

Obsessive-Compulsive Visual Search: A Reexamination of Presence–Absence Asymmetries

**Noam Sarna<sup>1</sup>**, **Matan Mazor<sup>2</sup>**, **and Reuven Dar<sup>1</sup>** b<sup>1</sup> The School of Psychological Sciences, Tel Aviv University, and <sup>2</sup>All Souls College and

Department of Experimental Psychology, University of Oxford

Clinical Psychological Science 1–9 © The Author(s) 2024

Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/21677026241258380 www.psychologicalscience.org/CPS



#### Abstract

In previous research, obsessive-compulsive tendencies were associated with longer search times in visual-search tasks. These findings, replicated and extended to a clinical sample, were specific to target-absent trials, with no effect on target-present trials. This selectivity was interpreted as checking behavior in response to mild uncertainty. However, an alternative interpretation is that individuals with high obsessive-compulsive (OC+) tendencies have a specific difficulty with inference about absence. In two large-scale, preregistered, online experiments (conceptual replication: N = 1,007; direct replication: N = 226), we sought to replicate the original finding and elucidate its underlying cause: an increased sensitivity to mild uncertainty or a selective deficiency in inference about absence. Both experiments showed no evidence of prolonged search times in target-absent trials for OC+ individuals. Taken together, our results do not support the notion that inducing mild uncertainty in the form of target absence leads to excessive checking among OC+ individuals.

#### **Keywords**

inference about absence, metacognition, obsessive compulsive disorder, visual search

Received 11/16/23; Revision accepted 4/30/24

Theories on obsessive compulsive disorder (OCD) emphasize the pivotal role of pathological doubt in the disorder's phenomenology (Dar, 2004; Dar et al., 2021; Rasmussen & Eisen, 1989; Reed, 1985). This persistent doubt is reflected in lowered confidence in memory, decisionmaking, perception, and other cognitive functions, which give rises to repetitive checking rituals that, paradoxically, only serve to intensify the doubt (van den Hout & Kindt, 2003). In the lab, doubt and checking behavior are commonly manifested in slow reaction times (e.g., Banca et al., 2015; Hauser et al., 2017; Sarig et al., 2012).

In the present study, we focused on the finding that participants with high obsessive compulsive tendencies (OC+) took more time than participants with low OC tendencies (OC-) to identify when a target was absent from a visual-search array, whereas no such difference was observed when the target was present (Toffolo et al., 2013). These findings have been replicated (Toffolo et al., 2014) and extended to a clinical sample, in which they were found to be specific to patients with OCD and absent in patients suffering from anxiety (Toffolo et al., 2016). In these experiments, checking behavior was operationalized as search time, and high and low uncertainty were operationalized by means of contrasting target-present and target-absent trials. Relatively longer search times for the OC+ group in target-absent trials were interpreted as perseverative checking behavior under mild uncertainty.

However, although deciding that a target is absent is indeed commonly accompanied by lower levels of subjective confidence compared with deciding that a target is present (Mazor et al., 2020, 2021), these type of decisions about absence are also qualitatively different from decisions about presence because they cannot be based on direct perceptual evidence. To determine that a target is absent, one must believe that if the target were present, one would have been able to perceive it: a form of inference that requires counterfactual

#### **Corresponding Author:**

Noam Sarna, School of Psychological Sciences, Tel Aviv University Email: noamsarna@mail.tau.ac.il thinking and reliance on self-knowledge (Mazor, 2021). Therefore, an alternative mechanism behind the longer search times in target-absent trials among OC+ participants could be a specific difficulty with inference about absence rather than simply heightened sensitivity to uncertainty.

Clinical observations provide some support for the idea that people with OCD struggle with inferences about absence. One example is "hit-and-run OCD," in which individuals feel compelled to mentally or physically retrace their driving route to ensure that they did not kill or injure someone while driving (Hyman & Pedrick, 2010). This phenomenon manifests key properties of inference about absence: To conclude that an accident has not happened, a person needs to rely on the belief that if it did happen, the person would have noticed it.

This clinical example raises the possibility that the increased search time for target-absent trials may be due to a specific difficulty in inferring absence rather than a general intolerance of uncertainty. To test this idea, in two preregistered online studies, we conducted a conceptual replication and a direct replication of the visual-search study by Toffolo et al. (2013). Participants high and low in OCD tendencies were presented with visual-search displays and asked to decide whether a target was absent or present. Experiment 1 aimed to elucidate whether the increased search times in targetabsent trials for OC+ individuals are attributable to a specific difficulty with inference about absence or a general difficulty with handling uncertainty. Following our failure to replicate the original findings in this first experiment, Experiment 2 was designed as a more direct replication of Toffolo et al. (2013) using the exact same stimuli and instructions.

#### **Experiment** 1

In Experiment 1, we sought to dissociate specific difficulties with inference about absence from more general difficulties with uncertainty by introducing an easy targetabsent condition. To our surprise, we observed no group differences in target-absent search times, even for search displays that elicit high levels of uncertainty. We therefore focus our report here on this replication failure.

# Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All analysis scripts and anonymized data are available at github.com/Noamsarna/ocd\_visual\_search. The order and timing of experimental events were determined pseudorandomly by the Mersenne Twister pseudorandom number generator, initialized to ensure registration time-locking (Mazor et al., 2019). A detailed preregistration document for Experiment 1 is available at github.com/Noamsarna/ocd\_visual\_search/tree/ main/experiments/Experiment1.

# Method

Participants. The research was approved by the Research Ethics Committee of Tel-Aviv University (Study ID No. 0004169-1). A total of 1,007 participants were recruited via Prolific (https://prolific.co/) and selected based on the following criteria: an acceptance rate above 95%, no participation in previous pilot studies, use of non-Safari browsers, and being native English speakers located in the UK. The median completion time for the entire experiment was 14 min. Participants were paid £2 for their participation, equivalent to an hourly wage of £8.57. Participants were divided into high (OC+) and low (OC-) OCD tendencies groups based on their scores in the Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002; see below). The OC+ group consisted of individuals in the highest quartile of the OCI-R scores distribution, and the OC- group comprised individuals in the lowest quartile of this distribution. The entire sample (N = 1,007) completed the visual-search task. Because of higher than expected exclusion rate and in deviation from our preregistered plan to collect 250 participants in each group, our final sample included 213 OC+ participants and 220 OC- participants.

The average age of the total sample was 30.41 years (SD = 5.7). Half of the sample identified as female. In terms of ethnicity, the majority (84%) identified as White, followed by Asian (7%), Black (4%), and mixed/ other (5%). The predominant nationality was UK (93%). Employment status was predominantly full-time (62%), followed by part-time (16%).

Visual-search task. The visual-search task consisted of four blocks, each containing 24 trials of searching for either a closed or an open square. The task began with a practice phase consisting of one block with six trials. Each display was presented for a maximum of 10 s or until a response was received. During the practice phase, feedback about accuracy was given after each trial: If the response was correct, the word "Correct!" appeared on the screen for 1 s; if the response was wrong, the word "Wrong" appeared on the screen for 5 s. In the main part of the experiment, no feedback was given, as was the case in the original paradigm (Toffolo et al., 2013). After completing the practice, participants looked for either a closed square among rotated open squares ("hard search"; Fig. 1, Main Part, right) or for a rotated open square among closed squares ("easy search"; Fig. 1, Main Part, left). The difference in difficulty between these two search types is due to a search asymmetry for open/

# Visual Search Trial structure: target present or absent? + Image: Search 1000-2000 ms 200 ms RT; up to 10,000 ms 1000/5000 ms; only in practice trials

# Practice

# Main Part



# **Difficulty Estimation**



**Fig. 1.** Overview of experimental design. (Top) Each visual-search trial started with a centered black fixation cross. (Middle) Practice: Participants completed practice trials, searching for a rotated "T" among rotated "L"s in six-trial blocks until they achieved a minimum accuracy of 0.83 (no more than one error). (Middle) Main part: The primary experiment comprised 96 trials in four blocks, with the target identity changing after two blocks. Each 24-trial block followed a 2 × 2 design, manipulating set size (nine or 25) and target presence (present/absent). (Bottom) Search difficulty estimation: Participants used their mouse to rate search difficulty on a continuous scale. In questions about target-present searches, the target was marked with a red square.

closed edges (Treisman & Gormican, 1988). We further manipulated target presence and set size, resulting in a  $2 \times 2 \times 2$  design (search type: easy search or hard search;

target: present/absent; set size: nine or 25). Block order was counterbalanced between participants, and trial order within individual blocks was fully randomized (Fig. 1).

#### Measures

**OCI-R.** The OCI-R is an 18-item self-report measure of OCD-symptom severity. Responders are asked to rate their level of distress pertaining to 18 statements in the past month on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*). The OCI-R has been shown to have good validity, test–retest reliability, and internal consistency in both clinical (Foa et al., 2002) and nonclinical samples (Hajcak et al., 2004).

Depression, Anxiety and Stress Scales-21. The Depression, Anxiety and Stress Scales-21 (DASS-21; Lovibond & Lovibond, 1995) is a 21- item self-report questionnaire that is divided into three seven-item subscales to measure dimensional components of depression, anxiety, and stress. Each individual item refers to the respondent's experiences over the past week and is evaluated on a 4-point scale, ranging from 0 (the item does not apply to me at all) to 3 (the item applies to me very much or most of the time). The DASS-21 has shown high reliability, validity, and internal consistency within both clinical groups and a community sample (Antony et al., 1998; Henry & Crawford, 2005). In this study, only the depression and anxiety scales were used. We used the depression and anxiety subscales to control for nonspecific effects associated with OCD tendencies.

### Procedure

A static version of Experiment 1 can be accessed at https://noamsarna.github.io/ocd\_visual\_search/experi ments/demos/exp1/. Participants were first instructed about the experiment's structure, which comprised three parts: a visual-search task, questions about the visual search, and the two inventories: OCI-R and DASS-21. Then, they received written instructions about the visual-search task. After completing the visual-search task, participants were asked to rate the difficulty of noticing the presence or absence of a certain target among different distractors (for more information about this, see the appendix in the Supplemental Material available online). Following the difficulty estimation, participants completed the OCI-R and DASS-21. We included two attention-check questions among the OCI-R items, asking participants to select a certain answer ("If you read this question, check the option 'Not at all'").

# Data analysis

Participants were excluded from the analysis if they made more than 15% errors in the main part of the experiment or for having extremely fast or slow reaction times (below 100 ms) in more than 25% of the trials. Participants were also excluded if they failed one or more of the attention checks. In total, 109 out of 1,007 participants were excluded from the analysis. For the remaining participants, error trials and trials with response times (RTs) below 100 ms were excluded from the response-time analysis.

#### Results

We focus our report here on our failure to replicate a group difference in target-absent search times, even for search displays that elicit high levels of uncertainty. For all additional analysis from our preregistered hypotheses, see the appendix in the Supplemental Material.

To directly replicate group differences in targetabsent response times (RTs; Toffolo et al., 2013, 2014, 2016), we focused on the difficult search with the larger set size (set size = 25). We conducted a mixed-effects analysis of variance (ANOVA) with mean RT as the dependent variable, group (OC+ vs. OC-) as a betweensubjects variable, and target presence (present vs. absent) as a within-subjects variable. Specifically, we examined the interaction effect testing the hypothesis that the mean RT difference between the OC+ and OCgroups would be significantly more pronounced in target-absent trials. Contrary to our expectations, the analysis did not reveal a significant interaction between group and target presence, F(1,431) = 1.62, p = .203, Cohen's d = 0.12 (Fig. 2, Experiment 1). A null result was also obtained in a correlation analysis, pooling data from all participants and treating OCI-R scores as a continuous variable (see preregistered Hypothesis 9 in the Supplemental Material). To quantify the evidence for the null, we conducted a Bayesian *t* test setting the scale at the averaged effect size found in Toffolo et al. (2013, 2014), reflecting a belief that if present, group differences should be negative and moderate in size (Rouder et al., 2009). A one-sided Bayesian independentsamples *t* test produced a Bayes factor of  $BF_{10} = 0.09$ , providing strong evidence for the null hypothesis of no group differences.

We conducted several additional analyses that examined the interaction between the OC groups and the presence of the target. First, at the group level, we performed multilevel regression, accounting for anxiety and depression. We found no interaction between the OC groups and the presence of the target (preregistered Hypothesis 10),  $\hat{\beta} = 8.62$ , 95% confidence interval [CI] = -21.50, 38.74, *t*(463.56) = 0.56, *p* = .575. Likewise, when we focused on the initial trials of the task, before any accumulated experience (preregistered Hypothesis 8), we found no interaction between group and target presence in a mixed-effects ANOVA, *F*(1,361) = 0.93,



**Fig. 2.** Results from Experiment 1, Experiment 2, and Toffolo et al. (2013, 2014). Mean response times for target-absent and target-present trials (*x*-axis). Error bars represent the standard error of the mean. Shapes represent the obsessive compulsive groups. Circle = individuals with high OCD tendencies (OC+); triangle = individuals with low OCD tendencies (OC-).

p = .335. Furthermore, at the group level, we observed no significant differences between the groups in their self-reported measures of task difficulty. A group difference in accuracy did reach significance such that the OC+ group (M = 0.94) was overall less accurate than the OC- group (M = 0.95), t(425.75) = 3.37, p < .001.This difference did not replicate in Experiment 2. To extend our analysis to the entire sample, encompassing the four OCI-R quartiles, we replaced the group variable (OC+, OC-) with the full range of OCI-R scores. In this analysis, we still found no interaction between OCI-R scores and the presence of the target (preregistered Hypothesis 9),  $\hat{\beta} = -0.07$ , 95% CI = -2.48, 2.35), t(941.54) = -0.05, p = .957. Detailed calculations and results for all these hypotheses are provided in the appendix in the Supplemental Material for further reference.

# **Experiment 2**

In Experiment 1, target-absent search times were not significantly slower in OC+ compared with OC- individuals. Although this stands in contrast to previous reports (Toffolo et al., 2013, 2014, 2016), our results differed from those of the original study in other respects as well. Most notably, search times in this study ( $\approx$ 4.5 s for target-absent and  $\approx$ 2.6 s for target-present) were overall shorter compared with those in Toffolo et al. (2013) ( $\approx$ 5.5 for target-absent and  $\approx$ 3.5s for target-present). We therefore considered the possibility that the task used in Experiment 1 may have been less challenging and potentially insufficient to elicit doubt and trigger checking behavior. To test this, Experiment 2 employed the original stimuli

from Toffolo et al. (2013). The preregistered analysis plan is available at https://github.com/Noamsarna/ocd\_ visual\_search/tree/main/experiments/Experiment2. In Experiment 2, we conducted a further power analysis mirroring the methods of Toffolo et al. (2013), using their data, and adopting a bootstrap approach to determine an adequately powered sample size, as detailed in the preregistration document for Experiment 2. We employed the Mersenne Twister pseudorandom number generator to ensure that our preregistration preceded data collection (Mazor et al., 2019).

#### Method

A total of 226 participants were recruited via Prolific. To maximize statistical power for a group comparison, we invited former participants whose OCI-R scores were in the top or bottom quartile in Experiment 1. In line with our preregistered stopping rule, we kept data collection until we had invited all participants in the first and fourth quartiles from our previous experiment (n = 220 and n = 213, respectively). Participants completed the OCI-R questionnaire again in the present study (the test-retest reliability for the OCI-R yielded a Pearson's correlation coefficient of r = .87, p < .001) and were assigned to the OC+/OC- groups based on the original cutoff scores from Toffolo et al. (2013; OCI-R total score  $\geq$  17 for the OC+ group; OCI-R total score  $\leq$  5 for the OC- group). Our final sample consisted of 110 OC+ participants and 68 OC- participants. The entire experiment took 12 min to complete, and participants were paid £1.8 for their participation, equivalent to an hourly wage of £9.

### Procedure

A static version of Experiment 2 is available at https:// noamsarna.github.io/ocd\_visual\_search/experiments/ demos/exp2/. Experiment 2 was similar to Experiment 1 with the following exceptions. First, we used the original stimuli from Toffolo et al. (2013). The visualsearch task consisted of one block of 50 individual search displays, each containing 25 elements. The search task was more challenging because of a larger search grid, which meant larger distances between stimuli and reduced stimulus size. Second, Experiment 2 did not include an assessment of perceived difficulty, comprising only the visual search followed by the same questionnaires as in Experiment 1. Third, to make it identical to Toffolo et al., practice trials in Experiment 2 (four per block) involved the same stimuli as the main blocks. Fourth, participants were instructed to press the spacebar to move from the fixation-cross screen to the search-display screen, at which point, the search display appeared immediately. Finally, the visual-search part of the experiment included only the hard-search type: detecting a closed square among open squares.

# Data analysis

Because Experiment 2 served as a direct replication, we adopted the same rejection criteria as Toffolo et al. (2013) so that participants were excluded if their error count exceeded 2.5 *SD* from the mean error rate of the entire sample. As in Experiment 1, participants were also excluded from the analysis if they failed to answer correctly one or more attention-check questions.

# Results

In contrast to Toffolo et al. (2013), in which presenceabsence differences in RT were more pronounced among OC+ participants, in our replication sample, the one-tailed t test of the interaction contrast (using the difference in search times as a dependent variable) was not significant, t(144.88) = 1.41, p = .081, Cohen's d =0.22, providing no evidence for the expected interaction. Note that the numeric trend of the interaction in our sample was driven by shorter RT in the OC+ group compared with the OC- group in target-present trials rather than by longer RT for target-absent responses (Fig. 2, Experiment 2). This pattern is different from that reported by Toffolo et al., in which OC+ participants were slower in both search types, but particularly in target-absent searches (Fig. 2; Toffolo et al., 2013). Unlike in Experiment 1, we observed no differences in accuracy between the groups (OC+: M = 0.83; OC-: M =(0.83), t(138.01) = -0.44, p = .664. Finally, a one-sided Bayesian independent-samples *t* test produced a Bayes factor of  $BF_{10} = 0.13$ , providing moderate evidence for the null hypothesis of no group differences.

# Discussion

In two preregistered, large-sample studies, we found no evidence of prolonged search time among OC+ participants in target-absent trials, contrary to previous findings by Toffolo and colleagues (2013, 2014, 2016).

The most notable difference between our experiments and those conducted by Toffolo and colleagues (2013, 2014, 2016) lies in our use of an online setting versus their use of in-person lab experiments. Completing tasks online as opposed to a laboratory setting generates more "technical noise," that is, unexplained variance driven by technical variation. However, previous studies have suggested that such noise has minimal impact on RT differences in perceptual tasks. In a study comparing RT distributions from a lab-based Matlab and an online JavaScript experiment, the results revealed near-identical RTs between the two setups (de Leeuw & Motz, 2016). The JavaScript experiment showed a consistent delay of around 25 ms, which had minimal impact on the sensitivity to RT changes because of experimental manipulations.

Furthermore, in our study, participants completed the visual-search task using a range of computers and displays rather than in a controlled lab environment with a fixed screen, as in Toffolo et al. (2013, 2014, 2016). Yet simulation studies have demonstrated minimal impact of technical variance on statistical power and the precision of effect-size estimates (Brand & Bradley, 2012). Key behavioral findings in psychology, including those observed in the Stroop and flanker tasks, and effects reliant on much smaller time constants, such as attentional blink and subliminal priming, have been successfully replicated in web-based studies (Crump et al., 2013). Specifically, a recent online visual-search study reported significant RT variations between experimental conditions, with a focus on smaller time constants than those anticipated in our paradigm (Mazor & Fleming, 2022). Particularly strong evidence for the comparability of lab-based versus web-based findings comes from a study that used a fully randomized design for RT effects (Hilbig, 2016). The results showed that a word-frequency effect (manifested in different RT) was comparable in magnitude across all three conditions. Taken together, these studies show that although some variations between settings in RT exist, they are minor, especially when the outcome measure is RT alterations because of experimental manipulations.

Additional differences between our research and Toffolo et al.'s (2013, 2014, 2016) studies that could interact with OCD tendencies are anonymity and demographic variations. It is plausible that the anonymity afforded by online studies could lead to participants feeling less responsible for study outcomes than identifiable psychology students who meet experimenters in person. Moreover, participants in Toffolo et al.'s studies were monitored by an eye-tracker camera, a factor that has been suggested to reduce reliance on internal cues, such as metacognitive experiences (Noah et al., 2018). Given the sensitivity of OC+ individuals to personal responsibility (Salkovskis, 1985) and the heightened sense of anonymity in online studies, the transition to an online setting may have attenuated group differences in checking behavior.

Our failure to find an association between obsessivecompulsive tendencies and inference about absence may appear inconsistent with well-known clinical manifestations of OCD, such as those observed in hit-andrun OCD. However, our experimental operationalization of inference about absence differed from these clinical manifestations in two important ways. First, we did not manipulate perceived responsibility, a key feature of hit-and-run OCD and one that is posited to play a key role in OCD more generally (Salkovskis, 1985). Second, in the clinical example of hit-and-run OCD, the compulsion is associated more with a recollection of an event rather than its direct experience. Indeed, most findings related to reduced confidence in individuals with OCD have been observed in relation to memory rather than perception (for a review, see Dar et al., 2022). More research is needed to elucidate the interaction of these two features with inference about absence in OCD.

Finally, this replication attempt puts into action several key features of replicable science (Tackett et al., 2017). Our study included detailed preregistration with hypotheses, power analysis, analysis plan, and exclusion criteria. We used the preregistration time-locking tool (Mazor et al., 2019), thereby guaranteeing that our registration preceded the data-collection process. Furthermore, our study represents the first independent replication attempt. Finally, we have made our raw data, analysis scripts, and task codes publicly available. Beyond a contribution to the experimental literature on OCD, we hope this report may serve as a helpful reference for reproducible and open clinical-psychological science.

# Conclusion

The presented findings diverge from those of previous studies by Toffolo and colleagues (2013, 2014, 2016) because we were unable to replicate the effect of prolonged search time for OC+ participants in target-absent trials. At the very least, this replication failure indicates that the original effect may be constrained to a specific setting, thus limiting its generalizability to other contexts. More broadly, our results advocate for the application of open-science practices in clinicalpsychology research to foster methodological integrity and ensure the reliability of findings.

#### Transparency

Action Editor: DeMond M. Grant

Editor: Jennifer L. Tackett

Author Contributions

**Noam Sarna:** Conceptualization; Formal analysis; Investigation; Methodology; Writing – original draft.

**Matan Mazor:** Conceptualization; Formal analysis; Investigation; Methodology; Writing – review & editing.

**Reuven Dar:** Conceptualization; Resources; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This study was supported by the Israeli Science Foundation (Grant 1156/11). MM is supported by a post-doctoral research fellowship from All Souls College, University of Oxford.

**Open Practices** 

This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psy chologicalscience.org/publications/badges.



#### **ORCID** iDs

Noam Sarna (D) https://orcid.org/0009-0008-2533-2016 Matan Mazor (D) https://orcid.org/0000-0002-3601-0644 Reuven Dar (D) https://orcid.org/0000-0003-2178-0483

#### Acknowledgments

We are grateful to Marieke Toffolo for her cooperation, data sharing, and support. Special thanks to Ori Levit for her assistance with the statistical analysis. The Israeli Science Foundation had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication. An extended version of this article was previously made available as a preprint, which can be accessed at 10.31234/osf.io/rmz7n.

#### **Supplemental Material**

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/21677026241258380

#### References

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, *10*(2), 176–181.

- Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., Castelo-Branco, M., & Voon, V. (2015). Evidence accumulation in obsessive-compulsive disorder: The role of uncertainty and monetary reward on perceptual decision-making thresholds. *Neuropsychopharmacology*, 40(5), 1192–1202. https://doi.org/10.1038/npp.2014.303
- Brand, A., & Bradley, M. T. (2012). Assessing the effects of technical variance on the statistical outcomes of web experiments measuring response times. *Social Science Computer Review*, 30(3), 350–357. https://doi.org/ 10.1177/0894439311415604
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE*, 8(3), Article e57410. https://doi.org/10.1371/journal.pone.0057410
- Dar, R. (2004). Elucidating the mechanism of uncertainty and doubt in obsessive-compulsive checkers. *Journal of Behavior Therapy and Experimental Psychiatry*, 35(2), 153–163. https://doi.org/10.1016/j.jbtep.2004.04.006
- Dar, R., Lazarov, A., & Liberman, N. (2021). Seeking proxies for internal states (SPIS): Towards a novel model of obsessive-compulsive disorder. *Behaviour Research and Therapy*, 147, Article 103987. https://doi.org/10.1016/j .brat.2021.103987
- Dar, R., Sarna, N., Yardeni, G., & Lazarov, A. (2022). Are people with obsessive-compulsive disorder underconfident in their memory and perception? A review and meta-analysis. *Psychological Medicine*, 52(13), 2404–2412. https://doi.org/10.1017/S0033291722001908
- de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a web browser? Comparing response times collected with JavaScript and psychophysics toolbox in a visual search task. *Behavior Research Methods*, 48(1), 1–12. https://doi .org/10.3758/s13428-015-0567-2
- 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.
- Hajcak, G., Huppert, J. D., Simons, R. F., & Foa, E. B. (2004). Psychometric properties of the OCI-r in a college sample. *Behaviour Research and Therapy*, 42(1), 115–123. https:// doi.org/10.1016/j.brat.2003.08.002
- Hauser, T. U., & Allen, M., NSPN Consortium, Rees, G., & Dolan, R. J. (2017). Metacognitive impairments extend perceptual decision making weaknesses in compulsivity. *Scientific Reports*, 7(1), 6614. https://doi.org/10.1038/ s41598-017-06116-z
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44(2), 227–239. https://doi.org/10.1348/ 014466505X29657
- Hilbig, B. E. (2016). Reaction time effects in lab-versus webbased research: Experimental evidence. *Behavior Research*

*Methods*, 48(4), 1718–1724. https://doi.org/10.3758/ s13428-015-0678-9

- Hyman, B. M., & Pedrick, C. (2010). *The OCD workbook: Your guide to breaking free from obsessive-compulsive disorder*. New Harbinger Publications.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343.
- Mazor, M. (2021). *Inference about absence as a window into the mental self-model*. PsyArXiv. http://doi.org/10.31234/osf.io/zgf6s
- Mazor, M., & Fleming, S. M. (2022). Efficient search termination without task experience. *Journal of Experimental Psychology: General*, 151(10), 2494–2510. https://doi.org/ 10.1037/xge0001188
- Mazor, M., Friston, K. J., & Fleming, S. M. (2020). Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *eLife*, 9, Article e53900. https://doi.org/10.7554/eLife.53900
- Mazor, M., Mazor, N., & Mukamel, R. (2019). A novel tool for time-locking study plans to results. *European Journal of Neuroscience*, 49(9), 1149–1156.
- Mazor, M., Moran, R., & Fleming, S. M. (2021). Stage 2 registered report: Metacognitive asymmetries in visual perception. *Neuroscience of Consciousness*, 2021(1), Article niab025. https://doi.org/10.1093/nc/niab025
- Noah, T., Schul, Y., & Mayo, R. (2018). When both the original study and its failed replication are correct: Feeling observed eliminates the facial-feedback effect. *Journal of Personality and Social Psychology*, *114*(5), 657–664. https://doi.org/10.1037/pspa0000121
- Rasmussen, S. A., & Eisen, J. L. (1989). *Clinical features and phenomenology of obsessive compulsive disorder*. SLACK Incorporated.
- Reed, G. F. (1985). Obsessional experience and compulsive behavior: A cognitive-structural approach. Academic Press.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237. https://doi.org/10.3758/PBR.16.2.225
- Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, 23(5), 571–583.
- Sarig, S., Dar, R., & Liberman, N. (2012). Obsessive-compulsive tendencies are related to indecisiveness and reliance on feedback in a neutral color judgment task. *Journal of Behavior Therapy and Experimental Psychiatry*, 43(1), 692–697.
- Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., Oltmanns, T. F., & Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, 12(5), 742–756. https://doi.org/10.1177/1745691617690042
- Toffolo, M., Hout, M. A., van den Engelhard, I. M., Hooge, I. T. C., & Cath, D. C. (2014). Uncertainty, checking, and intolerance of uncertainty in subclinical obsessive

compulsive disorder: An extended replication. *Journal of Obsessive-Compulsive and Related Disorders*, *3*(4), 338–344. https://doi.org/10.1016/j.jocrd.2014.08.004

- Toffolo, M., Hout, M. A., van den Engelhard, I. M., Hooge, I. T. C., & Cath, D. C. (2016). Patients with obsessivecompulsive disorder check excessively in response to mild uncertainty. *Behavior Therapy*, 47(4), 550–559.
- Toffolo, M., Hout, M. A., van den Hooge, I. T. C., Engelhard, I. M., & Cath, D. C. (2013). Mild uncertainty promotes

checking behavior in subclinical obsessive-compulsive disorder. *Clinical Psychological Science*, *1*(2), 103–109. https://doi.org/10.1177/2167702612472487

- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95(1), 15–48.
- van den Hout, M., & Kindt, M. (2003). Repeated checking causes memory distrust. *Behaviour Research and Therapy*, 41(3), 301–316.