Clinical Psychology Review 75 (2020) 101807

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Why check? A meta-analysis of checking in obsessive-compulsive disorder: Threat vs. distrust of senses

Asher Y. Strauss[a,*], Isaac Fradkin[a], Richard J. McNally[b], Omer Linkovski[c], Gideon Emanuel Anholt[d], Jonathan D. Huppert[e]

[a] Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel
[b] Department of Psychology, Harvard University, Cambridge, MA, USA
[c] Department of Psychiatry and Behavioral Sciences, Stanford University Medical Center, Stanford, CA, USA
[d] Department of Psychology, Ben-Gurion University of the Negev, Beer Sheva, Israel
[e] Helen and Sam Beber Chair of Clinical Psychology, Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel

HIGHLIGHTS

- This meta-analysis examines checking among OCD individuals and controls as measured experimentally.
- Effects of checking did not vary depending on whether conditions were threatening or neutral.
- Effects of checking were found in tasks requiring perceptual decision-making but not reasoning.
- Results support models which associate OCD checking with distrust of senses.

ARTICLE INFO

Keywords:
Obsessive-compulsive disorder
Checking
Information seeking
Threat
Reasoning
Perception

ABSTRACT

Compulsive checking is the most common ritual among individuals with obsessive-compulsive disorder (OCD). Yet, other than uncertainty, the variables prompting checking are not fully understood. Laboratory studies suggest that task conditions - whether threatening (anxiety-relevant) or neutral, and task type - whether requiring perceptual or reasoning decision-making – may be influential. The purpose of our meta-analysis was to compare OCD participants and healthy controls on experimental tasks involving uncertainty in which a behavioral measure of checking was obtained. Four databases were searched. Twenty-two studies met the inclusion criteria, including 43 conditions comparing 663 OCD participants to 614 healthy controls. Due to the dependent structure of the data a robust variance estimation analysis approach was used. Overall effects were similar for neutral and threatening conditions. However, OCD participants responded with greater checking compared to controls on perceptual tasks, but not on reasoning tasks. Results support previous reports suggesting that OCD checking can be observed in neutral conditions, possibly posing as a risk factor for a checking vicious cycle. In addition, our results support OCD models which focus on checking as stemming from interference with automatic processes and distrust of sensory modalities.

1. Introduction

Checking is the most common compulsion among individuals suffering from obsessive compulsive disorder (OCD), affecting as many as 80% of them at some point in their lives (Ruscio, Stein, Chiu, & Kessler, 2010). Many patients repetitively check features of the external environment such as stoves and locks (Rachman, 2002), but checking may involve other acts such as checking the news to make sure one has not been involved in a car accident, or counting text messages to one's partner to ascertain one's love (e.g., Doron, Derby, Szepsenwol, & Talmor, 2012). Checking often drives other clinical aspects of OCD. Indeed, seeking reassurance from others and excessively requiring information prior to decision-making can qualify as compulsive checking (Coleman, Pieterefesa, Holaway, Coles, & Heimberg, 2011; Foa et al., 2003).

By checking, people with OCD say they aim to prevent harm, reduce uncertainty (Rachman, 2002), or diminish feelings that things are not just right (Coles & Ravid, 2016). Paradoxically, repetitive checking

* Corresponding author at: Department of Psychology, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel.
E-mail address: asher.strauss@mail.huji.ac.il (A.Y. Strauss).

https://doi.org/10.1016/j.cpr.2019.101807
Received 19 February 2019; Received in revised form 19 October 2019; Accepted 20 November 2019
Available online 12 December 2019
0272-7358/

often increases uncertainty rather than decreasing it (e.g., van den Hout & Kindt, 2003; van den Hout, van Dis, van Woudenberg, & van de Groep, 2017). A model for the onset and maintenance of compulsive checking has been proposed (Nedeljkovic & Kyrios, 2007; Toffolo, van den Hout, Hooge, Engelhard, & Cath, 2013): In an uncertain situation, individuals with a vulnerability for feeling uncertainty experience greater uncertainty which leads them to engage in checking. Checking momentarily reduces the distress caused by the uncertainty, but paradoxically increases feelings of uncertainty, thereby initiating a vicious cycle: feelings of uncertainty prompt checking and checking perpetuates feelings of uncertainty. In a meta-analysis reviewing the effects of repeated checking van den Hout et al. (2017) analyzed 28 studies and reported aggregated effects of checking on memory confidence, vividness and detail. Since the focus of that review was the consequences of checking, all included studies examined checking as an *independent* variable. The focus of the current meta-analysis is to examine checking as a *dependent* variable, focusing on the conditions that prompt individuals with OCD to react with repeated checking.

1.1. Eliciting checking behaviors in the lab

Researchers have devised diverse tasks to elicit checking in the laboratory, reporting mixed results. In an early report, Milner, Beech, and Walker (1971) found that, relative to controls, obsessional participants requested more repetitions of an auditory signal before indicating whether they had heard it. In addition, Volans (1976) showed that OC checkers checked jars to a greater degree before deciding on the probabilistic beads task, though group differences were significant only after controlling for neuroticism. In this task, participants view two jars with opposing ratios of two colors of beads (e.g., jar A: 15% blue 85% red, jar B: 85% blue 15% red; Huq, Garety, & Hemsley, 1988). Based on the cumulative sampling of beads selected from one of the jars (unknown to the participants), participants are asked to decide from which jar the beads were selected. The main dependent variable in this task is the number of beads requested by a participant prior to deciding. Using the same task, Fear and Healy (1997), found that OCD participants requested nonsignificantly more beads compared to healthy controls. Later studies using the probabilistic beads task described above have generally failed to replicate Volans' early finding (e.g., Jacobsen, Freeman, & Salkovskis, 2012; Pélissier & O'Connor, 2002) and one report found the individuals with OCD requested significantly fewer beads (Grassi et al., 2015). Jacobsen et al. (2012) modified the beads task to pertain to threatening and OCD relevant information, by replacing the beads with emotionally salient words, but still no significant differences emerged.

Other tasks have also yielded conflicting results. When asked to judge whether two images were identical in a delayed matching-to-sample task, OCD participants checked images to a greater degree compared to healthy controls (Rotge et al., 2008). A subsequent study replicated this finding (Jaafari et al., 2013), whereas another did not (Rotge et al., 2015). Using a different task, Arntz, Voncken, and Goosen (2007) found that individuals with OCD checked excessively when requested to sort pills according to color, though only when they were led to believe that their sorting decisions had crucial implications for medication use in India. Foa et al. (2003) assessed the number information cards participants requested when asked to imagine making decisions in various scenarios. Although both healthy participants and those with OCD requested more cards in high-threat scenarios, only the OCD group did so for low-threat scenarios. Toffolo, van den Hout, Engelhard, Hooge, and Cath (2016) found that individuals with OCD checked more on a visual-search task compared to both anxiety and healthy controls. In this task, participants were required to decide if a visual target cue was either absent or present in a visual field. Checking behaviors were measured by counting the number of eye fixations. In sum, checking behaviors measured in the lab seem to vary considerably and may depend on the type of task and the conditions under which it is administered. The focus of our meta-analysis is to test whether two variables – condition valence and task type – moderate these effects.

1.2. Conditions that elicit checking behavior (neutral vs. threat)

According to Rachman's (2002) cognitive theory, compulsive checking will mainly occur under high-responsibility conditions – a prediction consistent with Salkovskis's model of OCD (1985, 1998) stressing the role of inflated responsibility in the etiology and maintenance of OCD symptoms. Indeed, Arntz et al.' (2007) finding that individuals with OCD checked excessively only in the high responsibility condition supports Rachman's assertion. Other conditions likely to provoke excessive checking are those related to risk of harm or other OCD-related themes not directly related to responsibility (e.g., disgust). However, as reviewed above, some studies show that checking may be elevated even under neutral conditions (e.g., Jaafari et al., 2013; Rotge et al., 2008). Even mild uncertainty may produce group differences in checking (Toffolo et al., 2016). Examining the moderating role of the condition in which the task was administered may clarify some of the mixed results above. Indeed, experimental investigations of checking in the lab provide an opportunity to dissociate checking behaviors from their typical clinical context.

1.3. Task types that elicit checking behaviors (reasoning vs. perceptual)

Whereas the conditions under which the task is administered, or the valence of the task stimuli may moderate checking, other characteristics pertaining to the task's cognitive requirements may play a role such as whether it requires a reasoning versus perceptual decision. Data gathering for making a deliberative stepwise decision, such as in the probabilistic beads task or Foa et al.'s (2003) card task, where participants must logically weigh probabilities or alternatives, may differ from data gathering for making a perceptual decision, such as in delayed matching-to-sample task (Rotge et al., 2008) or in the visual search task (Toffolo et al., 2016). In a perceptual task, data gathering involves using and trusting sensory modalities compared to reasoning tasks where data gathering involves deliberative thinking. Previous studies have shown that OCD participants exhibited lower confidence related to perception (Hermans et al., 2008) and that perseverative and compulsive staring produces paradoxical effects on perceptual confidence, as van den Hout, Engelhard, de Boer, du Bois, and Dek (2008) found. Indeed, distrust of sensory information is a key element of the theory proposed by O'Connor, Aardema, and Pélissier (2005) regarding OCD doubt, and plays a role in other theoretical accounts of OCD doubt, such as the Seeking Proxies for Internal States theory (SPIS; Lazarov, Liberman, Hermesh, & Dar, 2014). Whereas the SPIS model does not focus only on perception but pertains to all internal states, the most compelling empirical support concerns tasks involving subjective somatosensory perceptions (i.e., muscle tension; Lazarov, Cohen, Liberman, & Dar, 2015; Lazarov, Dar, Oded, & Liberman, 2010; Lazarov et al., 2014). Examining whether checking behavior may depend on task type in this manner is therefore important for elucidating reported mixed results.

1.4. Previous related meta-analyses

There have not been any meta-analytic reviews focusing on checking in OCD, distinguishing between neutral and threat provoking conditions and task type – whether the task demands a perceptual or reasoning decision. However, two meta-analyses examined questions related to checking in OCD. The first study examined variants of the probabilistic beads task to determine whether there were specific decision-making deficits in psychotic patients versus other groups. The beads task has been extensively used with psychotic and delusional disorders as a measure of making decisions in the absence of sufficient checking (potentially the opposite effect of what would be expected in

OCD, i.e., “jumping to conclusions” [JTC]). So, Siu, Wong, Chan, and Garety (2016) examined the JTC bias across studies using variants of the beads task and compared 3 groups: psychotic disorders, non-psychotic disorders, and healthy controls. They reported a moderate negative effect size (−0.60) for psychotic participants compared to healthy controls, indicating that participants with psychotic disorders overall gathered less data prior to decision-making. In addition, they reported a small, nonsignificant positive effect (0.24) for a comparison of a small subgroup of OCD participants to healthy controls. Unfortunately, this estimate was based solely on three studies (Fear & Healy, 1997; Jacobsen et al., 2012; Reese, McNally, & Wilhelm, 2011), and did not include other studies using the beads task (e.g., Grassi et al., 2015), other tasks measuring data gathering in decisions pertaining to reasoning (such as Foa et al., 2003), or any of the tasks requiring perceptual decision-making mentioned above.

In another meta-analysis, De Putter, Van Yper, and Koster (2017) reviewed different methods of inducing OCD symptoms in the lab across several symptom dimensions, reporting an effect size of 0.58 for inducing OCD symptoms in the checking dimension with negligible heterogeneity. This effect size comprised experiences related to OCD (e.g., urges to check, anxiety), not just checking per se. Moreover, given their focus on symptom induction, their search did not include terms that would detect some studies addressing perceptual checking (e.g., Clair et al., 2013) and reasoning checking (e.g., Jacobsen et al., 2012). Therefore, a comprehensive meta-analysis focusing on checking explicitly is warranted to elucidate mixed results and examine possible moderators.

1.5. Measuring checking behaviors in the lab

Checking can be measured in various ways. A common measure is to count the number of times participants request additional information prior to making their decision as in the beads task or in Foa et al. (2003) card task. A similar method is to count the number of inspections of a stimulus such as in the delayed-matching-to-sample task (Rotge et al., 2008). A variant involves using eye-tracking technology to count fixations on a stimulus (e.g., Bucarelli & Purdon, 2016; Toffolo et al., 2016) or counting the number of gaze movements between stimuli (Jaafari et al., 2013). Checking can also be measured in a virtual reality environment by coding virtual behaviors (e.g., Kim et al., 2008).

Potential confounds arise in two other important conditions when measuring checking: when feedback is provided and when checking comes at a cost. In some tasks, participants receive feedback while checking, such as in the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) where participants either win or lose money while drawing cards from different decks in search of the more advantageous deck. Although drawing more cards from less advantageous decks may reflect a tendency to check excessively, it may also reflect deficiencies in feedback processing or other decision-making processes. Alternatively, other tasks penalize participants for checking such as in a decrementing reward condition in an information sampling task (Chamberlain et al., 2007) where the number of points the participant stands to win decreases for every check. Here, too, group differences may not reflect checking per se, but rather other processes such as sensitivity to punishment. Therefore, in our meta-analysis we included only tasks in which checking was measured independently of feedback and with no cost.

1.6. The current meta-analysis

Our aim was to examine the extent of checking among individuals with OCD compared to healthy controls in experimental lab tasks, and to test whether condition valence and task type moderated effects of checking.

Condition valence denotes the emotional tone of the task (i.e., neutral versus threat). Threatening conditions were those designed to provoke anxiety, feelings of heightened responsibility, or other OCD-typical concerns or that involved decisions involving judgments about emotionally negative stimuli. Neutral conditions and stimuli lacked these features.

Task type refers to the main cognitive domain of the decision-making process. We have chosen the broad term “perceptual tasks” to describe tasks that require data-gathering as part of a low-level decision-making process involving the detection of sensory information (e.g., Toffolo et al.'s, 2016 visual search task). We have chosen the broad term “reasoning tasks” to describe tasks that require data gathering as part of a high-level decision-making processes such as using these data for logical weighing probabilities or to deliberate between different alternatives (e.g., Huq et al., 1988 probabilistic beads task).

An extensive search was performed to obtain all studies comparing checking across neutral and threatening conditions, including various tasks, pertaining to both perceptual and reasoning decision-making processes. We hypothesized that effect would be larger for threatening conditions compared to neutral conditions. In addition, given that different accounts of OCD have suggested that OCD doubt may pertain to the perceptual domain, and based on the non-systematic review above, we hypothesized that effects would be larger for perceptual tasks compared to reasoning tasks.

2. Method

2.1. Inclusion and exclusion criteria

Studies eligible for inclusion were required to include a direct comparison between adult participants with a diagnosis of OCD and healthy controls on a task involving direct measurement of checking. Only studies including a behavioral measure other than decision time (reaction time) were included, as decision time is not a process-pure measure of checking; it may also signify distraction, difficulties retaining attention or other processes.[1] We excluded studies involving provision of performance feedback to participants to eliminate confounds of learning. Therefore, we excluded tasks in which exploration and exploitation of options were not separate (such as in the IGT). Finally, checking needed to be free of cost to eliminate confounds of punishment sensitivity or other motivations to reduce checking. Treatment studies were included if tasks were administered before or during the initial stages of therapy. Study reports were required to be in English.

2.2. Search strategies

Guidelines for conducting meta-analyses were followed (Moher, Liberati, Tetzlaff, & Altman, 2009). The PubMed, Web of Science, PsycNet, and ProQuest databases were searched through March 2019 for published and unpublished studies including the following keywords (asterisk denotes truncation designed to capture grammatical variability): “obsessiv*, compulsiv*” or “OCD” with, “checking”, “uncertainty”, “decision-making”, “decision-making”, “reasoning bias*”, “data gathering”, “evidence gathering”, “liberal acceptance”, “draw* to decision*”, “DTD”, “jump* to conclusion*”,”JTC”, “bead* task*”, “fish* task*”, “survey task*”, “word task*”, “information sampling”, “eye track*”, “random-dot-motion”, “random dot motion”, “probabili* task*”. Databases' built-in categorization systems were used when appropriate (e.g., MeSH terms in PubMed).

[1] Some of these confounds can be controlled for using computational models. Indeed, Banca et al. (2015) reported that OCD patients required more evidence on a perceptual task compared to healthy controls using these methods. However, given that almost all studies measuring response times did not use computational modeling, we chose not to include reaction time as a measure of checking.

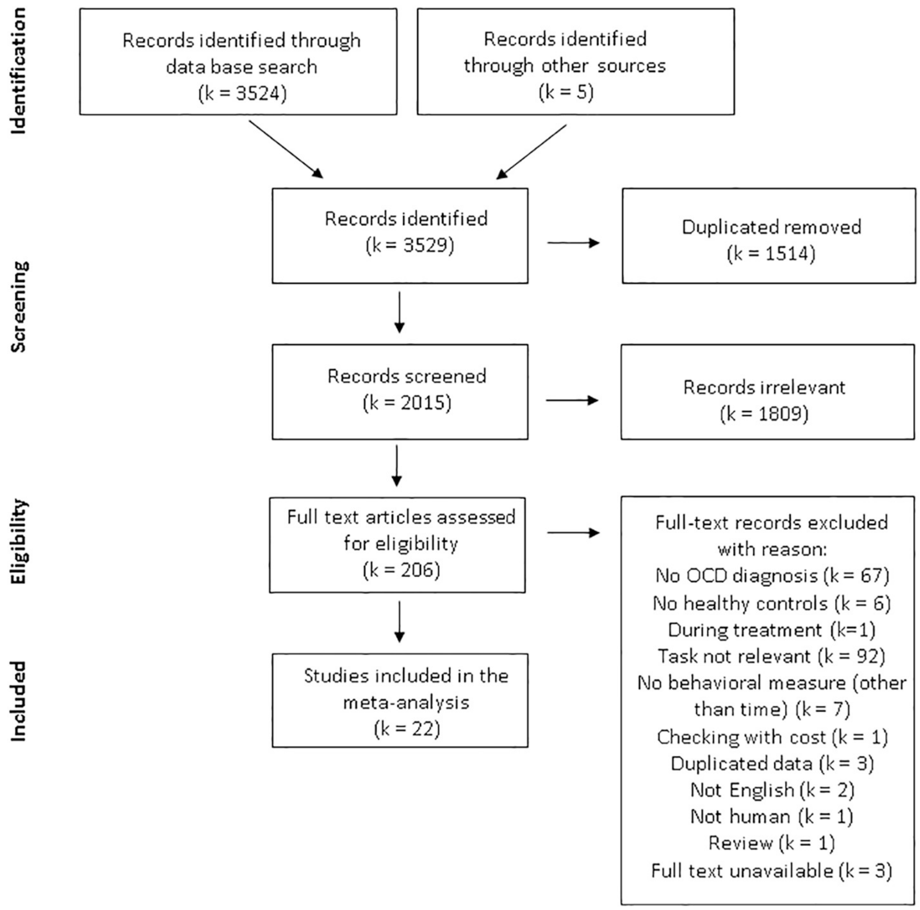

Fig. 1. PRISMA study selection flow diagram.

A PRISMA flow chart (Moher et al., 2009) depicting the study selection process can be found in Fig. 1. After the identification of potential studies was completed, abstracts were screened manually with the help of the Covidence systematic review software (Babineau, 2014), and studies that were clearly irrelevant (e.g., irrelevant populations, irrelevant tasks) were excluded, resulting in a total of 206 studies. Full texts were not found for 3 studies after an extensive search as well as direct contact with authors, leaving a total of 203 studies, 22 of which were found eligible (see Fig. 1). When crucial information was lacking, authors were contacted by e-mail (k = 8), with sufficient information attained from six authors. When data were unavailable via authors, but appeared in a figure in the article, we used WebPlotDigitizer (Drevon, Fursa, & Malcolm, 2017) to extract the data. This was done with two studies (k = 2): From Foa et al. (2003) only means were extracted from the plot, and standard deviations were later computed from the tables. For Rotge et al. (2008) both means and standard errors were extracted from the plot. For Rotge et al. (2008), 20 of the 50 patients with OCD were receiving treatment, though none had recovered (i.e., YBOCS < 8). We include both studies but conducted sensitivity analysis to see whether these studies affected our findings.

2.3. Data extraction and coding

We extracted the following information from studies when available: (a) technical details (publication status, year, country of study); (b) participants' demographic characteristics (age, gender, education); (c) participants' clinical characteristics (age at onset, illness duration, symptom severity, psychotropic medication use, comorbid diagnoses, and OCD subtype distribution); (d) condition valence (neutral versus threat); (e) type of checking task (reasoning vs. perceptual); and (e) effect size measures.

Coding was conducted by the first author. Reliability was tested via a second reviewer who screened a third of the full-text articles for eligibility, with interrater agreement reaching 98% (*Cohen's* κ = 0.9). These studies were later coded independently by this reviewer, with interrater agreement (intraclass correlation coefficient) regarding effect sizes (i.e., Hedges's g values) and categorization of tasks approaching 1.

2.4. Statistical analysis

Twenty-two studies were included in the analysis. Ten studies included more than one condition in which checking was measured (e.g., target-absent vs. target-present or responsibility vs. non-responsibility stimuli) yielding together 43 effect sizes (*mean number of conditions* = 1.95, *range* = 1–5). Hedges's g effect sizes, comparing the performance of OCD participants and healthy controls, were calculated with positive values indicating greater checking behavior in OCD participants. Three studies separated OCD checkers from OCD non-checkers (Clair et al., 2013; Kim et al., 2012; Rotge et al., 2008). For these studies, prior to calculating effect sizes, groups were collapsed to one OCD group and pooled means and standard deviations were computed using group sizes as weights. Analysis was conducted using a robust variance estimation method (RVE; Hedges, Tipton, & Johnson, 2010) with conditions nested within tasks. An RVE model was preferred due to multiple conditions within studies, which were assumed to be correlated. RVE produces valid point estimates, standard errors, confidence intervals and significance tests when effect sizes are non-independent without needing to model this correlation (Fisher & Tipton, 2015). This method is the most suitable for our analysis given that none of the studies with multiple conditions reported correlation statistics between conditions. It can be used even with a small number of studies when proper corrections to the degrees of freedom are used (Tipton, 2015). Using simulations, Tipton (2015) has shown that estimators perform best when degrees of freedom after corrections are four or greater. When the degrees of freedom are less than four, the probability of a Type I error may exceed 0.05 (confidence intervals are too narrow). Whereas RVE estimation requires specifying a value for the within-participant correlation (ρ), simulation analysis has shown that RVE coefficients are relatively insensitive to changes in ρ (e.g., Ishak, Platt, Joseph, & Hanley, 2008; Tipton, 2013). Nevertheless, Hedges et al. (2010) recommend follow-up sensitivity analyses to ensure that results are insensitive to changes in ρ. Analysis was conducted using R software version 3.4.4 and the "robumeta" package (Fisher & Tipton, 2015) version 2.0, a R-package for RVE meta-analysis. Two main moderators were examined: Condition valence (neutral vs. threat) and task type (perceptual vs. reasoning). These moderators were entered as predictors in the meta-regression RVE model. To estimate the percentage of heterogeneity explained by each of the moderators (R^2), we computed the relative reduction in between-study variance (τ^2) from an intercept-only model to a model with a moderator as a predictor. To test the moderating role of symptom severity, we modeled effect sizes of checking as predicted by the Yale Brown Obsessive Compulsive Scale (YBOCS) scores. To test the moderating role of gender, age, depression and anxiety, we modeled effect sizes of checking by standardized group differences in these variables (odds ratio for gender and Hedges's *g* for age, anxiety and depression). Studies' methodological quality was assessed independently by the first and second authors using a modified version of the Agency for Healthcare Research and Quality assessment tool (AHRQ; Williams, Plassman, Burke, Holsinger, & Benjamin, 2010; see supplementary material for the modified tool). Differences between raters were later resolved by consensus between both raters. To examine the possibility and potential effect of publication bias, we performed Egger regression tests accompanied by interpretation of funnel plots.

3. Results

3.1. Preliminary analysis and coding

Twenty-two studies were included. Studies originated in a variety of countries (Canada: 2, France: 5, Italy: 1, South Korea: 3, The Netherlands: 3, United Kingdom: 4, United States: 4). These studies included 43 effect sizes, 633 participants with OCD and 614 healthy controls. Gender statistics were missing for two studies. Among the 20 remaining studies 52% of the OCD participants were female, compared to 54% of the healthy controls. Age statistics were missing for one study. The average age of the 21 remaining studies was 35.94 for the OCD participants and 34.04 for the healthy controls. In all studies but one, participants were either matched or a lack of difference between groups was found for gender (k = 21), and all but four studies reported matching or lack of difference between groups in age (k = 18). Matching or lack of difference for education was reported in eleven studies (k = 11) and for IQ in eight (k = 8). Indeed, we found no significant difference between groups in gender (OR = 0.92, $z = -0.66$; p = .51, k = 20) or age (β = 0.29, z = 1.30, p = .19, k = 21). Diagnosis was obtained using a structural interview (e.g., Structured Clinical Interview for Diagnostic Statistical Manual [SCID], Mini-International Neuropsychiatric Interview [MINI]) in all studies but one (Fear & Healy, 1997). The YBOCS was administered in 19 studies, with a mean total score of 23.26 (SD = 2.75, *range* = 18.90–28.44). Twelve studies reported percentage of participants with checking compulsions with an average percentage of 65% (*range* = 33% - 100%, k = 12).Only some studies reported education (k = 8) comorbidity (k = 10), medication use (k = 9), duration of illness (k = 9) and age of onset (k = 6).

All 22 studies included in our analysis are presented in Table 1, along with the different conditions administered in the task, the checking dependent variable and classification of condition valence and task type. Twenty-nine of these effect sizes were obtained in conditions coded as *neutral*. Fourteen of these effect sizes were obtained in conditions coded as *threatening*.

Eleven studies included *reasoning tasks*: the beads task (k = 9); a survey task / word task (an emotional salient variation of the beads task) was administered in addition to the original beads task in three of the beads task studies (k = 3), an information sampling task (k = 1), and a card task (k = 1). Eleven other studies included *perceptual tasks*: a pill sorting task (k = 1), a delayed matching-to-sample task (k = 3), an image comparison task (k = 1), a visual search task (k = 2), a virtual reality checking task (k = 3), and a virtual reality game (k = 1).

Anxiety control groups were used in only five studies (Arntz et al., 2007; Hezel, Stewart, Riemann, & McNally, 2019; Jacoby, Abramowitz, Buck, & Fabricant, 2014; Pélissier & O'Connor, 2002; Toffolo et al., 2016), and no study included a depression control group. However, approximately half of the studies reported depression (k = 11) or anxiety measures (k = 10). Depression measures included: Beck Depression Inventory II (BDI-II; k = 4), Montgomery-Asberg Depression Rating Scale (MADRS; k = 1), Hamilton Depression Rating Scale (HDRS; k = 3), Hospital Anxiety and Depression Scale (HADS-D; k = 2) and the Center for Epidemiologic Studies Depression Scale Revised (CESD-R; k = 1). Anxiety measures included: state portion of the State–Trait Anxiety Inventory (STAI-S; k = 1), Beck Anxiety Inventory (BAI; k = 3), Hamilton Anxiety Rating Scale (HARS; k = 4), Hospital Anxiety and Depression Scale (HADS-A; k = 2). Depression scores for healthy controls were missing in two studies (Kim et al., 2010, 2012), and anxiety scores for healthy control were missing in three studies (Botta et al., 2018; Jaafari et al., 2013; Kim et al., 2012). To compute Hedges's *g* for studies when examining anxiety or depression as a moderator, we imputed these scores by using the average scores for healthy controls from other studies included in this analysis which used the same measure.

3.2. Overall mean effect size for checking behavior and moderator analysis

The overall mean effect size for checking across all studies was 0.26 (SE = 0.11, $t_{(21.4)}$ = 2.48, *95% CI* = [0.04, 0.48], p = .02) - a small effect relative to those reported in the meta-analyses reviewed above (De Putter et al., 2017; So et al., 2016). Considerable heterogeneity was found (τ^2 = 0.19, I^2 = 71%). We examined whether this heterogeneity was moderated by condition valence (neutral vs. threat), and task type (perceptual vs. reasoning). Condition valence did not account for any systematic variance (τ^2 = 0.19, I^2 = 71%, $R^2 \approx$ 0%), nor did it significantly predict checking (β = 0.05, $t_{(13.7)}$ = 0.60, p = .56). In contrast, adding task type to the model dramatically decreased heterogeneity (τ^2 = 0.11, I^2 = 56%, R^2 = 43%), and it significantly moderated checking (β = 0.26, $t_{(16.8)}$ = 2.82, p = .01). Given the moderating role of task type, we analyzed the two task types separately.

3.3. Perceptual tasks

The forest plot for perceptual decision tasks appears in Fig. 2. The overall mean effect size for checking across all perceptual decision-making tasks was 0.51 (SE = 0.10, *95% CI* = [0.28, 0.73], $t_{(10)}$ = 5.04, p < .01, k = 11) and some heterogeneity emerged (τ^2 = 0.06, I^2 = 44%). Within perceptual tasks, neutral condition mean effect size was 0.53 (SE = 0.14, *95% CI* = [0.18, 0.87], $t_{(5.71)}$ = 3.80, p < .01, k = 7), larger than the threat condition mean effect size of 0.40 (SE = 0.13, *95% CI* = [0.06, 0.75], $t_{(4.29)}$ = 3.21, p = .03, k = 6). However, the difference between the conditions was not significant ($\beta = -0.05$, $t_{(8.2)} = -0.24$, p = .82). Indeed, as apparent in Fig. 2, the overall effect for perceptual tasks, and the effects of threatening and neutral conditions were quite similar.

Table 1
Study characteristics.

Study	Task	Task type	Condition	Condition valence	Measure of checking
Fear and Healy (1997)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
Pélissier and O'Connor (2002)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
Foa et al. (2003)	Card task	Reasoning	Low risk	Neutral	Number of information cards requested
			High risk	Threat	
			OCD-relevant	Threat	
Arntz et al. (2007)	Pill sorting task	Perceptual	Low responsibility	Neutral	Average behavioral checklist
			High responsibility	Threat	
Chamberlain et al. (2007)	Information sampling task	Reasoning	Fixed reward	Neutral	Number of boxes opened
Kim et al. (2008)	Virtual reality checking task	Perceptual		Threat	Number of checks in the virtual reality
Rotge et al. (2008)	Delayed matching-to-sample task	Perceptual	Identical	Neutral	Number of returns to the first image
		Perceptual	Non-identical	Neutral	
Kim et al. (2010)	Virtual reality checking task	Perceptual		Threat	Number of checks in the virtual reality
Reese et al. (2011)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
			60–40	Neutral	
	Survey task	Reasoning	85–15 self-relevant	Threat	Number of words drawn
			60–40 self-relevant	Threat	
Jacobsen et al. (2012)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
			60–40	Neutral	
	Words task	Reasoning	60–40 Neutral	Neutral	Number of words drawn
			60–40 Carelessness	Threat	
			60–40 Social evaluation	Threat	
Kim et al. (2012)	Virtual reality checking task	Perceptual		Threat	Number of checks in the virtual reality
Clair et al. (2013)	Delayed matching-to-sample task	Perceptual		Neutral	Number of returns to the first image
Jaafari et al. (2013)	Image comparison task	Perceptual	Non-identical	Neutral	Number gaze moves between drawings
Jacoby et al. (2014)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
			60–40	Neutral	
			44–28-28	Neutral	
Grassi et al. (2015)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
			Identical	Neutral	
Rotge et al. (2015)	Delayed matching-to-sample task	Perceptual		Threat	Number of returns to the first image
Toffolo et al. (2016)	Visual search task	Perceptual	Target-present	Neutral	Number of fixations
			Target-absent	Neutral	
Morein-Zamir et al. (2017)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
van Bennekom et al. (2017)	Virtual reality game	Perceptual		Threat	Number of checks in the virtual reality
O'Connor et al. (2005)	Beads task	Reasoning	85–15	Neutral	
Botta et al. (2018)	Visual search task	Perceptual	Orthographic	Neutral	Number of fixations
			Semantic	Neutral	
			Neutral	Neutral	
			Obsession-related	Threat	
Hezel et al. (2019)	Beads task	Reasoning	85–15	Neutral	Number of beads drawn
			60–40	Neutral	
	Survey task	Reasoning	85–15 self-relevant	Threat	Number of words drawn
			60–40 self-relevant	Threat	

3.4. Reasoning tasks

The forest plot for reasoning tasks appears in Fig. 3. The overall mean effect size for checking across all reasoning studies was −0.02 (SE = 0.15, 95% CI = [−0.35, 0.31], $t_{(9.79)} = -0.137$, p = .89, k = 11), and considerable heterogeneity emerged ($\tau^2 = 0.17$, $I^2 = 67\%$). Within reasoning tasks, the mean effect size for neutral conditions was 0.01 (SE = 0.17, 95% CI = [−0.34, 0.39], $t_{(9.83)} = 0.08$, p = .94, k = 11), whereas the mean effect size for threatening conditions was 0.16 (SE = 0.19, 95% CI = [−0.45, 0.77], $t_{(2.93)} = 0.85$, p = .46, k = 4). Note that this last estimate was based on only four studies. Indeed, the difference between the conditions was not significant ($\beta = 0.30$, $t_{(4.02)} = 1.19$, p = .30). Nevertheless, the effect size for threatening reasoning tasks is only a third of that for neutral perceptual tasks, further strengthening the moderating role of task type.

3.5. Additional moderators

There was no overall significant effect of gender differences between groups on checking ($\beta = 0.02$, $t_{(7.68)} = 0.19$, p = .91, k = 20), nor was there a significant effect of group differences in age ($\beta = -0.09$, $t_{(1.23)} = -4.56$, p = .10, k = 21), symptom severity ($\beta = 0.02$, $t_{(8.27)} = 0.36$, p = .73, k = 19), depression ($\beta = 0.17$, $t_{(2.05)} = 0.58$, p = .62, k = 11) or anxiety ($\beta = 0.04$, $t_{(3.13)} = 0.63$, p = .57, k = 10). Similar findings were obtained when examining these moderators within the perceptual tasks (for gender: $\beta = -0.46$, $t_{(3.1)} = -1.69$, p = .19, k = 10; for age: $\beta = -0.41$, $t_{(2.4)} = -0.49$, p = .67, k = 10; for symptom severity: $\beta = -0.07$, $t_{(3.0)} = -1.62$, p = .20, k = 9; for depression $\beta = -0.08$, $t_{(1.73)} = -0.64$, p = .60, k = 6; and for anxiety $\beta = 0.03$, $t_{(2.94)} = 0.53$, p = .65, k = 7) and the reasoning tasks (for gender: $\beta = -0.15$, $t_{(4.5)} = -0.48$, p = .65, k = 10; for age: $\beta = -0.04$, $t_{(1.3)} = -0.77$, p = .55, k = 11; for symptom severity: $\beta = 0.09$, $t_{(4.6)} = 1.83$, p = .13, k = 10; for depression $\beta = -0.92$, $t_{(2.01)} = 2.43$, p = .14, k = 5; and for anxiety $\beta = 0.00$, $t_{(1.00)} = 0.00$, p = .99, k = 3).

3.6. Sensitivity analysis

As recommended by Hedges et al. (2010), we verified our main results were insensitive to changes in the within subject correlation (ρ). The finding that task type (perceptual vs. reasoning) significantly moderated checking was insensitive to alterations of ρ (*range of p-values* = .0108–0.0109), whereas the finding that condition valence (neutral vs. threat) was not a significant moderator was too insensitive to alterations in ρ (*range of p-values* = .5576–0.5615). Indeed, overall estimates of effect size for perceptual tasks and reasoning tasks were insensitive to these alterations in ρ (aggregated Hedge's g range for perceptual tasks: 0.5087–0.5088; for reasoning tasks:

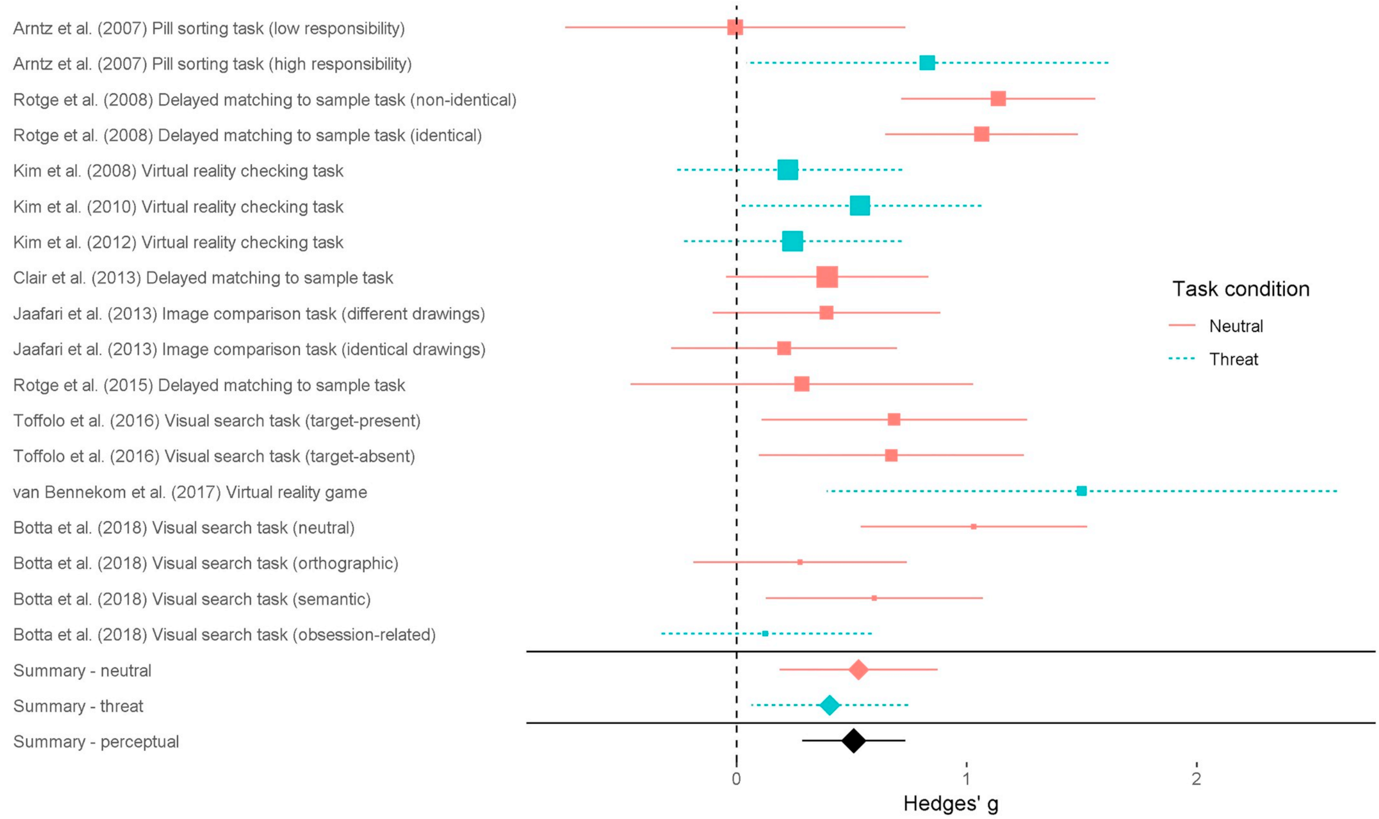

Fig. 2. Forest plot depicting effect sizes for checking in neutral and threatening perceptual decision-making studies (k = 11).

[−0.0205]–[−0.0201]).

Next, we examined if our main results were affected by including data from the Foa et al. (2003) and Rotge et al. (2008) studies. Because some statistics were unavailable in these publications, we extracted them via WebPlotDigitizer (Drevon et al., 2017). Results were generally insensitive to removing these two studies. Task type significantly moderated the effect ($\beta = 0.26$, $t_{(12.9)} = 3.98$, $p < .01$), but not condition valence ($\beta = 0.04$, $t_{(11.4)} = 0.69$, $p = .50$). The effect sizes for both types of tasks decreased, but led to a similar conclusion. Specifically, the overall effect size for perceptual tasks was 0.41 ($SE = 0.06$, *95% CI* = [0.27, 0.56], $t_{(8.4)} = 6.58$, $p < .01$, $k = 10$) and for reasoning tasks −0.16 ($SE = 0.13$, *95% CI* = [−0.41, 0.17], $t_{(8.6)} = -0.91$, $p = .39$, $k = 10$).

3.7. Methodological quality of studies and publication bias

A detailed description of each individual study's methodological quality assessed via the modified AHRQ assessment tool appears in Table S1 in the supplementary material. As reported above, all studies but one ($n = 21$) ascertained patients' diagnoses by using a structured diagnostic interview, and almost all studies matched or reported a lack of significant differences on gender ($n = 21$) and age ($n = 18$). Almost all studies ($n = 19$) either did not report missing data or took adequate steps to minimize missing data bias. Analytic methods were found adequate in most studies, apart from five studies, with small samples (under 20 participants per group) which might have been advised to use non-parametric statistical tests due to probable violations of normality. There were several consistent problems across many studies. All but two studies ($n = 20$) sampled OCD and healthy control participants from different sampling frames, thus increasing the risk of selection bias. Whereas this is a considerable threat, studies sampling both clinical participants and controls from the same sampling frame are rare in the field. All but one study ($n = 21$) did not provide a priori sample size justification, rendering the risk of being underpowered for detecting meaningful effects. This stress the need for aggregating data across studies via meta-analysis as done in this report. All but one study ($n = 21$) did not explicitly state that experimenters were blind to group condition, thus increasing the risk of expectancy effects. However, this risk can be reduced by administering computerized tasks. Indeed, the majority of the studies in our report ($n = 16$) did so.

To test for publication bias, we conducted an Egger regression test by regressing studies' standard errors on effect sizes for checking. This test was nonsignificant ($\beta = 2.06$, $t_{(6.4)} = 1.37$, $p = .22$). Fig. 4 presents funnel plots for both perceptual tasks and reasoning tasks. Egger regression tests were nonsignificant for perceptual task ($\beta = 1.05$, $t_{(3.3)} = 0.54$, $p = .62$) or reasoning tasks ($\beta = 3.70$, $t_{(4.37)} = 1.89$, $p = .13$). Nonetheless, the funnel plot for reasoning tasks (right plot, Fig. 4) depicts asymmetry consistent with our finding that task type affects checking, but that there is no overall difference between OCD and healthy controls in reasoning tasks. Given this publications bias, effects for reasoning tasks seem exaggerated, implying the presence of unpublished studies reporting null or negative results. A slight asymmetry emerged for perceptual tasks, driven mainly by van Bennekom, Kasanmoentalib, de Koning, and Denys's (2017) study which reported a large effect and a large standard error (bottom right point in the left plot in Fig. 4). To ensure our main results were not affected by this study, we re-ran the main analysis without this study. Task type continued to significantly moderate the effect ($\beta = 0.24$, $t_{(16.8)} = 2.67$, $p = .02$) but not condition valence ($\beta = 0.02$, $t_{(12.3)} = 0.24$, $p = .82$). The overall effect size for perceptual tasks without this study was 0.48 ($SE = 0.10$, *95% CI* = [0.25, 0.70], $t_{(9.36)} = 4.18$, $p < .01$, $k = 10$).

4. Discussion

The aim of this meta-analysis was to examine checking in OCD. Whereas checking is the most common compulsion reported among

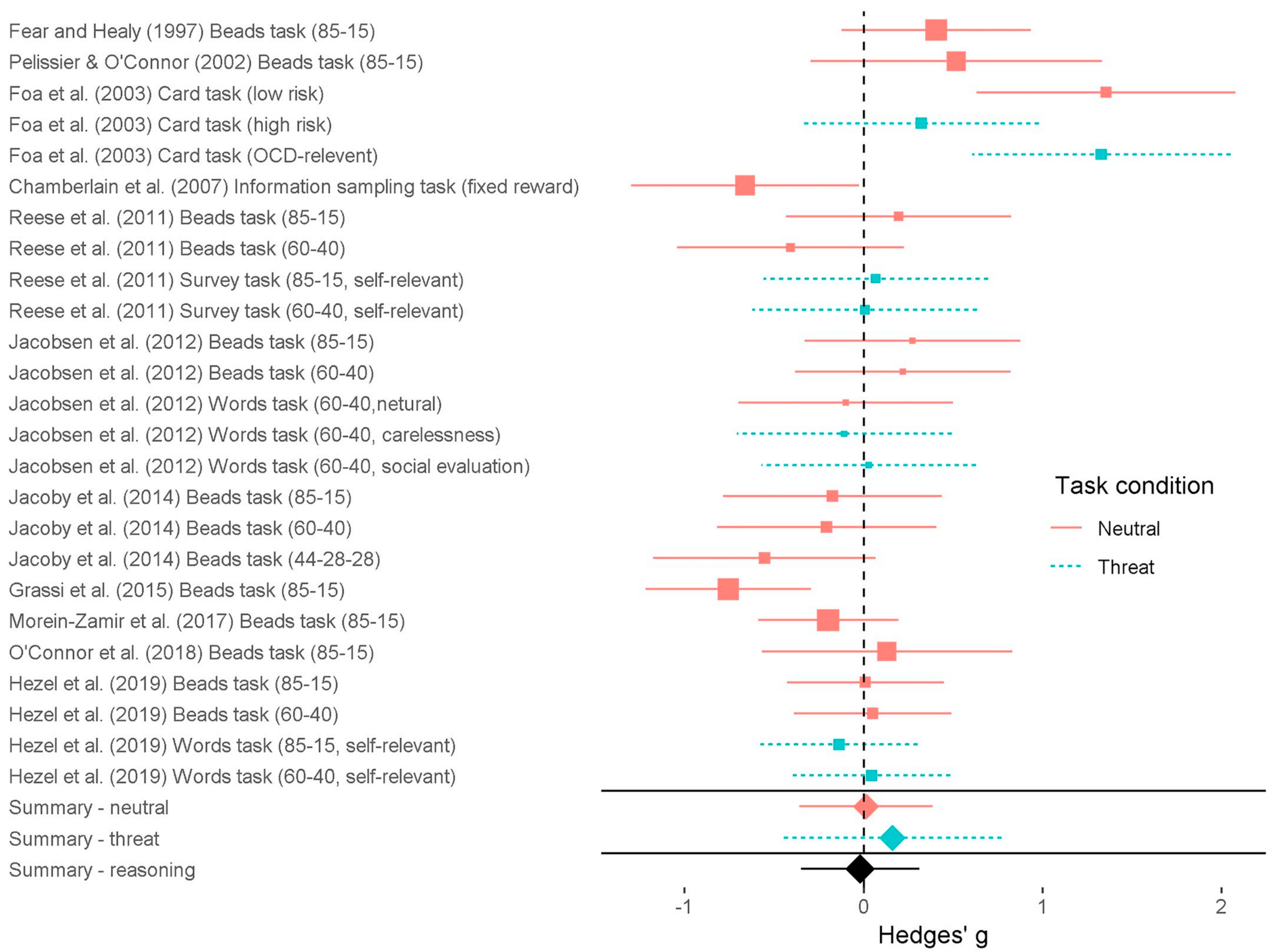

Fig. 3. Forest plot depicting effect sizes for checking in neutral and threatening reasoning decision-making studies (k = 11).

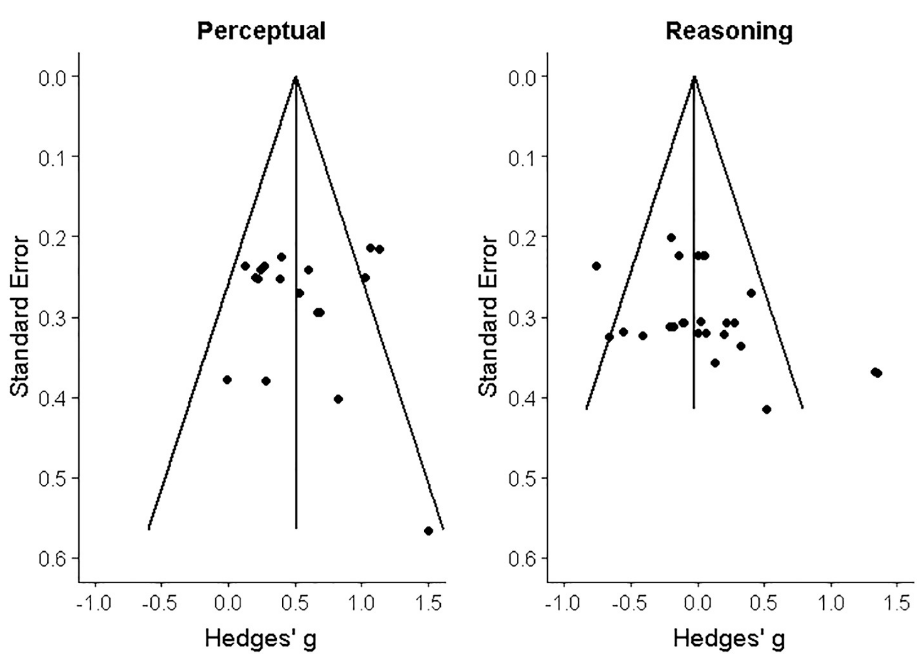

Fig. 4. Funnel plots for effect sizes of checking for perceptual (left plot) and reasoning (right plot). Funnels correspond to the overall model mean effect size (vertical black line) and 95% confidence intervals (diagonal black lines).

individuals suffering from OCD, the conditions under which it occurs remain uncertain. Our findings suggest individuals with OCD and healthy controls do not differ on neutral versus threat conditions, nor do the groups differ on tasks requiring reasoning. However, individuals with OCD do exhibit excessive checking in perceptual decision tasks.

According to Rachman's (2002) cognitive account of compulsive checking, compulsive checking should be elicited only in face of elevated sense of responsibility. In this sense, our finding that OCD participants checked more compared to controls even in conditions in which responsibility was not elevated appears to contradict Rachman's assertion. Instead, our findings are in line with recent suggestions that OCD checking may be elicited also in the mere presence of mild uncertainty (Toffolo et al., 2016). That being said, Rachman's account describes a phenomenological account of pathological checking in OCD, and the extrapolation to lab behavior on experimental tasks requires more inquiry and elaboration. In addition, Rachman (2002) suggested that checking may paradoxically increase the sense of responsibility. Therefore, even if mild uncertainty prompts checking, a sense of responsibility may be crucial for eliciting the self-perpetuating mechanism of compulsive checking. Indeed, an experimental design testing the hypothesized vicious cycle of checking, together with the conditions which set this cycle in to motion is warranted. Whereas this report attempts to review the conditions in which checking may begin, and other reports have documented the effects of checking on distrust of memory (see van den Hout et al., 2017 for a meta-analysis), no study to the best of our knowledge has tested the vicious cycle hypothesis in a single experiment. Whereas a tendency to check excessively one's perception may be a risk factor for pathological compulsive checking, how excessive checking in neutral conditions may evolve to pathological checking remains unresolved.

Whereas our findings suggest that individuals with OCD and healthy controls do not differ on neutral versus threat conditions, it is important to note that all studies used the same neutral stimuli for all participants. Only Clair et al. (2013) used idiosyncratic stimuli by obtaining an individualized score for the non-anxiogenic stimuli. Finding neutral tasks for OCD may be challenging, since some individuals with OCD may find neutral stimuli anxiety provoking, either by personal associations or by the mere fact that making wrong decisions is greatly aversive for them. Moreover, given the heterogeneity in the phenomenology of this disorder, even threating stimuli may be perceived by some individuals with OCD as less anxiogenic than what was intended. Finally, although the idiosyncratic approach has its virtues, it introduces uncontrolled "materials effects" as the stimuli subjects encounter vary in diverse ways that may affect performance. Ideally, investigators should include a manipulation check or concurrent physiological measures to assess how threatening each participant regards the stimuli.

The fact the individuals with OCD did check more in tasks requiring perceptual processing is in line with Salkovskis's (1998) assertion that obsessional thought is associated with an "attempt to monitor closely and take control over processes that would otherwise operate in automatic and well-practiced ways" (p.40). Perceptual tasks rely on automatic processes to a greater degree compared to the deliberative and deductive processes employed in reasoning tasks. Indeed, deficiencies stemming from over-control have been reported using lab experiments. For example, individuals with OCD have been found to exhibit difficulties in adjusting control on the Stroop task (Kalanthroff, Anholt, & Henik, 2014). In another study, using a modified Serial Reaction Time task, individuals with OCD were found to favor controlled over automatic processing (Soref, Liberman, Abramovitch, & Dar, 2018).

Results are also in line with two contemporary theories pertaining to OCD doubt: the Inferential Confusion (IC) theory (O'Connor et al., 2005) and the Seeking Proxies for Internal States (SPIS) theory (Lazarov et al., 2010). According to the IC model, OCD doubt stems from a *distrust of the senses* and is maintained by an inverse inference which follows an idiosyncratic subjective narrative rather than general formal reasoning. Data gathering according to O'Connor and colleagues is not about gathering more data in an accumulative, deliberative fashion (which would predict greater data gathering in the reasoning tasks), but about using data in an idiosyncratic fashion which is characterized by disregarding information already provided by the senses (which is in line with the results found in the perceptual tasks). An alternative view is that OCD patients doubt *internal representations* of external stimuli, leading them to require more confirmation by resampling the external stimulus itself. This view is consistent with the SPIS model by Lazarov and colleagues, which asserts that OCD doubt stems from an attenuated accessed to internals states, and that patients compensate by seeking external proxies. In this sense, increased checking in perceptual decision-making tasks reflects an attempt to use external proxies for noisier or difficult to access internal representations.

One important feature which distinguishes the tasks in our analysis is the form of information seeking underlying the checking behavior. Reasoning tasks in general concern gathering more information (e.g., in the beads task: drawing a new bead), whereas perceptual tasks in a broad sense concern reviewing the same information (e.g., in the delayed matching-to-sample task: reviewing the first image again). It is difficult to differentiate between the type of task and the form of information seeking as perceptual tasks are more about confirming one's senses (or internal representations of sensory stimuli) and therefore more about information reviewing whereas most reasoning tasks involve a stepwise process of logical reasoning, accumulating data sequentially. Therefore, although both types of task involve uncertainty, they differ in terms of the form of information seeking. Indeed, our results comport with our clinical experience; many OCD patients seek reassurance about the information they have at hand rather than embark on a quest for new information. More research is needed to attempt to disentangle these processes.

Whereas some have suggested that in a sense "OC checkers may seem to be ideal scientists" (p. 667; Dar, Rish, Hermesh, Taub, & Fux, 2000), our results show that these individuals gather the same amount of evidence when it comes to logical reasoning. If OCD doubt is more about interference with automatic process as suggested by Salkovskis (1998) or involves going beyond reasonable doubt and disregarding sensory information as suggested by O'Connor et al. (2005), not only are OC individuals not ideal scientists, but they are rather the opposite: distrusting information and adopting pathological skepticism (Ron, Oren, & Dar, 2016).

To what extent are the reported effects specific to OCD or do they reflect a general effect associated with anxiety or depression? Whereas checking effects were not moderated by self-report measures of anxiety and depression, these results were based only on half of the sample and should be interpreted with caution given the small number of studies. Further supporting the unique effects of OCD, in the only two studies that examined performance on perceptual tasks that used an anxious control group, both found unique effects for OCD vs. anxious controls (Arntz et al., 2007; Toffolo et al., 2016). Two out of three studies in which reasoning tasks were examined using OCD vs. anxious controls vs. healthy controls did not report any group differences (Hezel et al., 2019; Jacoby et al., 2014). However, the third study reported a unique effect in the OCD group compared to the controls (Pélissier & O'Connor, 2002). Clarifying the unique attributes of OCD checking and doubt in comparison with other disorders is of great importance. Whereas this report suggests that distrust of sensory information is an important feature of OCD checking and doubt, to what extent this is unique to OCD compared to other disorders (e.g., GAD) is still unknown. Although, most clinicians report that doubt manifests itself differently in OCD compared to GAD, efforts to delineate these differences experimentally have gained only limited success. Indeed, the intolerance of uncertainty scale correlates only slightly higher in GAD compared to OCD (Gentes & Ruscio, 2011), and a recent review did not find robust difference in cognitive confidence in OCD compared to other disorders (Ouellet-Courtois, Wilson, & O'Connor, 2018).

An important limitation to the reasoning versus perceptual comparison which has been proposed here is that the vast majority of identified reasoning tasks are variants of the beads task. Only two other tasks required high-level decision-making processes (the information sampling task [Chamberlain et al., 2007] and the card task [Foa et al., 2003]). Accordingly, our conclusions may not generalize to other reasoning tasks. The beads task has been used extensively for studying reasoning in psychotic disorders (see reviews by Dudley, Taylor, Wickham, & Hutton, 2016; So et al., 2016) with reports of a moderate aggregated effect size. Compared to the usefulness of the beads task for capturing cognitive features of psychotic disorders, our report concludes that this task fails to do so in OCD. Moreover, the beads task captures several aspects of high-level decision making but not others. This report was limited to *checking behaviors* performed in the context of decision making and does not preclude that individuals with OCD may exhibit a unique form of reasoning as O'Connor et al. (2005) suggested. It is important to explore whether individuals with OCD exhibit excessive checking in tasks that capture other features of reasoning in future research. In this sense, regarding the perceptual vs. reasoning comparison, our report *highlights* that individuals with OCD exhibit excessive checking in perceptual tasks and *questions* the presence of this excessive behavior in reasoning tasks examined to date.

This review has several other limitations. First, despite our comprehensive search, we identified 22 studies. Therefore, we were underpowered for testing differences between sub-groups within the perceptual and reasoning tasks and were forced to keep our meta-regression models simple. Second, our classification of tasks in the threat category pooled together inflated responsibility, high-threat and OCD-relevant conditions. Unfortunately, there were insufficient studies to enable us to isolate the effects of these variables. More studies observing checking under various threat and anxiety provoking conditions are needed. Third, tasks that penalized participants for checking were excluded to eliminate confounds such as punishment sensitivity. Whereas this avoids methodological confounds, checking for OCD patients can be quite costly in terms of time, effort, and adverse social consequences. Ideally, cost-free vs. cost should be examined as a possible moderator. Unfortunately, only two studies examined checking with cost, thereby prohibiting their inclusion in the meta-analysis. Future studies should examine checking under cost. Last, all perceptual tasks that met inclusion criteria pertained to the visual domain. Individuals suffering from OCD distrust other sensory modalities such as touch and sound. Future studies are needed to examine whether these findings generalize to other perceptual domains.

If replicated and extended, our findings might have clinical implications. Cognitive restructuring is used to a lesser degree in OCD treatments and cognitive change is less of a predictor of symptom change in OCD compared to anxiety disorders (Anholt & Kalanthroff, 2013; Lorenzo-Luaces, Keefe, & DeRubeis, 2016). This is most probably due to the fact that OCD often takes an ego-dystonic nature, where the suffering individual perceives the obsession as intrusive to his reasoning process and irrational. The current report suggests this concept by stressing that deficits in OCD are less about systematic weighing of probabilities and more about trusting perceptual information. The ego-dystonic nature of distrust of senses is understandable because individuals with OCD have contradictory evidence in front of them (e.g., that the light is off), yet do not seem to fully register it. This sensory information is not typically obtained in a deliberative process of data gathering, but rather automatically (though the patient then recruits deliberative, systematic processes to try to compensate for the deficits in sensory information; c.f., Lazarov et al., 2015). Whereas some have suggested to focus the client on areas in which the mind goes beyond sensory information (O'Connor et al., 2005), others suggest helping the client accept the uncertainty and the inability to gain complete sense of certainty (e.g., Grayson, 2014). Nevertheless, both seem to agree that it is often futile to try to reason one's way out of an obsession.

Author Disclosure

Preparation of this manuscript was supported by the Israel Science Foundation (grant # 1698/15) to Jonathan D. Huppert.

Asher Y. Strauss headed, designed and executed this project. Isaac Fradkin helped design the methods and statistical analysis, as well as the interpretation of the results. Richard J. McNally participated in interpretation of the results as well as writing the manuscript. Omer Linkovski and Gidon Anholt helped design the meta-analysis search as well as the writing of the manuscript. Jonathan D. Huppert supervised the project, was very influential in forming the interpretation, and aided in the writing process.

None of the authors have any interests that might be interpreted as influencing this research, and APA ethical standards were followed in the conduct of this study.

Acknowledgments

The authors would like to thank Elad Zlotnick for fruitful comments and Lilach Kinberg for screening and coding articles for reliability. Preparation of this manuscript was supported by the Israel Science Foundation (grant # 1698/15) to Jonathan D. Huppert. The authors report no conflicting interests.

Appendix A. Studies included in the meta-analyses

Arntz, A., Voncken, M., & Goosen, A. C. A. (2007). Responsibility and obsessive–compulsive disorder: An experimental test. *Behaviour Research and Therapy*, *45*(3), 425–435. doi:https://doi.org/10.1016/j.brat.2006.03.016

Botta, F., Vibert, N., Harika-Germaneau, G., Frasca, M., Rigalleau, F., Fakra, E., … Jaafari, N. (2018). Visual search for verbal material in patients with obsessive–compulsive disorder. *Psychiatry Research*, *264*, 244–253. doi:https://doi.org/10.1016/j.psychres.2018.03.054

Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive–compulsive disorder and trichotillomania. *Neuropsychologia*, *45*(4), 654–662. doi:https://doi.org/10.1016/j.neuropsychologia.2006.07.016

Clair, A.-H., N'Diaye, K., Baroukh, T., Pochon, J.-B., Morgiève, M., Hantouche, E., … Mallet, L. (2013). Excessive checking for non-anxiogenic stimuli in obsessive-compulsive disorder. *European Psychiatry*, *28*(8), 507–513. doi:https://doi.org/10.1016/j.eurpsy.2012.11.003

Fear, C. F., & Healy, D. (1997). Probabilistic reasoning in obsessive–compulsive and delusional disorders. *Psychological Medicine*, *27*(01), 199–208.

Foa, E. B., Mathews, A., Abramowitz, J. S., Amir, N., Przeworski, A., Riggs, D. S., … Alley, A. (2003). Do Patients With Obsessive–Compulsive Disorder Have Deficits in Decision-Making? *Cognitive Therapy and Research*, *27*(4), 431–445. doi:https://doi.org/10.1023/A:1025424530644

Grassi, G., Pallanti, S., Righi, L., Figee, M., Mantione, M., Denys, D., … Stratta, P. (2015). Think twice: Impulsivity and decision-making in obsessive–compulsive disorder. *Journal of Behavioral Addictions*, *4*(4), 263–272. doi:https://doi.org/10.1556/2006.4.2015.039

Hezel, D. M., Stewart, S. E., Riemann, B. C., & McNally, R. J. (2019). Standard of proof and intolerance of uncertainty in obsessive-compulsive disorder and social anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *64*, 36–44. doi:https://doi.org/10.1016/j.jbtep.2019.02.002

Jaafari, N., Frasca, M., Rigalleau, F., Rachid, F., Gil, R., Olié, J.-P., … Vibert, N. (2013). Forgetting what you have checked: A link between working memory impairment and checking behaviors in obsessive-compulsive disorder. *European Psychiatry*, *28*(2), 87–93. doi:https://doi.org/10.1016/j.eurpsy.2011.07.001

Jacobsen, P., Freeman, D., & Salkovskis, P. (2012). Reasoning bias and belief conviction in obsessive-compulsive disorder and delusions: Jumping to conclusions across disorders? *British Journal of Clinical Psychology*, *51*(1), 84–99. doi:https://doi.org/10.1111/j.2044-8260.2011.02014.x

Jacoby, R. J., Abramowitz, J. S., Buck, B. E., & Fabricant, L. E. (2014). How is the Beads Task related to intolerance of uncertainty in anxiety disorders? *Journal of Anxiety Disorders*, *28*(6), 495–503. doi:https://doi.org/10.1016/j.janxdis.2014.05.005

Kim, K., Kim, C.-H., Cha, K. R., Park, J., Han, K., Kim, Y. K., … Kim, S. I. (2008). Anxiety provocation and measurement using virtual reality in patients with obsessive-compulsive disorder. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *11*(6), 637–641. doi:https://doi.org/10.1089/cpb.2008.0003

Kim, K., Kim, S. I., Cha, K. R., Park, J., Rosenthal, M. Z., Kim, J.-J., … Kim, C.-H. (2010). Development of a computer-based behavioral assessment of checking behavior in obsessive-compulsive disorder. *Comprehensive Psychiatry*, *51*(1), 86–93. doi:https://doi.org/10.1016/j.comppsych.2008.12.001

Kim, K., Roh, D., Kim, C.-H., Cha, K. R., Rosenthal, M. Z., & Kim, S. I. (2012). Comparison of checking behavior in adults with or without checking symptom of obsessive-compulsive disorder using a novel computer-based measure. *Computer Methods and Programs in Biomedicine*, *108*(1), 434–441. doi:https://doi.org/10.1016/j.cmpb.2012.03.014

Morein-Zamir, S., Shahper, S., Camos, J. G., Deruix, A., Worbe, Y., Fineberg, N. A., & Robbins, T. W. (2017). Tolerating uncertainty in OCD, do patients require more information to make decisions? *European Neuropsychopharmacology*, *27*(6), 613–614. doi:https://doi.org/10.1016/j.euroneuro.2016.07.017

O'Connor, K., Wilson, S., Taillon, A., Pélissier, M.-C., & Audet, J.-S. (2018). Inductive reasoning and doubt in obsessive compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *59*, 65–71. doi:https://doi.org/10.1016/j.jbtep.2017.11.002

Pélissier, M.-C., & O'Connor, K. P. (2002). Deductive and inductive reasoning in obsessive-compulsive disorder. *British Journal of Clinical Psychology*, *41*(1), 15–27. doi:https://doi.org/10.1348/014466502163769

Reese, H. E., McNally, R. J., & Wilhelm, S. (2011). Probabilistic reasoning in patients with body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*(3), 270–276. doi:https://doi.org/10.1016/j.jbtep.2010.11.005

Rotge, J. Y., Clair, A. H., Jaafari, N., Hantouche, E. G., Pelissolo, A., Goillandeau, M., … Burbaud, P. (2008). A challenging task for assessment of checking behaviors in obsessive–compulsive disorder. *Acta Psychiatrica Scandinavica*, *117*(6), 465–473. doi:https://doi.org/10.1111/j.1600-0447.2008.01173.x

Rotge, J.-Y., Langbour, N., Dilharreguy, B., Bordessoulles, M., Guehl, D., Bioulac, B., … Burbaud, P. (2015). Contextual and behavioral influences on uncertainty in obsessive-compulsive disorder. *Cortex*, *62*, 1–10. doi:https://doi.org/10.1016/j.cortex.2012.12.010

Toffolo, M. B. J., van den Hout, M. A., Engelhard, I. M., Hooge, I. T. C., & Cath, D. C. (2016). Patients With Obsessive-Compulsive Disorder Check Excessively in Response to Mild Uncertainty. *Behavior Therapy*, *47*(4), 550–559. doi:https://doi.org/10.1016/j.beth.2016.04.002

van Bennekom, M. J., Kasanmoentalib, M. S., de Koning, P. P., & Denys, D. (2017). A Virtual Reality Game to Assess Obsessive-Compulsive Disorder. *Cyberpsychology, Behavior, and Social Networking*, *20*(11), 718–722. doi:https://doi.org/10.1089/cyber.2017.0107

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2019.101807.

References[2]

Anholt, G. E., & Kalanthroff, E. (2013). Letter to the editor: Recent advances in research on cognition and emotion in OCD: A review. *Current Psychiatry Reports, 15*(12), 416. https://doi.org/10.1007/s11920-013-0416-x.

Arntz, A., Voncken, M., & Goosen, A. C. A. (2007). Responsibility and obsessive–compulsive disorder: An experimental test. *Behaviour Research and Therapy, 45*(3), 425–435. https://doi.org/10.1016/j.brat.2006.03.016.

Babineau, J. (2014). Product review: Covidence (systematic review software). *Journal of the Canadian Health Libraries Association / Journal de l'Association Des Bibliothèques de La Santé Du Canada, 35*(2), 68–71. https://doi.org/10.5596/c14-016.

Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., ... 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.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3.

Botta, F., Vibert, N., Harika-Germaneau, G., Frasca, M., Rigalleau, F., Fakra, E., ... Jaafari, N. (2018). Visual search for verbal material in patients with obsessive–compulsive disorder. *Psychiatry Research, 264*, 244–253. https://doi.org/10.1016/j.psychres.2018.03.054.

Bucarelli, B., & Purdon, C. (2016). Stove checking behaviour in people with OCD vs. anxious controls. *Journal of Behavior Therapy and Experimental Psychiatry, 53*, 17–24. https://doi.org/10.1016/j.jbtep.2016.03.005.

Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive–compulsive disorder and trichotillomania. *Neuropsychologia, 45*(4), 654–662. https://doi.org/10.1016/j.neuropsychologia.2006.07.016.

Clair, A.-H., N'Diaye, K., Baroukh, T., Pochon, J.-B., Morgiève, M., Hantouche, E., ... Mallet, L. (2013). Excessive checking for non-anxiogenic stimuli in obsessive-compulsive disorder. *European Psychiatry, 28*(8), 507–513. https://doi.org/10.1016/j.eurpsy.2012.11.003.

Coleman, S. L., Pieterefesa, A. S., Holaway, R. M., Coles, M. E., & Heimberg, R. G. (2011). Content and correlates of checking related to symptoms of obsessive compulsive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders, 25*(2), 293–301. https://doi.org/10.1016/j.janxdis.2010.09.014.

Coles, M. E., & Ravid, A. (2016). Clinical presentation of not-just right experiences (NJREs) in individuals with OCD: Characteristics and response to treatment. *Behaviour Research and Therapy, 87*, 182–187. https://doi.org/10.1016/j.brat.2016.09.013.

Dar, R., Rish, S., Hermesh, H., Taub, M., & Fux, M. (2000). Realism of confidence in obsessive-compulsive checkers. *Journal of Abnormal Psychology, 109*(4), 673. https://doi.org/10.1037/0021-843X.109.4.673.

De Putter, L. M. S., Van Yper, L., & Koster, E. H. W. (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.

Doron, G., Derby, D. S., Szepsenwol, O., & Talmor, D. (2012). Flaws and all: Exploring partner-focused obsessive-compulsive symptoms. *Journal of Obsessive-Compulsive and Related Disorders, 1*(4), 234–243. https://doi.org/10.1016/j.jocrd.2012.05.004.

Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior Modification, 41*(2), 323–339. https://doi.org/10.1177/0145445516673998.

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2016). Psychosis, delusions and the "jumping to conclusions" reasoning bias: A systematic review and meta-analysis. *Schizophrenia Bulletin, 42*(3), 652–665. https://doi.org/10.1093/schbul/sbv150.

Fear, C. F., & Healy, D. (1997). Probabilistic reasoning in obsessive–compulsive and delusional disorders. *Psychological Medicine, 27*(1), 199–208.

Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. ArXiv:1503.02220. Retrieved from http://arxiv.org/abs/1503.02220.

Foa, E. B., Mathews, A., Abramowitz, J. S., Amir, N., Przeworski, A., Riggs, D. S., ... Alley, A. (2003). Do patients with obsessive–compulsive disorder have deficits in decision-making? *Cognitive Therapy and Research, 27*(4), 431–445. https://doi.org/10.1023/A:1025424530644.

Gentes, E. L., & Ruscio, A. M. (2011). A meta-analysis of the relation of intolerance of uncertainty to symptoms of generalized anxiety disorder, major depressive disorder, and obsessive–compulsive disorder. *Clinical Psychology Review, 31*(6), 923–933. https://doi.org/10.1016/j.cpr.2011.05.001.

Grassi, G., Pallanti, S., Righi, L., Figee, M., Mantione, M., Denys, D., ... Stratta, P. (2015). Think twice: Impulsivity and decision-making in obsessive–compulsive disorder. *Journal of Behavioral Addictions, 4*(4), 263–272. https://doi.org/10.1556/2006.4.2015.039.

Grayson, J. (2014). *Freedom from obsessive-compulsive disorder: A personalized recovery program for living with uncertainty.* New York, NY: Berkley Publishing Group.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5.

Hermans, D., Engelen, U., Grouwels, L., Joos, E., Lemmens, J., & Pieters, G. (2008). Cognitive confidence in obsessive-compulsive disorder: Distrusting perception,

[2] References of the studies included in the meta-analyses are listed in Appendix A.

attention and memory. *Behaviour Research and Therapy, 46*(1), 98–113. https://doi.org/10.1016/j.brat.2007.11.001.

Hezel, D. M., Stewart, S. E., Riemann, B. C., & McNally, R. J. (2019). Standard of proof and intolerance of uncertainty in obsessive-compulsive disorder and social anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 64*, 36–44. https://doi.org/10.1016/j.jbtep.2019.02.002.

van Bennekom, M. J., Kasanmoentalib, M. S., de Koning, P. P., & Denys, D. (2017). A Virtual Reality Game to Assess Obsessive-Compulsive Disorder. *Cyberpsychology, Behavior, and Social Networking, 20*(11), 718–722. https://doi.org/10.1089/cyber.2017.0107.

van den Hout, M., & Kindt, M. (2003). Repeated checking causes memory distrust. *Behaviour Research and Therapy, 41*(3), 301–316. https://doi.org/10.1016/S0005-7967(02)00012-8.

van den Hout, M. A., Engelhard, I. M., de Boer, C., du Bois, A., & Dek, E. (2008). Perseverative and compulsive-like staring causes uncertainty about perception. *Behaviour Research and Therapy, 46*(12), 1300–1304. https://doi.org/10.1016/j.brat.2008.09.002.

van den Hout, M. A., van Dis, E. A. M., van Woudenberg, C., & van de Groep, I. H. (2017). OCD-like checking in the lab: A meta-analysis and improvement of an experimental paradigm. *Journal of Obsessive-Compulsive and Related Disorders.*. https://doi.org/10.1016/j.jocrd.2017.11.006.

Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology Section A, 40*(4), 801–812. https://doi.org/10.1080/14640748808402300.

Ishak, K. J., Platt, R. W., Joseph, L., & Hanley, J. A. (2008). Impact of approximating or ignoring within-study covariances in multivariate meta-analyses. *Statistics in Medicine, 27*(5), 670–686. https://doi.org/10.1002/sim.2913.

Jaafari, N., Frasca, M., Rigalleau, F., Rachid, F., Gil, R., Olié, J.-P., ... Vibert, N. (2013). Forgetting what you have checked: A link between working memory impairment and checking behaviors in obsessive-compulsive disorder. *European Psychiatry, 28*(2), 87–93. https://doi.org/10.1016/j.eurpsy.2011.07.001.

Jacobsen, P., Freeman, D., & Salkovskis, P. (2012). Reasoning bias and belief conviction in obsessive-compulsive disorder and delusions: Jumping to conclusions across disorders? *British Journal of Clinical Psychology, 51*(1), 84–99. https://doi.org/10.1111/j.2044-8260.2011.02014.x.

Jacoby, R. J., Abramowitz, J. S., Buck, B. E., & Fabricant, L. E. (2014). How is the beads task related to intolerance of uncertainty in anxiety disorders? *Journal of Anxiety Disorders, 28*(6), 495–503. https://doi.org/10.1016/j.janxdis.2014.05.005.

Kalanthroff, E., Anholt, G. E., & Henik, A. (2014). Always on guard: Test of high vs. low control conditions in obsessive-compulsive disorder patients. *Psychiatry Research, 219*(2), 322–328. https://doi.org/10.1016/j.psychres.2014.05.050.

Kim, K., Kim, C.-H., Cha, K. R., Park, J., Han, K., Kim, Y. K., ... Kim, S. I. (2008). Anxiety provocation and measurement using virtual reality in patients with obsessive-compulsive disorder. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 11*(6), 637–641. https://doi.org/10.1089/cpb.2008.0003.

Kim, K., Kim, S. I., Cha, K. R., Park, J., Rosenthal, M. Z., Kim, J.-J., ... Kim, C.-H. (2010). Development of a computer-based behavioral assessment of checking behavior in obsessive-compulsive disorder. *Comprehensive Psychiatry, 51*(1), 86–93. https://doi.org/10.1016/j.comppsych.2008.12.001.

Kim, K., Roh, D., Kim, C.-H., Cha, K. R., Rosenthal, M. Z., & Kim, S. I. (2012). Comparison of checking behavior in adults with or without checking symptom of obsessive-compulsive disorder using a novel computer-based measure. *Computer Methods and Programs in Biomedicine, 108*(1), 434–441. https://doi.org/10.1016/j.cmpb.2012.03.014.

Lazarov, A., Cohen, T., Liberman, N., & Dar, R. (2015). Can doubt attenuate access to internal states? Implications for obsessive-compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 49*, 150–156. https://doi.org/10.1016/j.jbtep.2014.11.003.

Lazarov, A., Dar, R., Oded, Y., & Liberman, N. (2010). Are obsessive–compulsive tendencies related to reliance on external proxies for internal states? Evidence from biofeedback-aided relaxation studies. *Behaviour Research and Therapy, 48*(6), 516–523. https://doi.org/10.1016/j.brat.2010.02.007.

Lazarov, A., Liberman, N., Hermesh, H., & Dar, R. (2014). Seeking proxies for internal states in obsessive–compulsive disorder. *Journal of Abnormal Psychology, 123*(4), 695–704. https://doi.org/10.1037/abn0000004.

Lorenzo-Luaces, L., Keefe, J. R., & DeRubeis, R. J. (2016). Cognitive-behavioral therapy: Nature and relation to non-cognitive behavioral therapy. *Behavior Therapy, 47*(6), 785–803. https://doi.org/10.1016/j.beth.2016.02.012.

Milner, A. D., Beech, H. R., & Walker, V. J. (1971). Decision processes and obsessional behaviour. *British Journal of Clinical Psychology, 10*(1), 88–89.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.

Morein-Zamir, S., Shahper, S., Camos, J. G., Deruix, A., Worbe, Y., Fineberg, N. A., & Robbins, T. W. (2017). Tolerating uncertainty in OCD, do patients require more information to make decisions? *European Neuropsychopharmacology, 27*(6), 613–614. https://doi.org/10.1016/j.euroneuro.2016.07.017.

Nedeljkovic, M., & Kyrios, M. (2007). Confidence in memory and other cognitive processes in obsessive–compulsive disorder. *Behaviour Research and Therapy, 45*(12), 2899–2914. https://doi.org/10.1016/j.brat.2007.08.001.

O'Connor, K., Aardema, F., & Pélissier, M.-C. (2005). *Beyond reasonable doubt: Reasoning processes in obsessive-compulsive disorder and related disorders.* Chichester: John Wiley & Sons, Ltd.

Ouellet-Courtois, C., Wilson, S., & O'Connor, K. (2018). Cognitive confidence in obsessive-compulsive disorder: A systematic review and meta-analysis. *Journal of Obsessive-Compulsive and Related Disorders, 19*, 77–86. https://doi.org/10.1016/j.jocrd.2018.08.003.

Pélissier, M.-C., & O'Connor, K. P. (2002). Deductive and inductive reasoning in obsessive-compulsive disorder. *British Journal of Clinical Psychology, 41*(1), 15–27. https://doi.org/10.1348/014466502163769.

Rachman, S. (2002). A cognitive theory of compulsive checking. *Behaviour Research and Therapy, 40*(6), 625–639. https://doi.org/10.1016/S0005-7967(01)00028-6.

Reese, H. E., McNally, R. J., & Wilhelm, S. (2011). Probabilistic reasoning in patients with body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(3), 270–276. https://doi.org/10.1016/j.jbtep.2010.11.005.

Ron, O., Oren, E., & Dar, R. (2016). The doubt-certainty continuum in psychopathology, lay thinking, and science. *Journal of Behavior Therapy and Experimental Psychiatry, 53*, 68–74. https://doi.org/10.1016/j.jbtep.2015.08.005.

Rotge, J. Y., Clair, A. H., Jaafari, N., Hantouche, E. G., Pelissolo, A., Goillandeau, M., ... Burbaud, P. (2008). A challenging task for assessment of checking behaviors in obsessive–compulsive disorder. *Acta Psychiatrica Scandinavica, 117*(6), 465–473. https://doi.org/10.1111/j.1600-0447.2008.01173.x.

Rotge, J.-Y., Langbour, N., Dilharreguy, B., Bordessoulles, M., Guehl, D., Bioulac, B., ... Burbaud, P. (2015). Contextual and behavioral influences on uncertainty in obsessive-compulsive disorder. *Cortex, 62*, 1–10. https://doi.org/10.1016/j.cortex.2012.12.010.

Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry, 15*(1), 53–63. https://doi.org/10.1038/mp.2008.94.

Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy, 23*(5), 571–583. https://doi.org/10.1016/0005-7967(85)90105-6.

Salkovskis, P. M. (1998). Psychological approaches to the understanding of obsessional problems. In R. P. Swinson, M. M. Antony, S. Rachman, & M. A. Richter (Eds.). *Obsessive Compulsive Disorder – Theory, Research & Treatment.* New York: Guilford Press.

So, S. H., Siu, N. Y., Wong, H., Chan, W., & Garety, P. A. (2016). "Jumping to conclusions" data-gathering bias in psychosis and other psychiatric disorders — Two meta-analyses of comparisons between patients and healthy individuals. *Clinical Psychology Review, 46*, 151–167. https://doi.org/10.1016/j.cpr.2016.05.001.

Soref, A., Liberman, N., Abramovitch, A., & Dar, R. (2018). Explicit instructions facilitate performance of OCD participants but impair performance of non-OCD participants on a serial reaction time task. *Journal of Anxiety Disorders, 55*, 56–62. https://doi.org/10.1016/j.janxdis.2018.02.003.

Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research synthesis methods, 4*(2), 169–187. https://doi.org/10.1002/jrsm.1070.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393. https://doi.org/10.1037/met0000011.

Toffolo, M. B. J., van den Hout, M. A., Engelhard, I. M., Hooge, I. T. C., & Cath, D. C. (2016). Patients with obsessive-compulsive disorder check excessively in response to mild uncertainty. *Behavior Therapy, 47*(4), 550–559. https://doi.org/10.1016/j.beth.2016.04.002.

Toffolo, M. B. J., van den Hout, M. A., 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.

Volans, P. J. (1976). Styles of decision-making and probability appraisal in selected obsessional and phobic patients. *British Journal of Social and Clinical Psychology, 15*(3), 305–317. https://doi.org/10.1111/j.2044-8260.1976.tb00038.x.

Williams, J. W., Plassman, B. L., Burke, J., Holsinger, T., & Benjamin, S. (2010). Preventing Alzheimer's disease and cognitive decline. *Evidence Report/Technology Assessment No. 193. (Prepared by the Duke evidence-based practice center under contract No. HHSA 290–2007-10066-I)*Rockville, MD: Agency for Healthcare Research and Qualityhttps://www.choiceforum.org/docs/prevent.pdf (accessed Oct 02, 2019).