(97)10041-9)
- Rachman, S., Shafran, R., Mitchell, D., Trant, J., & Teachman, B. (1996). How to remain neutral: An experimental analysis of neutralization. *Behaviour Research and Therapy*, *34*, 889–898. [http://doi.org/10.1016/S0005-7967\(96\)00051-4](http://doi.org/10.1016/S0005-7967(96)00051-4)
- Raines, A. M., Vidaurri, D. N., Portero, A. K., & Schmidt, N. B. (2017). Associations between rumination and obsessive-compulsive symptom dimensions. *Personality and Individual Differences*, *113*, 63–67. <http://doi.org/10.1016/j.paid.2017.03.001>

- Rassin, E., Merckelbach, H., Muris, P., & Schmidt, H. (2001). The Thought-Action Fusion Scale: Further evidence for its reliability and validity. *Behaviour Research and Therapy, 39*, 537–544.
- Reynolds, M., & Salkovskis, P. M. (1991). The relationship among guilt, dysphoria, anxiety and obsessions in a normal population—an attempted replication. *Behaviour Research and Therapy, 29*, 259–265. [http://doi.org/10.1016/0005-7967\(91\)90116-K](http://doi.org/10.1016/0005-7967(91)90116-K)
- Rogosa, D. (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. *Educational and Psychological Measurement, 41*, 73–84. <http://doi.org/10.1177/001316448104100108>
- Salkovskis, P. M. (1985). Obsessional-compulsive problems—A cognitive-behavioral analysis. *Behaviour Research and Therapy, 23*, 571–583. [http://doi.org/10.1016/0005-7967\(85\)90105-6](http://doi.org/10.1016/0005-7967(85)90105-6)
- Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders, 10*, 379–391. [http://doi.org/10.1016/0887-6185\(96\)00018-7](http://doi.org/10.1016/0887-6185(96)00018-7)
- Speckens, A. E. M., Hackmann, A., Ehlers, A., & Cuthbert, B. (2007). Imagery special issue: Intrusive images and memories of earlier adverse events in patients with obsessive compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 411–422. <http://doi.org/10.1016/j.jbtep.2007.09.004>
- Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research, 27*, 247–259. <http://doi.org/10.1023/A:1023910315561>
- Van den Hout, M., Kindt, M., Weiland, T., & Peters, M. (2002). Instructed neutralization, spontaneous neutralization and prevented neutralization after an obsession-like thought. *Journal of Behavior Therapy and Experimental Psychiatry, 33*, 177–189. [http://doi.org/10.1016/S0005-7916\(02\)00048-4](http://doi.org/10.1016/S0005-7916(02)00048-4)
- Van den Hout, M., van Pol, M., & Peters, M. (2001). On becoming neutral: Effects of experimental neutralizing reconsidered. *Behaviour Research and Therapy, 39*, 1439–1448.
- Wahl, K., Ertle, A., Bohne, A., Zurowski, B., & Kordon, A. (2011). Relations between a ruminative thinking style and obsessive-compulsive symptoms in non-clinical samples. *Anxiety Stress and Coping, 24*, 217–225. <http://doi.org/10.1080/10615806.2010.482985>
- Wahl, K., Schönfeld, S., Hissbach, J., Küsel, S., Zurowski, B., Moritz, S., Hohagen, F., & Kordon, A. (2011). Differences and similarities between obsessive and ruminative thoughts in obsessive-compulsive and depressed patients: A comparative study. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 454–461. <http://doi.org/10.1016/j.jbtep.2011.03.002>
- Wahl, K., van den Hout, M., & Lieb, R. (2019). Rumination on unwanted intrusive thoughts affects the urge to neutralize in nonclinical individuals. *Journal of Obsessive-Compulsive and Related Disorders, 20*, 4–12. <http://doi.org/10.1016/j.jocrd.2018.02.002>
- Watkins, E. (2004). Adaptive and maladaptive ruminative self-focus during emotional processing. *Behaviour Research and Therapy, 42*, 1037–1052. <http://doi.org/10.1016/j.brat.2004.01.009>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect—The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. <http://doi.org/10.1037/0022-3514.54.6.1063>
- White, R., & Wild, J. (2016). “Why” or “how”: The effect of concrete versus abstract processing on intrusive memories following analogue trauma. *Behavior Therapy, 47*, 404–415. <http://doi.org/10.1016/j.beth.2016.02.004>
- Zanov, M. V., & Davison, G. C. (2010). A conceptual and empirical review of 25 years of cognitive assessment using the articulated thoughts in simulated situations (ATSS) think-aloud paradigm. *Cognitive Therapy and Research, 34*, 282–291. <http://doi.org/10.1007/s10608-009-9271-9>
- Zetsche, U., Ehring, T., & Ehlers, A. (2009). The effects of rumination on mood and intrusive memories after exposure to traumatic material: An experimental study. *Journal of Behavior Therapy and Experimental Psychiatry, 40*, 499–514. <http://doi.org/10.1016/j.jbtep.2009.07.001>

Author biographies

Martin Kollárik is a PhD candidate in the Department of Psychology, Division of Clinical Psychology and Epidemiology, at the University of Basel, Switzerland. His research focuses on the experience of unwanted intrusive thoughts in relation to obsessive-compulsive symptoms.

Marcel van den Hout is a Professor of Clinical Psychology and Experimental Psychopathology at Utrecht University, Netherlands. His research focuses on obsessive-compulsive disorder, anxiety disorders and Post Traumatic Stress Disorder.

Carlotta V. Heinzel is a PhD candidate in the Department of Psychology, Division of Clinical Psychology and Epidemiology, at the University of Basel, Switzerland. Her research interests include obsessive-compulsive disorder and rumination.

Patrizia D. Hofer is a research assistant and an assistant psychologist in Psychosomatics at the University Hospital in Basel. Her research interests include prevention of mental disorders in vulnerable individuals.

Roselind Lieb is Professor and Head of the Division of Clinical Psychology and Epidemiology in the Department of Psychology at the University of Basel, Switzerland. Her research concentrates on risk factors and pathways into various mental disorders.

Karina Wahl is a research scientist Post-Doc in the Department of Psychology, Division of Clinical Psychology and Epidemiology, at the University of Basel, Switzerland. Her research interests include obsessive-compulsive disorder, transdiagnostic processes such as repetitive negative thinking, and experimental psychopathology.

Appendix

Assessment of neutralizing

The behavioral neutralizing items were tailored to cover strategies that were identified by Rachman et al. (1996), including (a) canceling out (e.g., ripping

the paper with the sentence), (b) counter-balancing (e.g., changing the sentence so it says something positive), and (c) reassurance seeking (e.g., sending a message to the person written in the sentence). We also included a modified item from the *Yale-Brown Obsessive Compulsive Scale* (Goodman et al., 1989) assessing forms of superstitious behavior (e.g., knocking on wood, touching a crucifix, etc.). Further, we also used a self-developed item: leaving the room. With regard to the mental neutralizing items, we focused on mental strategies that a participant could employ to neutralize the activated UIT. The items were taken and modified primarily from the ROII Part 2 (Purdon . . . Clark, 1993, 1994), for example, trying to argue that everything was all right or praying. One item concerning the relevance of the intrusive thoughts was modified from the Cognitive Intrusions Questionnaire (Freeston et al., 1991) and one was self-developed: “. . . to inwardly calm oneself down.”

The investigator gave the paper with the written UIT to the participant and assessed whether he or she showed any behavioral strategies to neutralize the activated UIT. Subsequently, the experimenter continued to assess mental strategies.

Appendix B (Manuscript II)

Validity of the counter-app method in the assessment of intrusions

Martin Kollárik, Colette R. Hirsch, Carlotta V. Heinzl, Roselind Lieb, Karina Wahl

submitted to *Journal of Behavior Therapy and Experimental Psychiatry*

Validity of the Counter-App Method in the Assessment of Intrusions

Martin Kollárik^a, Colette R. Hirsch^b, Carlotta V. Heinzl^a, Roselind Lieb^a, Karina Wahl^{a*}

^aDepartment of Psychology, Division of Clinical Psychology and Epidemiology,
University of Basel, Switzerland

^bPsychology Department, Institute of Psychiatry, Psychology and Neuroscience,
King's College London, United Kingdom & South London and Maudsley NHS
Foundation Trust, United Kingdom

* corresponding author: Karina Wahl

Postal address: University of Basel, Department of Psychology, Division of Clinical Psychology and Epidemiology, Missionsstrasse 62a, 4055 Basel, Switzerland

Phone: (+41) 061 20 70654

E-Mail: karina.wahl@unibas.ch

Acknowledgments

We would like to thank all participants and other contributors of the present study.

Declarations of interest

Declarations of interest: none.

Abstract

Background and Objectives

The counter-app method is often used to assess the frequency of intrusions. The method requires the participants to press a button on a smartphone whenever an intrusion occurs during a predefined time period. We evaluated the convergent, discriminant, and predictive validity of the counter-app method in two studies.

Methods

In the first study, we assessed the frequency of intrusions with a counter-app method, thought-sampling method, and a retrospective measure in $N = 77$ students. Additionally, we assessed retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions. The second study ($N = 65$) was identical to the first except the thinking-aloud method replaced the thought-sampling method, and additionally we assessed behavioral neutralizing.

Results

The counter-app frequency was positively correlated with the convergent (thought-sampling and thinking-aloud frequencies, and retrospectively estimated frequencies of intrusions) and predictive (behavioral neutralizing) validity criteria. The correlations between counter-app frequency and discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions) were not smaller than the correlations between counter-app frequency and convergent validity criteria.

Limitations

We evaluated the predictive validity of the counter-app method using a criterion typical of obsessive-compulsive disorder research. Thus, the result for predictive validity might not transfer to other areas.

Conclusions

Our findings support convergent and predictive, but not discriminant, validity of the counter-app method. The counter-app method can validly assess the frequency of intrusions but is not appropriate if the study requires the differentiation of frequency from other constructs such as duration.

Keywords

Counter app, psychometric properties, validity, experimental study, intrusions

Introduction

Brief and sudden mental intrusions are commonplace in both clinical and nonclinical individuals, and are typically experienced as unwanted verbal thoughts, images, or impulses (Clark & Rhyno, 2005; Harvey et al., 2004; Purdon, 2005). Intrusions occur across many psychological disorders (e.g., Harvey et al., 2004; Purdon, 2005) as, for example, thoughts about harming others in psychotic patients (Morrison, 2005), contamination fears in obsessive-compulsive disorder (OCD; Abramowitz & Jacoby, 2014), or thoughts about a traumatic incident in posttraumatic stress disorder (PTSD; Falsetti et al., 2005). According to Purdon (2005), all intrusions, irrespective of their clinical (across all mental disorders) or nonclinical nature, share certain characteristics: they take attentional priority the moment they occur, are difficult to dismiss, and are perceived as unwanted. In addition, Clark and Rhyno (2005) asserted that intrusions are brief and sudden thoughts. One of the key aspects that differentiates clinical intrusions from their nonclinical counterparts is the frequency with which intrusions enter the conscious awareness. For example, findings have shown that individuals diagnosed with OCD or illness anxiety disorder experience intrusions more often than individuals with no diagnosis of mental disorder (e.g., Arnaez et al., 2021; Bouvard et al., 2017; Rachman & de Silva, 1978). Additionally, individuals at

high risk of suffering an eating disorder reported more intrusions than those at low risk (Belloch et al., 2016).

Researchers have employed various methodologies to assess the frequency of intrusions, such as self-report questionnaires and interview measures (Clark & Purdon, 1995). In experimental studies, a common method is event marking (e.g., Marcks & Woods, 2005, 2007; Najmi et al., 2009; Reynolds & Salkovskis, 1992) during which participants are typically instructed to record the occurrences of intrusions using a counter. A great advantage of this method is that the assessment takes place in real-time, which minimizes retrospective biases. In addition, it is straightforward to implement, requires little training, and allows simple collection of data. On the other hand, event marking might be subject to reactivity as participants' responses might be biased by various cognitive (e.g., active monitoring of the target thought) and behavioral (e.g., holding the device) processes involved in the methodology. Recently, Wahl et al. (2020) demonstrated that using a counter app on a smartphone (a contemporary event marking method) leads to overreporting of intrusions. Additionally, event marking might not appropriately account for long-lasting thoughts, since it assesses frequency but not duration of thoughts (Wahl et al., 2020).

Although event marking is typically used in experimental studies, surprisingly, little is known about its psychometric properties. Wahl et al. (2020) recently provided preliminary data about the validity and reliability of a counter-app method (installed on a smartphone) in assessment of intrusion frequency. The counter-app method had satisfactory convergent validity and satisfactory test–retest reliability, but rather poor discriminant validity, and inconclusive predictive validity. The study used retrospective measures to evaluate the convergent validity of the counter-app method. Although evidence supports retrospective measures as valid assessment tools of various constructs (e.g., Guan et al., 2006; Hardt & Rutter, 2004), they might be subject to

memory biases (Sato & Kawahara, 2011). Thus, employing other real-time measures of thought assessment to validate the counter-app method would further contribute to the evaluation of its psychometric properties.

Thought-sampling and thinking-aloud methods are alternative real-time assessments of intrusive thoughts (Smallwood & Schooler, 2015; Zanov & Davison, 2010). During the thought-sampling method, participants are typically probed several times during a predefined time period and asked to briefly report their thought content prior to each probe (Smallwood & Schooler, 2015). The thought-sampling method is widely used in mind wandering research (Smallwood & Schooler, 2006, 2015) and in research related to worry (e.g., Hirsch et al., 2013; Hirsch et al., 2015). Faber et al. (2018) and Steindorf and Rummel (2020) report findings that support the convergent validity of the thought-sampling method.

The thinking-aloud method requires participants to verbalize everything that is going through their minds during a predefined time period (Zanov & Davison, 2010). The articulated thoughts are recorded, transcribed, and later analyzed. Researchers have used the thinking-aloud method to examine thoughts related to various themes including anger and violence (Eckhardt, 2007), thought suppression (Erskine et al., 2017; Roemer & Borkovec, 1994), depression and anxiety (Molina et al., 1998), body dysmorphic disorder (Kollei & Martin, 2014), and eating disorders (Hilbert & Tuschen-Caffier, 2005). Studies provide evidence for the thinking-aloud method to have convergent validity (for discussion, see Davison et al., 1995; Zanov & Davison, 2010).

The main objective of this study is to further validate the counter-app method in assessment of intrusion frequency using two other real-time assessment methods as validation criteria. We provide information about the convergent, discriminant, and predictive validity of the method as evaluated in two independent studies with nearly identical designs. The studies differ in the method of real-time thought assessment

used as convergent validity criterion: the first study uses a thought-sampling method, whereas the second study uses a thinking-aloud method. Additionally, behavioral neutralizing is assessed at the end of the second study, as predictive validity criterion.

Study 1

Method

Participants

Undergraduate psychology students ($N = 77$) were recruited at the [blinded for review] to participate in the present study in exchange for course credit ($M_{age} = 22.84$, $SD = 5.82$; 80.52% female). To minimize the risk of substantial distress due to experimental priming of car-accident related intrusions, we excluded participants if they or a person close to them had ever experienced a car accident. The study was approved by the Ethics Committee of the Department of Psychology, [blinded for review].

Tasks

Activation and reactivation of intrusion

During the intrusion activation, participants were provided with a sheet of paper and a pen and asked to think of a loved person. Once they had a clear picture of that person in mind, they were instructed to write down and read aloud a sentence stating that they wished the loved person (whom they were instructed to name) would be involved in a horrible car accident that evening. Following this, the participants were told that any thoughts, images, or impulses related to the loved person being involved in a car accident would be referred to as the target intrusion throughout the study. Similar activation of intrusions has been successfully used in many studies (Rachman

et al., 1996; van den Hout et al., 2002; van den Hout et al., 2001). We reactivated the intrusion identically at the beginning of phase 2 (see Procedure Section below).

Thought task and methods

Focused-breathing task

Participants were asked to focus on their breathing during the counter-app method and the thought-sampling method (for more details, see both counter-app and thought-sampling methods below; adapted from Hirsch et al., 2013). This provided a similar background of cognitive activity in all participants against which the intrusive thoughts would occur.

Counter-app method

We used a counter app (FunCoolApps, 2016) on a smartphone to assess the frequency of intrusions (counter-app frequency). During the counter-app method, participants were instructed to press the volume-up button whenever the intrusion occurred, while focusing on their breathing. The smartphone display was covered to prevent visual feedback on the number of counts. Participants were provided with the following instruction on a computer screen:

During the next 5 minutes, I would like you to focus on your breathing. If your thoughts wander, try to refocus your attention back on your breathing again. However, if at any time you think of the target intrusive thought that you previously wrote down and read aloud, please record the occurrence of each thought by pressing the “+” key on the smartphone once for each occurrence.
(adapted from: Marcks & Woods, 2005; Wahl et al., 2020)

Thought-sampling method

In the thought-sampling method (adapted from: Hirsch et al., 2013; Hirsch et al., 2015), participants were prompted with a tone 12 times during the thought-sampling period (5 min) to indicate that they should report whether they were focused on their breathing or having other thoughts. If the latter, they used an answer sheet to categorize their thoughts as positive, neutral, or negative and provided one or two keywords to briefly describe the thoughts (e.g., positive – eating today). The participants received the answer sheet before the experimental session. At the end of the study, the participants were asked to elaborate on the keywords to describe what was going through their minds at that time. The participants' extended reports were recorded by the investigator using a smartphone. Later, three trained psychology students rated the participants' reports on whether the intrusion was experienced immediately before the presentation of the tone or not (thought-sampling frequency). The tones were created in E-Prime software (2016) and administered through Limesurvey (Questback GmbH, 2013). The distribution of tones across the thought-sampling period was randomly generated, then identical timing of tones was used for each participant. The instructions for the thought-sampling method, where "proceed as before" referred to the previously completed practice task, were as follows:

During the next 5 minutes, I would like you to focus on your breathing. If your thoughts wander, try to refocus your attention back on your breathing again. With each beep sound, please proceed as before and briefly document whether your thoughts were focused on your breathing, or if they were positive, negative, or neutral and provide one or two keywords to describe your thoughts, if necessary. For documentation, please use the answer sheet that is on the table. (adapted from: Hirsch et al., 2013; Hirsch et al., 2015)

Standardized questionnaires and other self-report measures

Standardized questionnaires

We administered the *Beck Depression Inventory-II* (BDI-II; Beck et al., 1996; German version: Hautzinger et al., 2006) and the *Beck Anxiety Inventory* (BAI; Beck et al., 1988; German version: Magraf & Ehlers, 2007) to assess depression and anxiety symptoms, respectively, the *Obsessive-Compulsive Inventory, Revised* (OCI-R; Foa et al., 2002; German version: Gönner et al., 2007) to assess obsessive-compulsive (OC) symptom severity, and the *Positive and Negative Affect Schedule* (PANAS; Watson et al., 1988; German version: Krohne et al., 1996) to assess current positive and negative affect. The psychometric properties of each scale are well established (Crawford & Henry, 2004; Gönner et al., 2007, 2008; Krohne et al., 1996; Magraf & Ehlers, 2007; Wang & Gorenstein, 2013).

Retrospectively estimated frequency, duration, percentage of time, intensity, and intrusiveness of the intrusions

Participants rated the frequency, duration, percentage of time, intensity, and intrusiveness of the intrusions directly after the counter-app or thought-sampling method as follows:

1. Frequency: "Please estimate (as accurately as possible) how many times thoughts about the car accident came to your mind during the last 5 minutes." Participants provided a number.
2. Duration: "Please indicate how long the target thoughts lasted on average" on a visual analogue scale (VAS) from 0 (*not at all/ only very briefly*) to 100 (*extremely long time*).
3. Percentage of time: "Please indicate during what percentage of time you were thinking of the target thoughts during the last 5 minutes" from 0 % (*no time at all*) to 100 % (*all the time*).

4. Intensity: "Please indicate the degree of *intensity* of your target thoughts during the last 5 minutes" from 0 (*not at all*) to 100 (*extreme*).
5. Intrusiveness: "Please indicate the degree of *intrusiveness* of your target thoughts during the last 5 minutes. Intrusiveness means that the target thoughts intruded into your mind with great force, persisted for a long time, occurred frequently, and possibly interrupted what you were thinking at that time" on a VAS ranging from 0 (*not at all*) to 100 (*extreme*).

The items were adapted from Wahl et al. (2020).

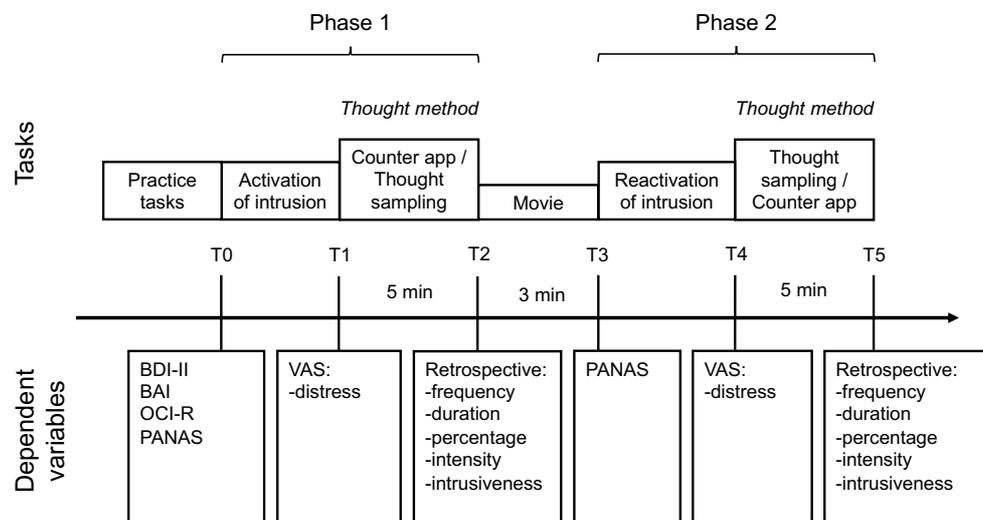
Manipulation check on the intrusion activation and reactivation

Participants rated their distress ("How distressed are you right now?") directly after the activation of the intrusion in phase 1 and directly after the reactivation of the intrusion in phase 2 on a VAS ranging from 0 (*not at all*) to 9 (*extremely high*), respectively. To consider the activation and reactivation of the intrusion successful, we expected the experienced distress to be comparable to that of other studies using the car accident sentence (e.g., Berman et al., 2010; Berman et al., 2012; Wahl et al., 2020).

Procedure

Figure 1 shows the experimental procedure of Study 1. Students gave their written, informed consent to participate in the study and were tested individually in a laboratory at the [blinded for review]. Prior to the experimental session, they were instructed on how to use the counter app that was installed on a study smartphone. In addition, they practiced the focused-breathing task, first alone and then combined with the thought-sampling method. Next, the participants completed standardized

measures of depression (BDI-II), anxiety (BAI), OC symptoms (OCI-R), and state affect (PANAS). This was followed by the activation of intrusion and immediately after that, the participants rated their level of distress (manipulation check). The study continued with one of two thought methods, depending on which experimental condition the participants were randomly assigned to. In the counter-app--thought-sampling (CA-TS) group ($n = 38$), participants were instructed to perform the counter-app method in phase 1 and the thought-sampling method in phase 2. In the thought-sampling--counter-app (TS-CA) group ($n = 39$), the order of the methods was reversed. The two experimental conditions orders were implemented to control for potential order effects. Directly after the first respective thought method, participants retrospectively rated the frequency, duration, percentage of time, intensity, and intrusiveness of the intrusions. Afterward, between phases 1 and 2, participants watched a short movie about visiting different places in Europe to reduce potential carry-over effects. Phase 2 started with the re-assessment of state affect (PANAS), continued with the reactivation of the intrusion, then with the measure of distress (manipulation check), followed by the counter-app or the thought-sampling method depending on which experimental group the participants were assigned to. Finally, phase 2 concluded with the retrospective measures of frequency, duration, percentage of time, intensity, and intrusiveness of the intrusions.

Figure 1*Experimental procedure of Study 1*

Note. BAI = Beck Anxiety Inventory; BDI-II = Beck Depression Inventory-II; OCI-R = Obsessive-Compulsive Inventory, Revised; PANAS = Positive and Negative Affect Schedule; VAS = visual analogue scale.

Statistical analysis

To evaluate the convergent and discriminant validity of the counter-app method, we calculated Pearson's correlation coefficients between the counter-app frequency and validation criteria (i.e. thought-sampling frequency and retrospectively estimated frequency, duration, percentage of time, intensity, and intrusiveness of intrusions). Convergent validity of a measure is defined as high correlation of the measure with other measures that assess the same construct (DeVon et al., 2007; Heale & Twycross, 2015). Many authors have argued that it is not possible to define a single appropriate cut-off for convergent validity (for discussion, see DeVon et al., 2007; Post, 2016). Thus, in this study positive correlations of the counter-app frequency with the thought-sampling frequency and retrospectively estimated frequency of intrusions were interpreted as indicators of convergent validity for the counter-app method.

Discriminant validity of a measure is defined as low correlation of the measure with other measures that assess different constructs (DeVon et al., 2007), yet with possibly overlapping features (Drost, 2011). Similar to convergent validity, defining a cut-off for discriminant validity is difficult (e.g., some authors suggest values of $< .30$, others $\leq .45$; DeVon et al., 2007; Post, 2016). In this study, indicators of discriminant validity for the counter-app method were demonstrated through smaller correlations (compared to those reflecting convergent validity) of the counter-app frequency with different but similar constructs: retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions. To examine the statistical significance of and differences between the correlations, we reported confidence intervals. A correlation coefficient with a confidence interval that does not include a zero was interpreted as significant. Confidence intervals of correlation coefficients that do not overlap were interpreted as statistically different.

Prior to correlation analysis, we detected outliers in the counter-app frequency, thought-sampling frequency, and retrospectively-estimated-intrusion frequency. To adjust for this we log transformed the values of these variables (Field, 2013). In addition, to account for potential order effects, we analyzed whether the correlations differed between the CA-TS and TS-CA groups and confirmed no such differences (see Appendix, Table 5). Thus, we combined the groups for the main analyses. We assessed the inter-rater reliability of the thought-sampling method with a two-way mixed, absolute agreement, average-measures intra-class correlation (ICC), per Hallgren (2012), and interpreted the ICC using the common cutoffs (Cicchetti, 1994).

Results

Sample characteristics, internal consistencies, inter-rater-reliability, and manipulation check

Table 1 shows the sample characteristics, internal consistencies of the standardized questionnaires, and manipulation check in Study 1. Depression, anxiety, and OC symptoms in this study seemed comparable to those of other studies with university students (e.g., Creamer et al., 1995; Grisham & Williams, 2009; Wang & Gorenstein, 2013). Participants experienced similar levels of positive affect directly before the intrusion activation and reactivation ($t(76) = -0.49, p = .623, d = 0.05$). Negative affect increased from phase 1 to phase 2 ($t(76) = -2.43, p = .018, d = 0.33$) but seemed still comparable to negative affect reported in other studies with a student sample (Rogatko, 2009) and with mixed samples (Terracciano et al., 2003; Watson et al., 1988). In this study, the internal consistencies of the standardized questionnaires ranged from acceptable (PANAS, negative affect subscale in phase 1) to excellent (e.g., BDI-II). The inter-rater reliability of the thought-sampling method was excellent, $ICC = 1.00$. The average distress experienced in phase 1 and phase 2 appeared comparable to that of other studies that used the car accident sentence paradigm to induce intrusions (e.g., Berman et al., 2010; Berman et al., 2012; Wahl et al., 2020). This suggests that the activation and reactivation of intrusion was successful.

Table 1

Sample characteristics, Cronbach's alpha, and manipulation check in Study 1

Variable	<i>M (SD)</i>	α
Depression (BDI-II)	7.74 (7.26)	.91
Anxiety (BAI)	8.61 (6.71)	.85
OC symptoms (OCI-R)	10.52 (9.23)	.90
PANAS		
Positive affect, phase 1	28.10 (6.53)	.86
Positive affect, phase 2	28.48 (8.06)	.91
Negative affect, phase 1	13.43 (3.65)	.76
Negative affect, phase 2	15.08 (6.00)	.88

Variable	<i>M (SD)</i>	α
Manipulation check		
Distress, phase 1	66.04 (22.50)	
Distress, phase 2	48.51 (29.00)	

Note. BAI = Beck Anxiety Inventory; BDI-II = Beck Depression Inventory-II; OC = obsessive-compulsive; OCI-R = Obsessive-Compulsive Inventory, Revised; PANAS = Positive and Negative Affect Schedule.

Convergent and discriminant validity of the counter-app method

Table 2 shows correlations of the counter-app frequency with the criteria of convergent and discriminant validity in Study 1. The counter-app frequency was positively correlated with the convergent validity criteria (thought-sampling frequency and retrospectively estimated frequency of intrusions). The confidence intervals show that the correlations of the counter-app frequency with the discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions) were not smaller than the correlations of the counter-app frequency with the convergent validity criteria.

Table 2

Convergent and discriminant validity of the counter-app method (Pearson correlations) in Study 1

Variable	Counter-app frequency	
	<i>r</i>	95% CI
Convergent validity		
Thought-sampling frequency	.38	[.171, .557]
Retrospectively estimated frequency	.32	[.103, .508]

Discriminant validity

Retrospectively estimated

Duration	.18	[-.046, .388]
Percentage of time	.22	[-.004, .423]
Intensity	.21	[-.015, .414]
Intrusiveness	.31	[.092, .499]

Note. CI = confidence interval.

Study 2**Method****Participants**

Undergraduate psychology students ($N = 65$, $M_{\text{age}} = 22.63$, $SD = 4.89$; 73.85% female) recruited at the [blinded for review] participated in the present study in exchange for course credit. Inclusion criteria were identical to those of Study 1. The Ethics Committee of the Department of Psychology, [blinded for review], approved the study.

Tasks

The activation and reactivation of intrusion, focused-breathing task, and counter-app method were identical to that of Study 1.

Thinking-aloud method

During the thinking-aloud method, participants were asked to “think aloud” (verbalize their stream of thoughts), while simultaneously focusing on their breathing for 5 min. The investigator was not present during the thinking-aloud method. Instructions for participants were presented on a computer screen:

During the next 5 min, I would like you to focus on your breathing. If your thoughts wander, try to refocus your attention back on your breathing again.

At the same time, I would like you to describe aloud everything that is going through your mind, from each moment to moment... Try to talk continuously, and don't be concerned about grammar or complete sentences. (adapted from: Molina et al., 1998)

Each participant's articulated stream of thoughts was recorded with a smartphone and later transcribed for further analysis. Both the first (MK) and the third (CVH) authors of this study independently rated the transcripts using a coding scheme on how many times the intrusion occurred (thinking-aloud frequency). The coding scheme, developed by the first author, is available in the Appendix. Subsequently, we calculated ICC between the two ratings to determine the inter-rater reliability (see Results section below).

Standardized questionnaires and other self-report measures

All standardized questionnaires and measures assessing retrospectively estimated frequency, duration, percentage of time, intensity, intrusiveness of the intrusions, and distress (manipulation check) were identical to those of Study 1.

Behavioral neutralizing

The investigator assessed behavioral neutralizing at the end of the study using a scale consisting of 11 items modified from previous work on neutralizing (Freeston et al., 1991; Goodman et al., 1989; Kollárik et al., 2020; Purdon & Clark, 1993, 1994; Rachman et al., 1996). Prior to the behavioral neutralizing assessment, the participants were given the car accident sentence written on paper (see activation and reactivation of intrusion in the previous section) and told that—if they wished—they could use the opportunity to neutralize the target thought by, for example, ripping the paper or

changing the name in the sentence. To avoid potential social desirability effects, the investigator was not present during the time that the participants could neutralize. Afterward, the investigator returned and asked whether they had done something to the paper with the written sentence used to provoke intrusions and noted each neutralizing behavior using the scale. Each item on the scale represents a different behavioral neutralizing strategy with two possible answers: 1 (*strategy occurred*) or 0 (*strategy did not occur*). The investigators operationalized behavioral neutralizing as the sum of all 11 items.

Procedure

The experimental procedure of Study 2 was identical to that of Study 1, except the thinking-aloud method replaced the thought-sampling method. Participants were randomly assigned to the counter-app--thinking-aloud (CA-TA) group ($n = 33$) or the thinking-aloud--counter-app (TA-CA) group ($n = 32$). At the beginning of the experiment, participants practiced the thinking-aloud method, first alone and then combined with the focused-breathing task. At the end of the study, we additionally assessed behavioral neutralizing.

Statistical analysis

The convergent and discriminant validity of the counter-app method in Study 2 was evaluated the same way as in Study 1, except the thinking-aloud method replaced the thought-sampling. Thus, positive correlation of the counter-app frequency with the thinking-aloud frequency would indicate convergent validity. We additionally evaluated the predictive validity of the counter-app method. Predictive validity of a measure is defined as high correlation of the measure with future criteria (DeVon et al., 2007; Heale & Twycross, 2015). As with convergent and discriminant validity, there are no

universal benchmarks for predictive validity (DeVon et al., 2007). In this study, a positive correlation between the counter-app frequency and subsequent behavioral neutralizing was interpreted as an indicator of predictive validity, since in the context of OCD research neutralizing behaviors are functionally related to the occurrence of intrusions (Rachman et al., 1996; Salkovskis, 1985). Visual inspection of the data showed outliers in the counter-app frequency, so we log transformed this data. As in Study 1, we did not confirm differences in correlations between the CA-TA and TA-CA randomization orders (see Appendix, Table 6). This indicates that order effects are unlikely, and we conducted correlation analysis on both groups combined. The calculation of the inter-rater reliability of the thinking-aloud method was identical to that of the thought-sampling method.

Results

Sample characteristics, internal consistencies, inter-rater reliability, and manipulation check

Table 3 shows the sample characteristics, internal consistencies of the standardized questionnaires, and manipulation check in Study 2. Anxiety, depression, and OC symptoms reported in this study appeared comparable to those reported in other studies with student samples (e.g., Creamer et al., 1995; Grisham & Williams, 2009; Wang & Gorenstein, 2013). Positive and negative affect in phase 1 did not differ from positive and negative affect in phase 2 (positive affect: $t(64) = -0.95$, $p = .345$, $d = 0.10$; negative affect: $t(64) = 0.65$, $p = .520$, $d = 0.08$). This finding suggests that carry-over effects from phase 1 on phase 2 were unlikely. The standardized questionnaires in this study demonstrated internal consistencies ranging from good (e.g., PANAS, negative affect subscale in phase 2) to excellent (e.g., BAI). The inter-rater reliability of the thinking-aloud method was excellent, $ICC = .95$. This suggests

that the two ratings were nearly identical and thus, we decided to use the first author's (MK) ratings for the subsequent analysis of thinking-aloud frequency. The experienced distress associated with the intrusion was comparable to that of other studies that employed similar intrusion activation (e.g., Berman et al., 2010; Berman et al., 2012; Wahl et al., 2020).

Table 3

Sample characteristics, Cronbach's alpha, and manipulation check in Study 2

Variable	<i>M (SD)</i>	α
Depression (BDI-II)	11.77 (7.57)	.89
Anxiety (BAI)	11.18 (10.33)	.93
OC symptoms (OCI-R)	12.62 (10.78)	.90
PANAS		
Positive affect, phase 1	27.86 (7.27)	.88
Positive affect, phase 2	28.65 (8.38)	.89
Negative affect, phase 1	15.92 (6.17)	.88
Negative affect, phase 2	15.49 (5.03)	.84
Manipulation check		
Distress, phase 1	61.91 (29.88)	
Distress, phase 2	43.30 (28.75)	

Note. BAI = Beck Anxiety Inventory; BDI-II = Beck Depression Inventory-II; OC = obsessive-compulsive; OCI-R = Obsessive-Compulsive Inventory, Revised; PANAS = Positive and Negative Affect Schedule.

Convergent, discriminant, and predictive validity of the counter-app method

Table 4 shows correlations of the counter-app frequency with convergent, discriminant, and predictive validity criteria in Study 2. The counter-app frequency was positively correlated with convergent (thinking-aloud frequency and retrospectively estimated frequency of intrusions) and predictive validity (behavioral neutralizing)

criteria. The correlations of the counter-app frequency with the discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions) were not smaller than the correlations between the counter-app frequency and convergent validity criteria.

Table 4

Convergent, discriminant, and predictive validity of the counter-app method (Pearson correlations) in Study 2

Variable	Counter-app frequency	
	<i>r</i>	95% CI
Convergent validity		
Thinking-aloud frequency ^a	.25	[.002, .469]
Retrospectively estimated frequency	.64	[.469, .765]
Discriminant validity		
Retrospectively estimated		
Duration	.43	[.208, .610]
Percentage of time	.60	[.417, .736]
Intensity	.52	[.316, .678]
Intrusiveness	.56	[.366, .707]
Predictive validity (OCD-context)		
Behavioral neutralizing	.27	[.028, .482]

Note. OCD = obsessive-compulsive disorder. CI = confidence interval.

^a Two participants had missing values.

Discussion

The current study aimed to further evaluate the validity of the counter-app method in assessing the frequency of intrusions using thought-sampling and thinking-

aloud methods as validation criteria. The counter-app frequency was positively associated with both the thought-sampling and thinking-aloud intrusion frequencies. Thus, the counter-app method seems to assess the real-time frequency of intrusions in a valid way. In addition, and consistent with the Wahl et al. (2020) study, the counter-app frequency was positively associated with the retrospective estimations of intrusion. In brief, our findings provide indications of convergent validity for the counter-app method.

In both studies, the associations of counter-app frequency with the discriminant validity criteria (retrospectively estimated duration, percentage of time, intensity, and intrusiveness of intrusions) were not smaller than the associations of counter-app frequency with the convergent validity criteria (thought-sampling and thinking-aloud frequencies, and retrospectively estimated frequencies of intrusions). Thus, the counter-app method is not an appropriate methodology if the discrimination of frequency from duration, percentage of time, intensity, and intrusiveness is a focus of research. The results on discriminant validity of the counter-app method reflect those of Wahl et al. (2020). Finally, the counter-app frequency was positively associated with future behavioral neutralizing, indicating predictive validity of the counter-app method in the context of OCD-related research. This is partially consistent with Wahl et al. (2020), who reported some evidence of predictive validity for the counter-app method.

There are several limitations of the present research. First, to evaluate the validity of counter-app method, we focused on correlations but did not address whether the counter-app method led to overreporting of intrusions, as indicated in Wahl et al. (2020). Second, to activate the intrusion, we used a paradigm that is typical of OCD-related research. However, as mentioned in the Introduction Section, intrusions across mental disorders and also in non-clinical samples are commonly unwanted, brief, and sudden, take attentional priority and are hard to dismiss (Clark & Rhyno, 2005; Purdon,

2005). Thus, we believe that our findings may also be relevant to intrusions that share these aspects, though this needs to be investigated in future research. Third, to evaluate the predictive validity of the counter-app method, we used a validation criterion typical of OCD-related research (behavioral neutralizing). Hence, this finding might not generalize to research related to other areas, such as PTSD or illness anxiety. Finally, our findings might not generalize beyond our sample to other populations, so future research should examine utility of the counter-app method in other samples, for example, in those diagnosed with OCD.

To conclude, our findings support convergent and predictive, but not discriminant, validity of the counter-app method. We recommend the use of the counter-app method in experimental studies to assess changes in the real-time frequency of intrusions in short time intervals. For example, recent evidence suggests that there is an association between maladaptive daydreaming (extensive fantasy activity that interferes with functioning; Bigelsen et al., 2016; Somer, 2002) and OC symptoms (Salomon-Small et al., 2021). Researchers could extend existing studies by using the counter-app method to examine frequency of intrusions (e.g., in a 5 min time interval) prior to and post daydreaming, to investigate the causal effects of daydreaming on intrusions. The counter-app method might be less appropriate for studies that aim to differentiate frequency from other constructs such as duration, for example, when examining whether daydreaming differentially affects the frequency and duration of intrusions. Future studies should investigate whether the validity of the counter-app method varies with the type of intrusion (e.g., intrusions related to traumatic incident in the context of PTSD or food-related intrusions in the context of eating disorders; Belloch et al., 2016; Harvey et al., 2004).

References

- Abramowitz, J. S., & Jacoby, R. J. (2014). Obsessive-compulsive and related disorders. In D. H. Barlow (Ed.), *The Oxford handbook of clinical psychology* (pp. 313-338). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199328710.001.0001>
- Arnaez, S., Garcia-Soriano, G., Lopez-Santiago, J., & Belloch, A. (2021). Illness-related intrusive thoughts and illness anxiety disorder. *Psychology and psychotherapy*, 94(1), 63–80. <https://doi.org/10.1111/papt.12267>
- Beck, A. T., Brown, G., Epstein, N., & Steer, R. A. (1988). An inventory for measuring clinical anxiety - psychometric properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893-897. <https://doi.org/10.1037/0022-006x.56.6.893>
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the beck depression inventory-II*. Psychological Corporation.
- Belloch, A., Roncero, M., & Perpina, C. (2016). Obsessional and Eating Disorder-related Intrusive Thoughts: Differences and Similarities Within and Between Individuals Vulnerable to OCD or to EDs. *European Eating Disorders Review*, 24(6), 446-454. <https://doi.org/10.1002/erv.2458>
- Berman, N. C., Abramowitz, J. S., Pardue, C. M., & Wheaton, M. G. (2010). The relationship between religion and thought-action fusion: use of an in vivo paradigm. *Behaviour Research and Therapy*, 48(7), 670-674. <https://doi.org/10.1016/j.brat.2010.03.021>
- Berman, N. C., Wheaton, M. G., & Abramowitz, J. S. (2012). The "Arnold Schwarzenegger Effect": is strength of the "victim" related to misinterpretations of harm intrusions? *Behaviour Research and Therapy*, 50(12), 761-766. <https://doi.org/10.1016/j.brat.2012.09.002>
- Bigelsen, J., Lehrfeld, J. M., Jopp, D. S., & Somer, E. (2016). Maladaptive daydreaming: Evidence for an under-researched mental health disorder. *Consciousness and Cognition*, 42, 254-266. <https://doi.org/10.1016/j.concog.2016.03.017>
- Bouvard, M., Fournet, N., Denis, A., Sixdenier, A., & Clark, D. (2017). Intrusive thoughts in patients with obsessive compulsive disorder and non-clinical participants: a comparison using the International Intrusive Thought Interview Schedule. *Cognitive Behaviour Therapy*, 46(4), 287-299. <https://doi.org/10.1080/16506073.2016.1262894>

- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Clark, D. A., & Purdon, C. L. (1995). The assessment of unwanted intrusive thoughts: a review and critique of the literature. *Behaviour Research and Therapy*, 33(8), 967-976. [https://doi.org/10.1016/0005-7967\(95\)00030-2](https://doi.org/10.1016/0005-7967(95)00030-2)
- Clark, D. A., & Rhyno, S. (2005). Unwanted intrusive thoughts in nonclinical individuals: Implications for clinical disorders. In D. A. Clark (Ed.), *Intrusive thoughts in clinical disorders: Theory, research, and treatment* (pp. 1-29). Guilford Press.
- Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43, 245-265. <https://doi.org/10.1348/0144665031752934>
- Creamer, M., Foran, J., & Bell, R. (1995). The Beck Anxiety Inventory in a nonclinical sample. *Behaviour Research and Therapy*, 33(4), 477-485. [https://doi.org/10.1016/0005-7967\(94\)00082-U](https://doi.org/10.1016/0005-7967(94)00082-U)
- Davison, G. C., Navarre, S. G., & Vogel, R. S. (1995). The Articulated Thoughts in Simulated Situations Paradigm - a Think-Aloud Approach to Cognitive Assessment. *Current Directions in Psychological Science*, 4(1), 29-33. <https://doi.org/10.1111/1467-8721.ep10770963>
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, 39(2), 155-164. <https://doi.org/10.1111/j.1547-5069.2007.00161.x>
- Drost, E. (2011). Validity and Reliability in Social Science Research. *Education research and perspectives*, 38(1), 105-123.
- Eckhardt, C. I. (2007). Effects of alcohol intoxication on anger experience and expression among partner assaultive men. *Journal of Consulting and Clinical Psychology*, 75(1), 61-71. <https://doi.org/10.1037/0022-006X.75.1.61>
- Erskine, J. A. K., Georgiou, G. J., Joshi, M., Deans, A., & Colegate, C. (2017). Ageing and thought suppression performance: Its relationship with working memory capacity, habitual thought suppression and mindfulness. *Consciousness and Cognition*, 53, 211-221. <https://doi.org/10.1016/j.concog.2017.06.019>

- Faber, M., Bixler, R., & D'Mello, S. K. (2018). An automated behavioral measure of mind wandering during computerized reading. *Behavior Research Methods*, *50*(1), 134-150. <https://doi.org/10.3758/s13428-017-0857-y>
- Falsetti, S. A., Monnier, J., & Resnick, H. S. (2005). Intrusive thoughts in posttraumatic stress disorder. In D. A. Clark (Ed.), *Intrusive thoughts in clinical disorders: Theory, research, and treatment* (pp. 30-53). Guilford Press.
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). SAGE Publications Ltd.
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment*, *14*(4), 485-496. <https://doi.org/10.1037//1040-3590.14.4.485>
- Freeston, M. H., Ladouceur, R., Thibodeau, N., & Gagnon, F. (1991). Cognitive intrusions in a non-clinical population. I. Response style, subjective experience, and appraisal. *Behaviour Research and Therapy*, *29*(6), 585-597. <https://www.ncbi.nlm.nih.gov/pubmed/1759957>
- FunCoolApps. (2016). *Click Counter* (Version 1.4) [Mobile app]. Google Play Store. <https://play.google.com/store/apps/details?id=com.useless.counter>
- Gönner, S., Leonhart, R., & Ecker, W. (2007). Das Zwangsinventar OCI-R - die Deutsche Version des Obsessive-Compulsive Inventory-Revised: Ein kurzes Selbstbeurteilungsinstrument zur mehrdimensionalen Messung von Zwangssymptomen [The German version of the Obsessive-Compulsive Inventory-Revised: A brief self-report measure for the multidimensional assessment of obsessive-compulsive symptoms]. *Psychotherapie Psychosomatik Medizinische Psychologie*, *57*(9-10), 395-404. <https://doi.org/10.1055/s-2007-970894>
- Gönner, S., Leonhart, R., & Ecker, W. (2008). The Obsessive-Compulsive Inventory-Revised (OCI-R): Validation of the German version in a sample of patients with OCD, anxiety disorders, and depressive disorders. *Journal of Anxiety Disorders*, *22*(4), 734-749. <https://doi.org/10.1016/j.janxdis.2007.07.007>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry*, *46*(11), 1006-1011. <https://doi.org/10.1001/archpsyc.1989.01810110048007>

- Grisham, J. R., & Williams, A. D. (2009). Cognitive control of obsessional thoughts. *Behaviour Research and Therapy*, 47(5), 395-402. <https://doi.org/10.1016/j.brat.2009.01.014>
- Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). *The validity of the stimulated retrospective think-aloud method as measured by eye tracking*. CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal Québec Canada. <https://doi.org/10.1145/1124772.1124961>
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23-34. <https://doi.org/10.20982/tqmp.08.1.p023>
- Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *Journal of child psychology and psychiatry, and allied disciplines*, 45(2), 260-273. <https://doi.org/10.1111/j.1469-7610.2004.00218.x>
- Harvey, A., Watkins, E., Mansell, W., & Shafran, R. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press. <https://doi.org/10.1093/med:psych/9780198528883.001.0001>
- Hautzinger, M., Keller, F., & Kühner, C. (2006). *Das Beck Depressionsinventar II. Deutsche Bearbeitung und Handbuch zum BDI II*. Harcourt Test Services.
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66-67. <https://doi.org/10.1136/eb-2015-102129>
- Hilbert, A., & Tuschen-Caffier, B. (2005). Body-related cognitions in binge-eating disorder and bulimia nervosa. *Journal of Social and Clinical Psychology*, 24(4), 561-579. <https://doi.org/10.1521/jscp.2005.24.4.561>
- Hirsch, C. R., Mathews, A., Lequertier, B., Perman, G., & Hayes, S. (2013). Characteristics of worry in Generalized Anxiety Disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 44(4), 388-395. <https://doi.org/10.1016/j.jbtep.2013.03.004>
- Hirsch, C. R., Perman, G., Hayes, S., Eagleson, C., & Mathews, A. (2015). Delineating the Role of Negative Verbal Thinking in Promoting Worry, Perceived Threat, and Anxiety. *Clinical Psychological Science*, 3(4), 637-647. <https://doi.org/10.1177/2167702615577349>

- Kollárik, M., van den Hout, M., Heinzl, C. V., Hofer, P. D., Lieb, R., & Wahl, K. (2020). Effects of rumination on unwanted intrusive thoughts: A replication and extension. *Journal of Experimental Psychopathology*, 11(1). <https://doi.org/10.1177/2043808720912583>
- Kollei, I., & Martin, A. (2014). Body-related cognitions, affect and post-event processing in body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 45(1), 144-151. <https://doi.org/10.1016/j.jbtep.2013.09.005>
- Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Investigations with a German version of the positive and negative affect schedule (PANAS). *Diagnostica*, 42(2), 139-156. <https://doi.org/10.1037/t49650-000>
- Magraf, J., & Ehlers, A. (2007). *Beck Angst-Inventar (BAI) - Manual*. Harcourt Test Services GmbH.
- Marcks, B. A., & Woods, D. W. (2005). A comparison of thought suppression to an acceptance-based technique in the management of personal intrusive thoughts: a controlled evaluation. *Behaviour Research and Therapy*, 43(4), 433-445. <https://doi.org/10.1016/j.brat.2004.03.005>
- Marcks, B. A., & Woods, D. W. (2007). Role of thought-related beliefs and coping strategies in the escalation of intrusive thoughts: An analog to obsessive-compulsive disorder. *Behaviour Research and Therapy*, 45(11), 2640-2651. <https://doi.org/10.1016/j.brat.2007.06.012>
- Molina, S., Borkovec, T. D., Peasley, C., & Person, D. (1998). Content analysis of worrisome streams of consciousness in anxious and dysphoric participants. *Cognitive Therapy and Research*, 22(2), 109-123. <https://doi.org/10.1023/A:1018772104808>
- Morrison, A. P. (2005). Psychosis and the phenomenon of unwanted intrusive thoughts. In D. A. Clark (Ed.), *Intrusive thoughts in clinical disorders: Theory, research, and treatment* (pp. 175-198). Guilford Press.
- Najmi, S., Riemann, B. C., & Wegner, D. M. (2009). Managing unwanted intrusive thoughts in obsessive-compulsive disorder: Relative effectiveness of suppression, focused distraction, and acceptance. *Behaviour Research and Therapy*, 47(6), 494-503. <https://doi.org/10.1016/j.brat.2009.02.015>
- Post, M. W. (2016). What to Do With "Moderate" Reliability and Validity Coefficients? *Arch Phys Med Rehabil*, 97(7), 1051-1052. <https://doi.org/10.1016/j.apmr.2016.04.001>

- Psychology Software Tools, Inc. (2016). *E-Prime* (Version 3.0) [Computer software].
<https://support.pstnet.com/>
- Purdon, C. (2005). Unwanted intrusive thoughts: Present status and future directions. In D. A. Clark (Ed.), *Intrusive thoughts in clinical disorders: Theory, research, and treatment* (pp. 226-244). Guilford Press.
- Purdon, C., & Clark, D. A. (1993). Obsessive intrusive thoughts in nonclinical subjects. Part I. Content and relation with depressive, anxious and obsessional symptoms. *Behaviour Research and Therapy*, 31(8), 713-720.
[https://doi.org/10.1016/0005-7967\(93\)90001-B](https://doi.org/10.1016/0005-7967(93)90001-B)
- Purdon, C., & Clark, D. A. (1994). Obsessive intrusive thoughts in nonclinical subjects. Part II. Cognitive appraisal, emotional response and thought control strategies. *Behaviour Research and Therapy*, 32(4), 403-410.
<https://www.ncbi.nlm.nih.gov/pubmed/8192639>
- Questback GmbH. (2013). *Unipark* [Survey software]. <https://www.unipark.com/>
- Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy*, 16(4), 233-248. [https://doi.org/10.1016/0005-7967\(78\)90022-0](https://doi.org/10.1016/0005-7967(78)90022-0)
- Rachman, S., Shafran, R., Mitchell, D., Trant, J., & Teachman, B. (1996). How to remain neutral: An experimental analysis of neutralization. *Behaviour Research and Therapy*, 34(11-12), 889-898. [https://doi.org/10.1016/S0005-7967\(96\)00051-4](https://doi.org/10.1016/S0005-7967(96)00051-4)
- Reynolds, M., & Salkovskis, P. M. (1992). Comparison of positive and negative intrusive thoughts and experimental investigation of the differential effects of mood. *Behaviour Research and Therapy*, 30(3), 273-281.
[https://doi.org/10.1016/0005-7967\(92\)90073-p](https://doi.org/10.1016/0005-7967(92)90073-p)
- Roemer, L., & Borkovec, T. D. (1994). Effects of Suppressing Thoughts About Emotional Material. *Journal of Abnormal Psychology*, 103(3), 467-474.
<https://doi.org/10.1037/0021-843x.103.3.467>
- Rogatko, T. P. (2009). The Influence of Flow on Positive Affect in College Students. *Journal of Happiness Studies*, 10(2), 133-148. <https://doi.org/10.1007/s10902-007-9069-y>
- Salkovskis, P. M. (1985). Obsessional-compulsive problems - a cognitive-behavioral analysis. *Behaviour Research and Therapy*, 23(5), 571-583.
[https://doi.org/10.1016/0005-7967\(85\)90105-6](https://doi.org/10.1016/0005-7967(85)90105-6)

- Salomon-Small, G., Somer, E., Harel-Schwarzmann, M., & Soffer-Dudek, N. (2021). Maladaptive daydreaming and obsessive-compulsive symptoms: A confirmatory and exploratory investigation of shared mechanisms. *Journal of Psychiatric Research*, 136, 343-350. <https://doi.org/10.1016/j.jpsychires.2021.02.017>
- Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety Stress Coping*, 24(4), 359-367. <https://doi.org/10.1080/10615806.2010.543132>
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132(6), 946-958. <https://doi.org/10.1037/0033-2909.132.6.946>
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66, 487-518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Sommer, E. (2002). Maladaptive daydreaming: A qualitative inquiry. *Journal of Contemporary Psychotherapy*, 32, 197-212. <https://doi.org/10.1023/A:1020597026919>
- Steindorf, L., & Rummel, J. (2020). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods*, 52(1), 162-176. <https://doi.org/10.3758/s13428-019-01214-4>
- Terracciano, A., McCrae, R. R., & Costa, P. T. (2003). Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). *European Journal of Psychological Assessment*, 19(2), 131-141. <https://doi.org/10.1027//1015-5759.19.2.131>
- van den Hout, M., Kindt, M., Weiland, T., & Peters, M. (2002). Instructed neutralization, spontaneous neutralization and prevented neutralization after an obsession-like thought. *Journal of Behavior Therapy and Experimental Psychiatry*, 33(3-4), 177-189. [https://doi.org/10.1016/S0005-7916\(02\)00048-4](https://doi.org/10.1016/S0005-7916(02)00048-4)
- van den Hout, M., van Pol, M., & Peters, M. (2001). On becoming neutral: effects of experimental neutralizing reconsidered. *Behaviour Research and Therapy*, 39(12), 1439-1448. [https://doi.org/10.1016/S0005-7967\(00\)00109-1](https://doi.org/10.1016/S0005-7967(00)00109-1)
- Wahl, K., Lieb, R., Kollárik, M., & Purdon, C. (2020). The appropriateness of using a counter app in experimental studies assessing unwanted intrusive thoughts.

- Journal of Obsessive-Compulsive and Related Disorders*, 26.
<https://doi.org/10.1016/j.jocrd.2020.100550>
- Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Braz J Psychiatry*, 35(4), 416-431.
<https://doi.org/10.1590/1516-4446-2012-1048>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect - the Panas Scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
<https://doi.org/10.1037/0022-3514.54.6.1063>
- Zanov, M. V., & Davison, G. C. (2010). A Conceptual and Empirical Review of 25 Years of Cognitive Assessment Using the Articulated Thoughts in Simulated Situations (ATSS) Think-Aloud Paradigm. *Cognitive Therapy and Research*, 34(3), 282-291. <https://doi.org/10.1007/s10608-009-9271-9>

Appendix

Coding scheme for the thinking-aloud method

We determined 21 keywords that were relevant to the activation of intrusion: car, car accident, accident, sentence, person, name of the person, he, him, his, she, her, hers, thought, think, write, written, say, read, happen, imagine, and image. Two raters used the search function in Word to detect the keywords in the transcripts and rated them on whether they were related to the activated intrusion or not. If yes, the keywords were counted as intrusion occurrences. Additional rating rules were: One keyword (related to the activated intrusion) in one sentence counted as one intrusion occurrence, while multiple keywords (related to the activated intrusion) in one sentence also only counted as one intrusion occurrence. In some transcripts, only commas separated the phrases (e.g., “I am eating, what happened, was that a bird, is it late already...”) so periods could not be used to distinguish sentences. In these cases,

consecutive keywords counted as a single intrusion, and keywords separated by non-keyword phrases counted separately.

Table 5

Convergent and discriminant validity of the counter-app method (Pearson correlations) separately for experimental groups in Study 1

Variable	Order of thought methods			
	CA-TS group		TS-CA group	
	Counter-app frequency		Counter-app frequency	
	<i>r</i>	95% CI	<i>r</i>	95% CI
Convergent validity				
Thought-sampling frequency	.499	[.213, .706]	.252	[-.069, .526]
Retrospectively estimated frequency	.168	[-.160, .463]	.420	[.120, .649]
Discriminant validity				
Retrospectively estimated				
Duration	.082	[-.244, .391]	.276	[-.043, .544]
Percentage of time	.124	[-.204, .427]	.309	[-.007, .569]
Intensity	.016	[-.305, .334]	.368	[.059, .612]
Intrusiveness	.117	[-.211, .421]	.460	[.169, .677]

Note. Prior to the analysis, due to outliers, we log transformed the values of counter-app frequency and retrospectively estimated frequency of intrusions in both experimental groups, thought-sampling frequency in counter-app--thought-sampling (CA-TS) group, and retrospectively estimated percentage of time in thought-sampling-counter-app (TS-CA) group. CI = confidence interval.

Table 6

Convergent, discriminant, and predictive validity of the counter-app method (Pearson correlations) separately for experimental groups in Study 2

Variable	Order of thought methods			
----------	--------------------------	--	--	--

	CA-TA group		TA-CA group	
	Counter-app frequency		Counter-app frequency	
	<i>r</i>	95% CI	<i>r</i>	95% CI
Convergent validity				
Thinking-aloud frequency ^a	.438	[.105, .682]	.077	[-.285, .420]
Retrospectively estimated frequency	.614	[.343, .791]	.685	[.442, .834]
Discriminant validity				
Retrospectively estimated				
Duration	.493	[.180, .715]	.409	[.070, .663]
Percentage of time	.628	[.363, .799]	.588	[.301, .777]
Intensity	.549	[.253, .751]	.458	[.130, .696]
Intrusiveness	.610	[.337, .788]	.524	[.214, .738]
Predictive validity (OCD-context)				
Behavioral neutralizing	.386	[.049, .644]	.134	[-.225, .461]

Note. Prior to the analysis, due to outliers, we log transformed the values of counter-app frequency in both counter-app--thinking-aloud (CA-TA) and thinking-aloud--counter-app (TA-CA) groups. CI = confidence interval; OCD = obsessive-compulsive disorder.

^a For thinking-aloud frequency, we had one missing value in each experimental group.

Appendix C (Manuscript III)

Exam-related unwanted intrusive thoughts and related neutralizing behaviors:

Analogues to obsessions and compulsions

Martin Kollárik, Carlotta V. Heinzl, Marcel Miché, Roselind Lieb, Karina Wahl

submitted to *PLOS ONE*

Exam-related unwanted intrusive thoughts and related neutralizing behaviors: Analogues to obsessions and compulsions

Short Title: Exam-related unwanted intrusive thoughts and related neutralizing behaviors

Martin Kollárik, Carlotta V. Heinzl, Marcel Miché, Roselind Lieb, Karina Wahl*

Department of Psychology, Division of Clinical Psychology and Epidemiology,
University of Basel, Basel, Switzerland

* Corresponding author

E-mail: karina.wahl@unibas.ch (KW)

Abstract

Exam-related unwanted intrusive thoughts (UITs) and related neutralizing behaviors are common experiences among students. The present study investigated in what ways these UITs and behaviors are analogues to clinical obsessions and compulsions. Twenty-nine students completed three ecological momentary assessment surveys per day over 7 consecutive days, assessing the severity of exam-related UITs and related neutralizing behaviors, obsessive-compulsive (OC) symptoms, anxiety, distress, urge to neutralize, depressed mood, and stress in the week immediately before an exam period. Multilevel analysis demonstrated that the severity of exam-related UITs and related neutralizing behaviors was positively associated with OC symptoms, anxiety, distress, urge to neutralize, and stress but was not related to depressed mood. During the study period, the exam-related UITs occurred on average 7 times, and the related neutralizing behaviors on average 6 times. Overall, they were experienced with mild severity, low distress, and low urge to neutralize. Findings indicate that some aspects of exam-related UITs and related neutralizing behaviors (e.g., association with distress and urge to neutralize) might be analogous to OC symptoms but not all (e.g., no relation to depressed mood). We discuss how research on obsessive-compulsive disorder could benefit from considering exam-related UITs and related behaviors.

Introduction

Almost everyone occasionally experiences thoughts that enter the mind abruptly and are perceived as unwanted [1, 2]. Such unwanted intrusive thoughts (UITs) occur on a variety of themes such as contamination fears, doubts about possible mistakes, harm, immoral sexual thoughts, religion, and symmetry or order [2-4]. Previous studies have demonstrated that UITs are ubiquitous and occur irrespective of cultural

differences [2]. The cognitive model of obsessive-compulsive disorder (OCD) postulates that UITs lie on a continuum with obsessions [5, 6], since UITs share certain commonalities with clinically relevant obsessions [1]. Both UITs and obsessions are by definition repetitive intrusive thoughts, images, or impulses that are unwanted and disruptive [1, 7, 8]. Additionally, they tend to be followed by overt or covert neutralizing behaviors (termed compulsions in the context of OCD) with the aim of reducing the associated anxiety [7, 9, 10]. Previous studies suggest that the differences between UITs/neutralizing behaviors and obsessions/compulsions are a matter of degree rather than kind: The content is comparable [1, 3, 4, 11, however, see 12 for different findings], but UITs and neutralizing behaviors occur less frequently [1, 3, 11, 13] and are less distressing [1, 11, 13], and UITs interfere less with daily life than obsessions [3, 14].

It is common to study obsessive-compulsive (OC) phenomena in samples that do not necessarily include individuals with clinically relevant OC symptoms [15, 16]. This is mainly for practical reasons (easy to assess and affordable) but there are certain advantages as well [e.g., more precise experimental control; 15]. Two research practices are common. In the first, nonclinical participants are typically provided with a list of UITs and neutralizing behaviors and are asked to indicate which of these UITs and behaviors they have personally experienced [17-19]. One limitation of this method is that the participants tend to experience the selected UITs very rarely [e.g., a few times a year; 17, 20, 21], which might compromise the transfer of results to obsessions and compulsions. In the second practice, researchers temporarily induce UITs and neutralizing behaviors in the laboratory [16, 22]. For example, writing and then saying out loud a sentence such as "I hope [insert loved person] is in a car accident" can be an effective method of inducing UITs and neutralizing behaviors that resemble obsessions and compulsions [10, 16, 23, 24]. However, the experimentally induced

UITs and neutralizing behaviors are not necessarily idiosyncratic (person specific) and do not occur naturally in everyday life. This might also limit the transfer of results to obsessions and compulsions, which are highly idiosyncratic [25] and typically occur spontaneously in various situations [21, 26]. Finding a way to study more frequently occurring, idiosyncratic UITs and related neutralizing behaviors in their natural settings would be a further benefit to analogue research on OCD, as it would address the limitations of the aforementioned two methods.

Exam taking in students might be a promising starting point for investigating naturally occurring idiosyncratic UITs and related neutralizing behaviors, since exam periods are characterized by several key factors that play a role in OCD: high levels of perceived anxiety, depression, and stress [27, 28]. During exam periods, students are faced with the threat of failing an important exam and it is their responsibility to prevent this threat from coming true; a critical mistake can have important consequences. The cognitive-behavioral conceptualization of OCD asserts that more anxiety and stress will lead to more frequent intrusions [5, 29]. This model also suggests that higher levels of depression will increase the probability of misinterpreting these intrusions, which in turn is likely to trigger neutralizing responses [29]. According to Salkovskis and Millar (30), perceived responsibility for the consequences of the intrusions is a prerequisite for a neutralizing behavior. Thus, it is plausible to assume that UITs and neutralizing behaviors that are analogues to obsessions and compulsions might occur frequently during exam periods. Previous studies found indications that university students experienced exam-related UITs during exam periods [31, 32], for example, thoughts related to inadvertently making critical mistakes [32], and that they performed related neutralizing behaviors, such as repetitively checking or superstitious behaviors [32, 33]. However, previous research did not address in what way these thoughts and behaviors might be analogous to obsessions and compulsions.

The main purpose of the present study was to evaluate in what ways naturally occurring, idiosyncratic exam-related UITs and related neutralizing behaviors can be considered analogues to clinical obsessions and compulsions. We consider mainly two aspects to be important for the evaluation: (a) the characteristics of the respective thoughts and behaviors, which are made explicit in their definitions, and (b) their associations with other theoretically relevant constructs. The strength of the associations would indicate the degree of overlap among these constructs. Thus, exam-related UITs and related neutralizing behaviors should share some typical characteristics with obsessions (e.g., intrusiveness) and compulsions (e.g., repetitiveness), respectively. In addition, they should be associated with OC symptoms and the core aspects of OCD [e.g., anxiety, distress, and urge to engage in neutralizing; 7, 29] to a medium to high degree and also with other relevant factors thought to play a role in OCD, such as depressed mood and stress [5, 29], to a somewhat lesser degree.

In this study, the key characteristics of exam-related UITs and related neutralizing behaviors were defined in a way that maximizes the chances of high conceptual overlap with clinically relevant obsessions and compulsions. We defined the idiosyncratic exam-related UITs as “intrusive, short thoughts, images, or impulses that pop into your mind repeatedly, are related to the exam and its preparations ... [and are considered as] irrational, unwanted, or exaggerated” [32, p. 363]. Exam-related neutralizing behaviors were defined as any behaviors “dealing with such thoughts and/or the associated anxiety that was experienced as irrational or exaggerated” [32, p. 363]. Thus, by definition, exam-related UITs overlap with obsessive thoughts: Both are idiosyncratic, unwanted, intrusive, and irrational [7], enter the mind repeatedly, and can take the form of thoughts, images, or impulses. Similarly, exam-related neutralizing behaviors overlap with compulsions by definition: They tend to occur as a

consequence of their respective unwanted intrusions as a way to manage the associated anxiety. In this study, we mainly focused on the second aspect to evaluate the analogy, that is, the associations of exam-related UITs and related neutralizing behaviors with OC symptoms and the core aspects of OCD, as well as other OCD-relevant factors.

To gain more insight into the experience of exam-related UITs and related neutralizing behaviors, we first collected descriptive information (type, frequency, degree of severity, distress, and urge to neutralize). Second, to evaluate the analogy, we investigated the associations of the exam-related UITs and related neutralizing behaviors with OC symptoms and the core aspects of OCD (anxiety, distress, and urge to neutralize) as well as other OCD-relevant factors (depressed mood and perceived stress). We hypothesized that the analogy would be reflected in medium to large positive associations of exam-related UITs and related neutralizing behaviors with OC symptoms, anxiety, distress, and urge to neutralize. In addition, we hypothesized that the exam-related UITs and related neutralizing behaviors would be positively associated with depressed mood and perceived stress but with lower than medium effect sizes, since depressed mood and stress are not symptoms of OCD but rather factors that might foster OCD. Additionally, we provide exploratory information about the potential adaptiveness of exam-related UITs and related neutralizing behaviors in terms of confidence in passing the exam, and about the context in which the exam-related UITs and related neutralizing behaviors typically occur.

For brevity, in the Methods and Results sections, we use the abbreviations ER for exam related and N for neutralizing behaviors, for example, ER-UITs-N when referring to exam-related UITs and related neutralizing behaviors as one construct and ER-UITs and ER-N for exam-rated UITs and exam-rated neutralizing behaviors, respectively.

Methods

Participants and measures

We recruited 29 undergraduate psychology students ($M_{\text{age}} = 21.36$ years, $SD = 1.87$; 92.86% female) over 3 consecutive years at the end of the participants' 1st year at the University of Basel, 1 week before a stressful exam period. Participants were a subsample of the Wahl, Hofer (32) study who agreed to participate in the ecological momentary assessment (EMA) part. The study was approved by the Ethics Committee of the Department of Psychology, University of Basel. Participants received course credit and monetary compensation (20 Swiss francs) for their participation.

Baseline measures

Depression and anxiety (baseline). We used the Beck Depression Inventory [BDI; 34, German version: 35] and the Beck Anxiety Inventory [BAI; 36, German version: 37] to assess depressive and anxiety symptoms, respectively. Each scale consists of 21 items. The psychometric properties of the scales are well established [36-38].

OC symptoms (baseline). The Obsessive-Compulsive Inventory, Revised [OCI-R; 39, German version: 40], an 18-item self-report measure, was used to assess OC symptoms. The scale has demonstrated good validity and reliability [40, 41].

Perceived stress (baseline). The Symptoms subscale (13 items) of the Stress and Coping Inventory [SCI; 42] was used to assess perceived stress at baseline. SCI is a self-report questionnaire regarding perceived stress, stress symptoms, and coping with stress. The SCI Symptoms subscale measures physiological (e.g., abdominal pain, headache) and psychological (e.g., concentration and sleep problems) stress symptoms in the last 6 months. The subscale has good validity and reliability [32, 42].

Types of most frequent ER-UITs and ER-N (baseline). We identified the two most frequent ER-UITs and the two most frequent ER-N for each participant using the Stress-Related Thoughts and Behavior List [StressRTBL; 32]. Using this list, the interviewer read the definition of ER-UITs-N and then administered a checklist of 47 items assessing the occurrence of ER-UITs and ER-N over the past 2 weeks. The ER-UITs and ER-N were categorized into three and five groups of similar themes, respectively. The thought categories were fear of forgetting something important (e.g., accidentally missing a whole book chapter), fear of making a critical mistake (e.g., accidentally mixing up documents), and superstitious thoughts (e.g., must study for a specific amount of time per day or the exam will not go well). The behavior categories were checking (e.g., checking exam material to make sure nothing important was left out), ordering (e.g., arranging the desk), superstitious behavior (e.g., always using the same pen while studying), rigid rules or rituals (e.g., organizing every day in detail), and reassurance seeking (e.g., seeking reassurance that the exams will go well). Participants were asked to choose from the checklist the two ER-UITs and two ER-N that they had experienced most frequently in the last 2 weeks. These most frequent UITs and behaviors were then used for the EMA.

EMA measures (administered with a self-developed EMA app on a smartphone)

Severity of ER-UITs-N (EMA). We administered a modified Yale–Brown Obsessive Compulsive Scale [Y-BOCS; 43, German version: 44] to assess the severity of ER-UITs (modified Y-BOCS Obsessions subscale; items 1–5) and ER-N (modified Y-BOCS Compulsions subscale; items 6–10) in the last 30 min. Participants were asked to think about their first or second most frequent ER-UIT or ER-N, as identified in the StressRTBL interview, when answering. This procedure was preferred over asking about “any ER-UIT/ER-N” to provide a clear reference and thus increase

the validity of the method. The answers were recorded on a scale of 0 (*not at all severe*) to 4 (*extremely severe*). It is common to investigate obsessions and compulsions as one construct [45], since they cooccur [46]. Thus, we operationalized the EMA severity of ER-UITs-N as the total scale of the modified Y-BOCS (sum score of items 1–10).

Frequencies of ER-UITs and ER-N (EMA). We assessed frequency by referring to the two most frequent ER-UITs and ER-N in the last 30 min, as identified during the baseline StressRTBL interview (“How often did you experience thoughts such as [most frequent ER-UIT] or [second most frequent ER-UIT] in the last 30 minutes?” “How often did you engage in behaviors such as [most frequent ER-N] or [second most frequent ER-N] in the last 30 minutes?”). Both items were rated on a scale of 0 (not at all) to 5 (always).

Perceived anxiety, depressed mood, distress, and urge to neutralize (EMA). We assessed anxiety (“How afraid were you in the last 30 minutes?”) and depressed mood (“To what degree did you feel depressed or despondent in the last 30 minutes?”) with one item each on a scale of 0 (*not at all*) to 100 (*extremely*). Distress (“To what degree did you feel distressed in response to your [first most frequent ER-UIT] or [second most frequent ER-UIT] in the last 30 minutes?”) and urge to neutralize (“To what degree did you experience the urge to give in to your intrusive exam-related thoughts or to reassure yourself with an action such as [first most frequent ER-N] or [second most frequent ER-N] in the last 30 minutes?”) were both rated on a scale of 0 (*very low*) to 5 (*extremely high*).

OC symptoms (EMA). We used the modified OCI-R to assess OC symptoms in the last 30 min. In contrast to the baseline measure, we did not include the Hoarding subscale. Included subscales were washing, obsessing, ordering, checking, and neutralizing (mental counting rituals). The answer categories ranged from 0 (*not at all*) to 5 (*all the time*).

Perceived stress (EMA). We assessed overall level of stress (“How stressed were you in the last 30 min?”) with one item, rated on a scale of 0 (*not at all*) to 5 (*extremely*).

Perceived likelihood of exam success, studying for exams, and presence of other people (EMA). To assess the perceived likelihood of exam success, the following question was administered: “How likely is it that you will pass the upcoming exams?” The answers were provided on a scale of 0% to 100%. Studying for exams (“Are you currently reviewing materials for the upcoming exams?”) and the presence of other people (“Are you currently alone?”) were both answered with either 1 (*yes*) or 0 (*no*).

Procedure

The study took place during the week before a critical 2-week exam period at the end of the participants’ 1st year at the University of Basel. If students fail any of the six exams more than once, they are not allowed to study psychology anywhere in Switzerland. Therefore, this time period represents a very stressful life event for the students. Our testing started with a baseline measure and was followed by a 1-week EMA. For the baseline assessment, participants were tested individually in a laboratory at the University of Basel. All participants gave their written informed consent to participate in the study. Participants were then asked to rate their anxiety, depressed mood, OC symptoms, and perceived stress on a set of standardized questionnaires (BAI, BDI, OCI-R, and SCI). Following this, they were administered the StressRTBL by trained master’s students to identify their two most frequent ER-UITs and two most frequent ER-N (these were used as the target thoughts and behaviors for the EMA survey). After completing the baseline assessment, they were introduced to the EMA app, which was installed on a smartphone. Subsequently, they completed a short

practice trial to ensure they understood how to respond correctly to the EMA surveys. The investigator created an EMA schedule for every participant. During the next 7 days (the EMA period), the participants were prompted three times a day (Time Point 1: 7–10 a.m., Time Point 2: noon–5 p.m., and Time Point 3: 4–11 p.m.) to self-evaluate the frequency, distress, urge to neutralize, and severity of their two most frequent ER-UITs and two most frequent ER-N as well as anxiety, depressed mood, OC symptoms, perceived stress, and additional variables (perceived likelihood of exam success, studying for exams, and presence of other people). For two participants, the scheduled EMA surveys for the morning hours conflicted with their private schedule. Thus, they were prompted in the early afternoon (noon–1 p.m.), late afternoon (3–5 p.m.), and evening (6–10 p.m.).

Statistical analysis

All data were analyzed with IBM SPSS Statistics version 25; the figure was created using GraphPad Prism version 8.2.0 for Windows. First, we provide descriptive information about the type, frequency, and severity of ER-UITs-N as well as distress and urge to neutralize. To determine the frequency of ER-UITs during the EMA period, we transformed the frequency measure of these UITs into a dummy variable with ratings of 0 (*not at all*) coded as 0 (*ER-UIT did not occur*) and ratings of 1–5 coded as 1 (*ER-UIT occurred*). Subsequently, we built a sum score across all 21 time points for each participant. The frequency of ER-N during the EMA period was operationalized the same way as the frequency of ER-UITs.

Second, we conducted multilevel analyses to examine the associations of ER-UITs-N with the OC symptoms, OCD-relevant factors, and additional variables (perceived likelihood of exam success, studying for exams, and presence of other

people). Specifically, we calculated separate multilevel models for the EMA severity of ER-UITs-N (modified Y-BOCS total scale) as the outcome variable with each of the nine predictors. Predictors were EMA OC symptoms, anxiety, distress, urge to neutralize, depressed mood, perceived stress, perceived likelihood of exam success, studying for exams, and presence of other people.

Consistent with current recommendations [47, 48], we kept the random effects at the justifiable maximum. Thus, for each separate model, we defined the intercept and slope as random effects. For each multilevel analysis, we constructed a two-level model with the measurement occasions ($n = 21$ per participant; Level 1) nested within individuals (Level 2). Each model contained the respective predictor plus the variable time (in days) to account for linear time trends. As both of these variables were time varying, we added a random slope parameter for both of them. However, if the model failed to converge, we removed the random slope of the main predictor. For brevity, we do not report the time variable in the Results section.

To estimate all parameters, we used the maximum likelihood estimation method. Unstandardized estimates are reported in the Results section. Consistent with recent recommendations [49, 50], we calculated Cohen's f^2 [51] as the indicator of effect size for the fixed effects. Cohen's f^2 can be interpreted as the proportion of variance accounted for by the given predictor relative to a null model. We report the f^2 value unique to each main predictor, that is, over and above the effect of the variable time on the outcome variable. To ensure that the reduction of variance was accounted for only by the fixed effects and not by the random effects, we held the random effects constant (i.e., we defined only a random intercept) for each model [50] when calculating f^2 . The magnitude of f^2 can be interpreted using Cohen's guidelines [51]: 0.02 for a small effect, 0.15 for a medium effect, and 0.35 for a large effect; the in-between f^2 values were interpreted as a range (e.g., a value of 0.08 would be

interpreted as a small to medium effect; values above 0.35 were interpreted as large effects). The continuous EMA predictors were group-mean centered to facilitate interpretation, such that the obtained effects represented changes in the outcome due to changes on the individualized predictor scale, that is, both slope and intercept. Visual inspection of the data did not indicate a violation of the models' assumptions (functional form, normality of residuals, and homoscedasticity). The level of statistical significance was set at $p < .05$.

Following the generalizability approach [52] and consistent with current recommendations [53], we computed between- (R_{KF}) and within-person (R_C) reliability coefficients for the EMA scales (modified Y-BOCS and modified OCI-R). While R_{KF} represents the reliability of a measure across all days, R_C is the reliability of change in ratings over time across individuals. The reliability coefficients from .61 to .80 were interpreted as moderate and from .81 to 1.00 as substantial [54].

Results

Sample characteristics, compliance with the EMA method, and reliability coefficients

Two participants were excluded from the main analyses because they reported intrusive thoughts that were more typical of worrisome thoughts [55; e.g., "What will happen if I don't pass the exam?" or "What is my plan B?"] instead of ER-UITs during the baseline interview (StressRTBL). Table 1 shows means, standard deviations, and reliability coefficients for baseline measures of OC symptoms, anxiety, depression, and perceived stress. The OC symptoms appeared comparable to those reported in other studies with university students [56-58, however, see 59, 60 for different findings]. Anxiety and depressive symptoms in this sample seemed comparable to those in

another student sample under exam stress [61]. The internal consistency of the baseline measures ranged from good (SCI symptoms) to excellent (BAI). The participants completed a total of 92.06% of the possible EMA prompts, which suggests a high compliance with the study design [for comparison, see 62, 63]. The reliability coefficients for both EMA scales were substantial (modified Y-BOCS: $R_{KF} = 1.00$, $R_C = .90$; modified OCI-R: $R_{KF} = .99$, $R_C = .82$).

Table 1. Means, standard deviations, and Cronbach's alpha of the baseline measures.

Baseline variable	<i>M</i> (<i>SD</i>)	α
OC symptoms (OCI-R)	11.04 (9.54)	.89
Anxiety (BAI)	14.15 (9.29)	.90
Depression (BDI)	8.04 (6.88)	.89
Perceived stress (SCI symptoms)	24.81 (6.82)	.84

Note. BAI = Beck Anxiety Inventory; BDI = Beck Depression Inventory; OC = obsessive-compulsive; OCI-R = Obsessive-Compulsive Inventory, Revised; SCI = Stress and Coping Inventory.

Descriptive information about the ER-UITs and ER-N (type, frequency, degree of severity, distress, and urge to neutralize)

Fig 1 shows the most frequent types of ER-UITs and the most frequent types of ER-N as indicated by the participants during the baseline StressRTBL interview. Each participant could report a maximum of two ER-UITs and two ER-N. In total, 53 ER-UITs and 54 ER-N were reported. The most prominent ER-UITs were the fear of forgetting something important, followed by superstitious thoughts and the fear of making a critical mistake. The most prominent ER-N were rigid rules or rituals, followed

by checking, superstitious behavior, reassurance seeking, and ordering. During the EMA period, on average (across all subjects and time points), participants experienced seven ER-UITs ($M = 6.67$, $SD = 5.06$) and six ER-N ($M = 5.96$, $SD = 4.85$).

Fig 1. Frequencies (as percentages) of the most frequent exam-related unwanted intrusive thoughts (ER-UITs) and exam-related neutralizing behaviors (ER-N) as reported by the participants on the Stress-Related Thoughts and Behavior List at baseline. The frequency represents the number of each reported UIT and behavior within the total number reported (for ER-UITs: $n = 53$; for ER-N: $n = 54$) by the participants. Each participant could report a maximum of two UITs as well as two behaviors.

Table 2 shows means and standard deviations of the severity of ER-UITs-N, distress, and urge to neutralize, and the number of occurrences of low, mild, moderate, and high scores of the severity of ER-UITs-N during the EMA period. On average, participants experienced the ER-UITs-N with mild severity, low distress, and low urge to neutralize.

Table 2. Means and standard deviations of severity of ER-UITs-N, distress, and urge to neutralize, and number of occurrences of low, mild, moderate, and high scores of the severity of ER-UITs-N during the EMA period.

EMA variable	$M (SD)^a$	$N (\%)^b$
Severity of ER-UITs-N ^c	8.79 (6.37)	
Distress	1.49 (1.08)	
Urge to neutralize	1.31 (1.04)	
Severity score of ER-UITs-N		
Low (scores of 0–7)		52 (44.07)

EMA variable	<i>M (SD)</i> ^a	<i>N (%)</i> ^b
Mild (scores of 8–15)		50 (42.37)
Moderate (scores of 16–23)		13 (11.02)
High (scores of 24–40)		3 (2.54)

Note. EMA = Ecological momentary assessment; ER-UITs-N = exam-related unwanted intrusive thoughts and related neutralizing behaviors. Y-BOCS = Yale–Brown Obsessive Compulsive Scale.

^a To calculate the means, we included only those EMA prompts in which the participants reported experiencing the ER-UIT together with the ER-N in the last 30 min.

^b To calculate the number of occurrences of low, mild, moderate, and high scores of the severity of ER-UITs-N, we included only those EMA prompts in which the participants reported experiencing the ER-UIT together with the ER-N in the last 30 min. The benchmarks for the severity were consistent with the common Y-BOCS cut-offs [45].

^c Scoring is consistent with Y-BOCS, e.g., 8–15 = mild symptom severity [45].

The associations of ER-UITs-N with OC symptoms, anxiety, distress, urge to neutralize, depressed mood, perceived stress, perceived likelihood of exam success, studying for exams, and presence of other people (EMA measures)

Table 3 displays the results of the multilevel analyses. The severity of ER-UITs-N was positively associated with OC symptoms, anxiety, distress, and urge to neutralize, respectively. The associations with OC symptoms and anxiety were small to medium in size, while the associations with distress and urge to neutralize were large. The severity of ER-UITs-N was positively associated with perceived stress but was not related to depressed mood. The degree of association with perceived stress was small to medium.

Table 3. Associations between the EMA severity of ER-UITs-N (outcome) and other EMA variables (predictors).

EMA predictor variable (fixed effects)	Severity ER-UITs-N (modified Y-BOCS total)		
	Coefficient (SE)	<i>p</i>	<i>f</i> ²
OC symptoms	0.46 (0.13)	.004	.06
Anxiety	0.07 (0.02)	.002	.14
Distress	2.73 (0.34)	< .001	.58
Urge to neutralize	2.68 (0.34)	< .001	.50
Depressed mood	0.04 (0.02)	.052	.03
Perceived stress	1.17 (0.37)	.004	.12

Note. Each fixed effect represents a separate multilevel analysis with EMA severity of ER-UITs-N as the outcome. EMA = Ecological momentary assessment; ER-UITs-N = exam-related unwanted intrusive thoughts and related neutralizing behaviors; OC = obsessive-compulsive; Y-BOCS = Yale–Brown Obsessive Compulsive Scale.

The severity of ER-UITs-N was negatively associated with the perceived likelihood of exam success (coefficient = -0.07, *SE* = 0.02, *p* < .001, *f*² = .03) and positively associated with studying for exams (coefficient = 1.15, *SE* = 0.38, *p* = .005, *f*² = .02). There was no association between the severity of ER-UITs-N and presence of other people (coefficient = -0.04, *SE* = 0.47, *p* = .940, *f*² < .001).

Discussion

To present a more comprehensive picture of the experience of exam-related UITs and related neutralizing behaviors, we first provide descriptive information about this phenomenon. The most common type of exam-related UITs reported by the students was the fear of forgetting something important, followed by superstitious thoughts and the fear of making a critical mistake. The most common exam-related neutralizing behaviors was rigid rules or rituals closely followed by checking, then superstitious behaviors and reassurance seeking, and finally ordering. During the study period, students reported experiencing the exam-related UITs on average seven times and engaging in related neutralizing behaviors on average six times. Compared to OC symptoms, which occurred several times a day in other studies [1, 13], exam-

related UITs and related neutralizing behaviors seem to be less frequent. However, when we compare them roughly to the occurrence of subclinical OC symptoms, which were experienced on average once or twice a month in Morillo, Giménez (64), exam-related UITs and related neutralizing behaviors seem to be more frequent than subclinical OC symptoms.

The exam-related UITs and related neutralizing behaviors were experienced on average with mild severity, low distress, and low urge to neutralize. Since the scoring of the severity of these UITs and related behaviors is consistent with the Y-BOCS [45], exam-related UITs and related neutralizing behaviors might be comparable to mild OC symptom severity. In other studies, OC symptoms were on average associated with moderate to high distress and moderate to high urge to neutralize [1, 65]. In light of this, exam-related UITs and related neutralizing behaviors appear to be less distressing, and the urge to neutralize is lower than for OC symptoms. We emphasize that the previous studies used different methods to assess the frequency, severity, distress, and urge to neutralize related to OC symptoms, and thus our comparisons should be considered only rough estimations. Future studies are needed to directly compare exam-related UITs and related neutralizing behaviors with OC symptoms on frequency, severity, distress, and urge to neutralize.

The main goal of the present study was to investigate in what ways idiosyncratic exam-related UITs and related neutralizing behaviors can be considered analogues to clinical obsessions and compulsions. To evaluate the analogy, we primarily focused on the associations of the exam-related UITs and related neutralizing behaviors with OC symptoms and the core aspects of OCD, as well as with other OCD-relevant factors. The positive associations of the severity of exam-related UITs and related neutralizing behaviors with OC symptoms, anxiety, distress, urge to neutralize, and perceived stress generally support a conceptual overlap of these UITs and behaviors

with obsessions and compulsions. As predicted, the associations with distress and urge to neutralize were medium to large in their effect sizes. However, the associations with OC symptoms were smaller than expected. A possible explanation might be that common types of exam-related UITs and related neutralizing behaviors reported in this study did not exactly match the symptom dimensions included in the OCI-R (the scale we used to assess OC symptoms). For example, the OCI-R does not assess superstitious thoughts that were prominent exam-related UITs and are also common in OCD [4, 66]. Future studies might want to investigate the associations of exam-related UITs and related neutralizing behaviors also with other measures of OC symptoms, such as the Obsessional Intrusive Thoughts Inventory [4], which also includes a superstition dimension. The relatively small association of the severity of exam-related UITs and related neutralizing behaviors with anxiety was also not consistent with our hypothesis. Previous studies have demonstrated medium to large associations between OC symptoms and anxiety [40, 43]. The severity of exam-related UITs and related neutralizing behaviors was not associated with depressed mood. Other studies have shown OC symptoms to be related to depressed mood [67-69]. Overall, our hypotheses were partially supported.

Additionally, we investigated factors that might be associated with the exam-related UITs and related neutralizing behaviors. An increase in the severity of these UITs and behaviors was associated with a decrease in the perceived likelihood of passing the upcoming exams. This indicates that exam-related UITs and related neutralizing behaviors might not be adaptive for students in terms of passing the exams. Future studies might want to investigate the effect of these UITs and behaviors on *actual* exam outcomes. Students who were currently studying for the upcoming exams experienced more severe exam-related UITs and related neutralizing behaviors than those who were not currently studying. This indicates that the occurrence of

exam-related UITs and related neutralizing behaviors might not be evenly distributed during exam periods but might be more frequent when actually studying. However, this finding must be interpreted with caution, since we did not measure the duration of exam preparations. Finally, the severity of exam-related UITs and related neutralizing behaviors did not change depending on the presence of other people.

There are several limitations of the current study to acknowledge. First, although we investigated important associations of exam-related UITs and related neutralizing behaviors with OC symptoms and the core aspects of OCD, we did not directly address the degree of overlap in some of the defining characteristics, such as the degree of intrusiveness, irrationality, or ego-dystonicity [21]. Second, the findings might not generalize beyond our sample to the population. Next, we did not control for reactivity, that is, participants' biased responses solely due to the EMA methodology. However, recent studies indicated that the effect of smartphone assessment on the frequency of intrusions is rather low [62, 70]. Finally, the psychometric properties of the single-item measures we used to assess a number of relevant constructs (e.g., anxiety, depression, etc.) during the study period should be explored.

To conclude, exam-related UITs and related neutralizing behaviors might be considered analogous to obsessions and compulsions in some aspects: they were positively associated with distress and urge to neutralize to a large degree, and with perceived stress to a small to medium degree. However, in some other aspects, they might not be analogous: Their association with anxiety was smaller than expected, while their relation to depressed mood was not significant. In addition, exam-related UITs and related neutralizing behaviors were associated with OC symptoms to a smaller degree than predicted. However, as discussed above, this might be explained by the methodology of the present study. Nevertheless, we suggest that exam-related UITs and related neutralizing behaviors could extend analogue research by providing

an opportunity to study relatively frequent, idiosyncratic UITs and related neutralizing behaviors in their natural environment. Relative to questionnaire studies, which typically use UITs with a very low frequency, exam-related UITs have the advantage of being more frequent, and their variation over time could be used to study covariations with potential factors that might foster OCD and their mutual or converging courses in longitudinal studies. Exam-related UITs and neutralizing behaviors might be less useful in studies that investigate questions related to anxiety or depression in OCD, since their associations might not reflect the relation of OC symptoms with these constructs. Additionally, the application of exam-related UITs might be less relevant for laboratory-based experimental studies, since our preliminary data suggest that they occur more often when studying for exams.

Acknowledgments

We would like to thank all participants and other contributors of the present study.

References

1. Rachman S, de Silva P. Abnormal and normal obsessions. *Behav Res Ther.* 1978;16(4):233-48. doi: 10.1016/0005-7967(78)90022-0.
2. Radomsky AS, Alcolado GM, Abramowitz JS, Alonso P, Belloch A, Bouvard M, et al. Part 1—You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders.* 2014;3(3):269-79. doi: 10.1016/j.jocrd.2013.09.002.
3. Bouvard M, Fournet N, Denis A, Sixdenier A, Clark D. Intrusive thoughts in patients with obsessive compulsive disorder and non-clinical participants: a comparison using the International Intrusive Thought Interview Schedule. *Cognitive Behaviour Therapy.* 2017;46(4):287-99. doi: 10.1080/16506073.2016.1262894.
4. Garcia-Soriano G, Belloch A, Morillo C, Clark DA. Symptom dimensions in obsessive-compulsive disorder: from normal cognitive intrusions to clinical obsessions. *Journal of Anxiety Disorders.* 2011;25(4):474-82. doi: 10.1016/j.janxdis.2010.11.012.

5. Rachman S. A cognitive theory of obsessions. *Behav Res Ther.* 1997;35(9):793-802. doi: 10.1016/S0005-7967(97)00040-5.
6. Rachman S. A cognitive theory of obsessions: elaborations. *Behav Res Ther.* 1998;36(4):385-401. doi: 10.1016/S0005-7967(97)10041-9.
7. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington, DC: Author; 2013.
8. Rachman S. Part I. Unwanted intrusive cognitions. *Advances in Behaviour Research and Therapy.* 1981;3(3):89-99. doi: 10.1016/0146-6402(81)90007-2.
9. Freeston MH, Ladouceur R. What do patients do with their obsessive thoughts? *Behav Res Ther.* 1997;35(4):335-48. doi: 10.1016/S0005-7967(96)00094-0.
10. Rachman S, Shafran R, Mitchell D, Trant J, Teachman B. How to remain neutral: An experimental analysis of neutralization. *Behav Res Ther.* 1996;34(11-12):889-98. doi: 10.1016/S0005-7967(96)00051-4.
11. Muris P, Merckelbach H, Clavan M. Abnormal and normal compulsions. *Behav Res Ther.* 1997;35(3):249-52. doi: 10.1016/S0005-7967(96)00114-3.
12. Berry LM, Laskey B. A review of obsessive intrusive thoughts in the general population. *Journal of Obsessive-Compulsive and Related Disorders.* 2012;1(2):125-32. doi: 10.1016/j.jocrd.2012.02.002.
13. Morillo C, Belloch A, Garcia-Soriano G. Clinical obsessions in obsessive-compulsive patients and obsession-relevant intrusive thoughts in non-clinical, depressed and anxious subjects: Where are the differences? *Behav Res Ther.* 2007;45(6):1319-33. doi: 10.1016/j.brat.2006.11.005.
14. Garcia-Soriano G, Belloch A. Symptom dimensions in obsessive-compulsive disorder: Differences in distress, interference, appraisals and neutralizing strategies. *J Behav Ther Exp Psy.* 2013;44(4):441-8. doi: 10.1016/j.jbtep.2013.05.005.
15. Abramowitz JS, Fabricant LE, Taylor S, Deacon BJ, McKay D, Storch EA. The relevance of analogue studies for understanding obsessions and compulsions. *Clin Psychol Rev.* 2014;34(3):206-17. doi: 10.1016/j.cpr.2014.01.004.
16. De Putter LM, Van Yper L, Koster EH. Obsessions and compulsions in the lab: A meta-analysis of procedures to induce symptoms of obsessive-compulsive disorder. *Clin Psychol Rev.* 2017;52:137-47. doi: 10.1016/j.cpr.2017.01.001.
17. Purdon C, Clark DA. Perceived control and appraisal of obsessional intrusive thoughts: A replication and extension. *Behavioural and Cognitive Psychotherapy.* 1994;22(4):269-85. doi: 10.1017/S1352465800013163.
18. Purdon C, Clark DA. Obsessive intrusive thoughts in nonclinical subjects. Part I. Content and relation with depressive, anxious and obsessional symptoms. *Behav Res Ther.* 1993;31(8):713-20. doi: 10.1016/0005-7967(93)90001-B.
19. Freeston MH, Ladouceur R, Thibodeau N, Gagnon F. Cognitive intrusions in a non-clinical population. I. Response style, subjective experience, and appraisal. *Behav Res Ther.* 1991;29(6):585-97.
20. Belloch A, Morillo C, Lucero M, Cabedo E, Carrio C. Intrusive thoughts in non-clinical subjects: The role of frequency and unpleasantness on appraisal ratings and control strategies. *Clin Psychol Psychot.* 2004;11(2):100-10. doi: 10.1002/cpp.397.
21. Julien D, O'Connor KP, Aardema F. Intrusive thoughts, obsessions, and appraisals in obsessive-compulsive disorder: A critical review. *Clin Psychol Rev.* 2007;27(3):366-83. doi: 10.1016/j.cpr.2006.12.004.

22. Gagné JP, Kelly-Turner K, Radomsky AS. From the laboratory to the clinic (and back again): How experiments have informed cognitive-behavior therapy for obsessive-compulsive disorder. *J Exp Psychopathol*. 2018;9(4). doi: 10.1177/2043808718810030.
23. van den Hout M, van Pol M, Peters M. On becoming neutral: effects of experimental neutralizing reconsidered. *Behav Res Ther*. 2001;39(12):1439-48. doi: 10.1016/S0005-7967(00)00109-1.
24. van den Hout M, Kindt M, Weiland T, Peters M. Instructed neutralization, spontaneous neutralization and prevented neutralization after an obsession-like thought. *J Behav Ther Exp Psy*. 2002;33(3-4):177-89. doi: 10.1016/S0005-7916(02)00048-4.
25. McKay D, Abramowitz JS, Calamari JE, Kyrios M, Radomsky A, Sookman D, et al. A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. *Clin Psychol Rev*. 2004;24(3):283-313. doi: 10.1016/j.cpr.2004.04.003.
26. Clark DA, Inozu M. Unwanted intrusive thoughts: Cultural, contextual, covariational, and characterological determinants of diversity. *Journal of Obsessive-Compulsive and Related Disorders*. 2014;3(2):195-204. doi: 10.1016/j.jocrd.2014.02.002.
27. Austin EJ, Saklofske DH, Mastoras SM. Emotional intelligence, coping and exam-related stress in Canadian undergraduate students. *Aust J Psychol*. 2010;62(1):42-50. doi: 10.1080/00049530903312899.
28. Zunhammer M, Eberle H, Eichhammer P, Busch V. Somatic symptoms evoked by exam stress in university students: the role of alexithymia, neuroticism, anxiety and depression. *Plos One*. 2013;8(12). doi: 10.1371/journal.pone.0084911.
29. Salkovskis PM. Obsessional-compulsive problems - a cognitive-behavioral analysis. *Behav Res Ther*. 1985;23(5):571-83. doi: 10.1016/0005-7967(85)90105-6.
30. Salkovskis PM, Millar JFA. Still cognitive after all these years? Perspectives for a cognitive behavioural theory of obsessions and where we are 30 years later. *Aust Psychol*. 2016;51(1):3-13. doi: 10.1111/ap.12186.
31. Lepore SJ. Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *J Pers Soc Psychol*. 1997;73(5):1030-7. doi: 10.1037/0022-3514.73.5.1030.
32. Wahl K, Hofer PD, Meyer AH, Lieb R. Prior beliefs about the importance and control of thoughts are predictive but not specific to subsequent intrusive unwanted thoughts and neutralizing behaviors. *Cognitive Ther Res*. 2019;44:360-75. doi: 10.1007/s10608-019-10046-7.
33. Keinan G. The effects of stress and desire for control on superstitious behavior. *Pers Soc Psychol B*. 2002;28(1):102-8. doi: 10.1177/0146167202281009.
34. Beck AT, Rush AJ, Shaw BF, Emery G. *Cognitive therapy of depression*. New York: Guilford Press; 1979.
35. Hautzinger M, Bailer M, Worall H, Keller F. *Beck-Depressions-Inventar (BDI)*. Testhandbuch. 2nd rev. ed. Bern, Switzerland: Hans Huber; 1995.
36. Beck AT, Brown G, Epstein N, Steer RA. An inventory for measuring clinical anxiety - psychometric properties. *J Consult Clin Psych*. 1988;56(6):893-7. doi: 10.1037/0022-006x.56.6.893.
37. Magraf J, Ehlers A. *Beck Angst-Inventar (BAI) - Manual*. Frankfurt am Main: Harcourt Test Services GmbH; 2007.
38. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory - 25 Years of evaluation. *Clin Psychol Rev*. 1988;8(1):77-100. doi: 10.1016/0272-7358(88)90050-5.

39. Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, et al. The obsessive-compulsive inventory: Development and validation of a short version. *Psychol Assessment*. 2002;14(4):485-96. doi: 10.1037//1040-3590.14.4.485.
40. Gönner S, Leonhart R, Ecker W. [The German version of the Obsessive-Compulsive Inventory-Revised: A brief self-report measure for the multidimensional assessment of obsessive-compulsive symptoms]. *Psychotherapie Psychosomatik Medizinische Psychologie*. 2007;57(9-10):395-404. doi: 10.1055/s-2007-970894.
41. Gönner S, Leonhart R, Ecker W. The Obsessive-Compulsive Inventory-Revised (OCI-R): Validation of the German version in a sample of patients with OCD, anxiety disorders, and depressive disorders. *Journal of Anxiety Disorders*. 2008;22(4):734-49. doi: 10.1016/j.janxdis.2007.07.007.
42. Satow L. Stress- und Coping-Inventar (SCI): Vollständige Test- und Skalendokumentation 2012. Available from: <http://www.drSATOW.de/tests/stress-und-coping-inventar/>.
43. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry*. 1989;46(11):1006-11. doi: 10.1001/archpsyc.1989.01810110048007.
44. Hand I, Büttner-Westphal H. [The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS): A semistructured interview for assessing severity of compulsive cognitions and behavior]. *Verhaltenstherapie*. 1991;1:223-5. doi: 10.1159/000257972.
45. Abramowitz JS, Jacoby RJ. Obsessive-compulsive and related disorders. In: Barlow DH, editor. *The Oxford handbook of clinical psychology*. Oxford University Press; 2014. p. 313-38.
46. Leonard RC, Riemann BC. The co-occurrence of obsessions and compulsions in OCD. *Journal of Obsessive-Compulsive and Related Disorders*. 2012;1(3):211-5. doi: 10.1016/j.jocrd.2012.06.002.
47. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang*. 2013;68(3):255-78. doi: 10.1016/j.jml.2012.11.001.
48. Schmidt-Catran AW, Fairbrother M. The random effects in multilevel models: Getting them wrong and getting them right. *Eur Sociol Rev*. 2016;32(1):23-38. doi: 10.1093/esr/jcv090.
49. Lorah J. Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-Scale Assess E*. 2018;6. doi: 10.1186/s40536-018-0061-2.
50. Selya AS, Rose JS, Dierker LC, Hedeker D, Mermelstein RJ. A practical guide to calculating Cohen's f^2 , a measure of local effect size, from PROC MIXED. *Front Psychol*. 2012;3. doi: 10.3389/fpsyg.2012.00111.
51. Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd ed: Routledge; 1988.
52. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Pers Soc Psychol B*. 2006;32(7):917-29. doi: 10.1177/0146167206287721.
53. Trull TJ, Ebner-Priemer UW. Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. *Journal of Abnormal Psychology*. 2020;129(1):56-63. doi: 10.1037/abn0000473.
54. Shrout PE. Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research*. 1998;7(3):301-17. doi: 10.1177/096228029800700306.

55. Borkovec TD, Robinson E, Pruzinsky T, Depree JA. Preliminary exploration of worry - some characteristics and processes. *Behav Res Ther.* 1983;21(1):9-16. doi: 10.1016/0005-7967(83)90121-3.
56. Zermatten A, van der Linden M, Jermann F, Ceschi G. Validation of a French version of the Obsessive–Compulsive Inventory-Revised in a non-clinical sample. *European Review of Applied Psychology.* 2006;56(3):151-5. doi: 10.1016/j.erap.2005.07.003.
57. Grisham JR, Williams AD. Cognitive control of obsessional thoughts. *Behav Res Ther.* 2009;47(5):395-402. doi: 10.1016/j.brat.2009.01.014.
58. Smári J, Olason DT, Eyporsdottir A, Frolunde MB. Psychometric properties of the Obsessive Compulsive Inventory-Revised among Icelandic college students. *Scandinavian Journal of Psychology.* 2007;48(2):127-33. doi: 10.1111/j.1467-9450.2007.00574.x.
59. Hajcak G, Huppert JD, Simons RF, Foa EB. Psychometric properties of the OCI-R in a college sample. *Behav Res Ther.* 2004;42(1):115-23. doi: 10.1016/j.brat.2003.08.002.
60. Woo CW, Kwon SM, Lim YJ, Shin MS. The Obsessive-Compulsive Inventory-Revised (OCI-R): psychometric properties of the Korean version and the order, gender, and cultural effects. *J Behav Ther Exp Psy.* 2010;41(3):220-7. doi: 10.1016/j.jbtep.2010.01.006.
61. Creamer M, Foran J, Bell R. The Beck Anxiety Inventory in a nonclinical sample. *Behav Res Ther.* 1995;33(4):477-85. doi: 10.1016/0005-7967(94)00082-U.
62. Rattel JA, Grunberger LM, Reichenberger J, Liedlgruber M, Miedl SF, Blechert J, et al. Frequency of intrusions and appraisal of related distress after analogue trauma: A comparative ecological momentary assessment methods study. *Cognitive Ther Res.* 2019;43(1):174-84. doi: 10.1007/s10608-018-9941-6.
63. Rosenkranz T, Takano K, Watkins ER, Ehring T. Assessing repetitive negative thinking in daily life: Development of an ecological momentary assessment paradigm. *Plos One.* 2020;15(4). doi: 10.1371/journal.pone.0231783.
64. Morillo C, Giménez A, Belloch A, Lucero M, Carrió C, Cabedo E. Pensamientos intrusos en obsesivos subclínicos: contenidos, valoraciones, y estrategias de control. *Revista de psicopatología clínica.* 2003;8(1):23-38.
65. Wahl K, van den Hout M, Heinzl CV, Kollárik M, Meyer A, Benoy C, et al. Rumination about obsessive symptoms and mood maintains obsessive-compulsive symptoms and depressed mood: An experimental study. 2020; Forthcoming.
66. Katerberg H, Delucchi KL, Stewart SE, Lochner C, Denys DAJP, Stack DE, et al. Symptom dimensions in OCD: Item-level factor analysis and heritability estimates. *Behav Genet.* 2010;40(4):505-17. doi: 10.1007/s10519-010-9339-z.
67. Bhar SS, Kyrios M. Obsessions and compulsions are associated with different cognitive and mood factors. *Behav Change.* 2005;22(2):81-96. doi: 10.1375/bech.2005.22.2.81.
68. Spinella M. Mood in relation to subclinical obsessive-compulsive symptoms. *Int J Neurosci.* 2005;115(4):433-43. doi: 10.1080/00207450590522838.
69. Ricciardi JN, McNally RJ. Depressed mood is related to obsessions, but not to compulsions, in obsessive-compulsive disorder. *Journal of Anxiety Disorders.* 1995;9(3):249-56. doi: 10.1016/0887-6185(95)00006-A.
70. Kleindienst N, Priebe K, Petri M, Hecht A, Santangelo P, Bohus M, et al. Trauma-related memories in PTSD after interpersonal violence: an ambulatory assessment study. *European Journal of Psychotraumatology.* 2017;8(1):1409062. doi: 10.1080/20008198.2017.1409062.