How checking breeds doubt: Reduced performance in a simple working memory task

Ben Harkin and Klaus Kessler

Introduction

Checking is the most common symptom exhibited in obsessive-compulsive disorder (OCD; Ball, Baer, & Otto, 1996) with 81% of patients having checking compulsions (Antony, Downue, & Swinson, 1998). Ritualistic checking appears to be fuelled by a lack of confidence that checkers have in their memory, thus they check and re-check to increase their certainty, yet with effect of actually decreasing memory accuracy (Rachman & Shafran, 1998). Rachman (2002) proposed that any cognitive theory of pathological checking must account for the observed memory problems. In its simplest interpretation this could suggest a general memory deficit as the underlying cause. However, since empirical findings are inconsistent regarding such a general mnestic deficit (e.g. Tallis, 1997 vs. MacDonald et al., 1997 and McNally and Kohlbeck, 1993) it seems likely that memory uncertainty is rather domain- and episode-specific (e.g., “Did I turn my stove off this morning?”) (van den Hout and Kindt, 2003a and van den Hout and Kindt, 2003b).

Nelson and Narens (1994) proposed that the reactivation of memory traces can never provide an entirely veridical representation of the original input. That is, our memories will never be as vivid and “true” as the original experience. In certain domains compulsive checkers might be overly aware of these natural short-falls of memory traces, strongly experiencing doubt and starting to check the same memory trace over and over again, yet, without the possibility to enhance certainty. Accordingly, a paradox of memory research is that repeated checking of memory traces actually serves to further undermine memory certainty, memory vividness and confidence (e.g., Coles et al., 2006, van den Hout and Kindt, 2003a, van den Hout and Kindt, 2003b, van den Hout and Kindt, 2004 and Radomsky et al., 2006). For example, van den Hout and Kindt (2003b) asked participants to repeatedly turn on, off or check a computer simulation of a six-burner gas stove for 20 trials after which they were asked to report the vividness, detail and memory confidence for their last check of the stove. In the checking, compared to the control-condition checkers had a significant decrease in the three aforementioned metacognitive measures and the authors concluded that checking breeds doubt, not certainty. van den Hout and Kindt (2004) proposed that repetitive checking only increases the familiarity of the checked stimuli resulting in a shift from detailed episodic recollection to vague familiarity judgments along with reduced memory confidence (van den Hout & Kindt, 2003a).

While it appears that episodic long-term memory is mostly affected by checking, the question as to the stage at which checking interferes with the memory traces remains unresolved (Coles et al., 2006). Is the effect purely confined to episodic long-term memory or does it operate already at an earlier stage? Checking may already interfere with episodic representations in short-term working-memory (WM) and affect their transfer into long-term-memory (LTM). “Did I turn ALL the burners off?” – such a question could already arise seconds after leaving the kitchen and could strongly affect how we remember the state of the stove hours later. Although Shimamura (2000) suggested that WM representations are likely targets for compulsive checking, the relationship between WM and checking is poorly understood (e.g. Henseler et al., 2008). Here we propose that compulsive checking in the sense of repeatedly questioning the veridicality of memory representations (cf. Nelson & Narens, 1994) will reduce performance already at the stage of WM.

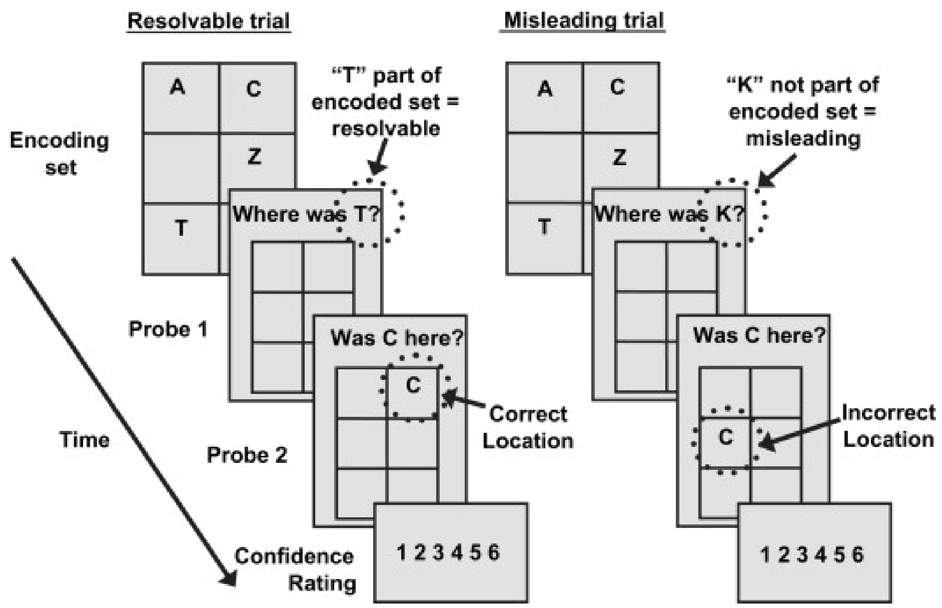
These reflections also implicate the “active” component of WM according to Baddeley's (1986) classic model, namely the “Central Executive”, since task-related but irrelevant thoughts (i.e., intrusive thoughts; Salkovskis, 1998) that question memory veridicality cannot be easily suppressed by top-down executive control in OCD. It has been reported that the primary neuropsychological profile of OCD/checking is one of executive dysfunction (Olley, Malhi, & Sachdev, 2007) and patients reveal deficits in relation to inhibition, cognitive flexibility, and multi-tasking (Olley et al., 2007, Omori et al., 2007 and Savage et al., 2000). Intrusive thoughts that doubt the veridicality of memory traces (e.g., “Did I REALLY turn all the burners off?”) could therefore be more detrimental for compulsive checkers because they cannot easily prevent these thoughts from affecting ongoing conscious processing (cf. Bannon et al., 2008 and Salkovskis, 1998).

Here we claim that Baddeley's (2000) WM model provides a suitable framework for how intrusive thoughts could already undermine short-term representations prior to consolidation into episodic LTM. Specifically we suggest that detrimental checking fuelled by unsuppressed intrusive thoughts might strongly interfere with representations in the so-called episodic buffer (EB) of WM (Baddeley, 2000). To briefly re-iterate, Baddeley's (2000) extension of his classic WM model (1986) introduced the EB to allow for multimodal, temporarily integrated representations and to serve as an interface to episodic LTM (also compare Fig. 7). Baddeley (2003) later emphasized the parallels between his EB and the concept of a “global workspace” (Baars, 2002 and Dehaene and Naccache, 2001), which is a formal neuro-cognitive approach to conscious/aware processing. In short, this embraces the notion that compulsive checking affects the current stream of consciousness (e.g. Salkovskis, 1998).

In the “stove” example, the EB explains how we can recall the states of the burners by allowing for multimodal bindings between visual, spatial, and temporal information over short periods of time. The EB is also thought to interface with episodic LTM providing the memory traces that become consolidated for longer periods of time. Accordingly, if EB functionality was disrupted by detrimental checking, then this would extend and refine the explanation of memory symptoms of compulsive checking in general. Such a finding would help clarify the etiological starting point for pathological checking behaviour (e.g. Coles et al., 2006).

The present studies set out to (1) engage the EB by using stimuli which required multimodal conjunctions between phonological (letters) and spatial (locations) features and (2) hamper EB functionality by inducing checking during the delay/maintenance period of the WM task. We used a delayed-match-to-sample WM task (a variation of the Sternberg (1969) paradigm), in which we induced checking during WM maintenance (see Fig. 1). The participants' primary task was to encode a set of 4 letters and their locations into working memory and indicate seconds later if a test letter was presented in the correct location. This primary task was intended to be easy and average performance was over 90%.

Fig. 1. - Schematic procedure of resolvable and misleading trials. A set of 4 letters presented randomly in 4 out of 6 possible locations had to be encoded within 2 s. Encoding was then followed after 500 ms by a first probe letter (probe-1) which was either part or not part of the encoded set, i.e., was resolvable or misleading. Subsequently participants had to indicate if the probe-2 letter was correctly or incorrectly located with respect to the encoded set, which was the actual memory test. Finally confidence in the probe-2 response had to be indicated on a scale from 1 (highly certain) to 6 (highly uncertain). Further explanations in the text.



The checking manipulation between encoding and the actual memory test was induced by presenting an intermediate, potentially misleading question (probe-1). Participants were asked explicitly where a specific letter had been, while this letter, e.g. “T” or “K” (see Fig. 1), either was or was not part of the encoded set. For the latter case the question of where the letter had been was not solvable and we expected that this misleading information would induce repeated checking of the veridicality of the encoded representations especially in high checkers. This, we argue, is analogous to the process of having just completed a task (e.g., turning off the stove) and then almost immediately starting to check the maintained WM representations for their veridicality. Such a check is likely to be driven by a thought or an external stimulus that is task-related but irrelevant to the successful recall of the memory trace (e.g., thought: “I think I left it on”; external cue: “Did you turn the oven off?”). Compulsive checkers have been reported to show a deficit in inhibiting intrusive thoughts and distracting stimuli (e.g., Olley et al., 2007, Omori et al., 2007 and Savage et al., 2000). We therefore expected that the presence of a misleading but irrelevant probe-1 question would especially interfere with WM representations of higher checkers.

In general we expected that with a misleading probe-1, participants (checkers in particular) would sequentially and repeatedly compare the visually presented probe letter to the memorised set, yet, without success. At the representational level this would lead to a competition between a strong visual stimulus and weaker, memorised letter-location bindings. The stronger this competition is (=lack of suppression of misleading information) and the more often this competition is repeated (=checking) the more strongly the originally encoded memories are weakened – ultimately resulting in a performance deficit on the actual memory test (probe-2).

In particular we expected that individuals with a high pre-disposition for checking (Vancouver Obsessional Compulsive Inventory (VOCI); Thordarson et al., 2004) would be more susceptible to our misleading-probe manipulation than “non-checkers”. We chose an easy primary task (4 letters in 6 locations) to avoid group differences due to differences in WM capacity at high load (see Van der Wee et al., 2003) and we included a control condition without intermediate probe (no probe-1) to obtain a baseline indication of capacity. We also employed two different blocks of trials, one that consisted predominantly (2/3) of resolvable probe-1 trials and one consisting of 2/3 misleading trials. This was a manipulation to investigate expectations regarding specific trial types and their potentially different impact on high and low checkers. Low checkers might more readily learn to ignore any misleading probe-1, if these trials are very frequent (mostly misleading block), while high checkers might not be able to suppress intrusive thoughts triggered by the misleading probe even if the context of the experiment (mostly misleading block) actually emphasizes their irrelevance.

Experiment 1

Method

Participants

40 Participants (mean 22.7 years: 12 males, 26 females) from the University of Glasgow gave written informed consents. British Psychological Society ethical requirements were met, including that of participant debriefing. The Vancouver Obsessional Compulsive Inventory, VOCI (Thordarson et al., 2004) was employed to evaluate all participants regarding their checking tendencies. The VOCI is a 55 item, self-report questionnaire for assessing the severity of OCD symptoms. The checking subscale was used in the present study. A median split of checking scores was used to obtain two groups: low (mean = 1.11, SD = 1.10) and high (mean = 9.53, SD = 5.49) “checkers”.

Stimuli and procedure

Participants sat 90 cm from a 19″ computer screen ran at 800 × 600 resolution with their head on a chin rest. Stimuli were capital letters in font Arial, size 18 and were presented against a black background within a 2 (columns) by 3 (rows) matrix covering an area of 300 × 420 pixels. After 1000 ms fixation, 4 letters were presented randomly in 4 of the 6 possible locations and participants had 2000 ms to encode the identity and the location of each letter (Fig. 1). After 500 ms, the probe-1 question requested the location of a specific letter. Participants indicated the location through a 2 × 3 spatially mapped keypad and were instructed to respond within 4000 ms (to keep the WM delay constant). Whether the probe-1 letter had or had not been part of the encoded set created the resolvable vs. misleading (irresolvable) trials. In a baseline condition probe-1 was omitted to measure WM performance on the primary task under ideal conditions.

A 1000 ms interstimulus interval (ISI) separated probe-1 and probe-2. Since baseline trials did not include the intermediate probe-1 a black screen was shown for 5500 ms between encoding and probe-2 (equaling the ISI between encoding and probe-2 on the other trial types). Probe-2 was the actual memory test for each trial and required participants to indicate if a letter was correctly located with respect to the originally encoded set. In all trials the probe-2 letter had been part of the encoded set in terms of identity while the probe location was correct only on 50% of the trials. Finally, a scale was displayed prompting participants to indicate their degree of confidence in their probe-2 response (6 levels: 1 = totally certain to 6 = totally uncertain). Three self-paced breaks were included and the experiment lasted approximately 90 min.

Design

A two (group: low vs. high checkers) by two (block type: mostly resolvable vs. mostly misleading block) by three (probe-1 trial type: resolvable, misleading, no probe-1) mixed design was employed with group as the between- and block and probe-1 as the within-subjects factors. The resolvable block comprised 180 trials with 120 resolvable trials, 40 misleading trials, and additional 20 baseline trials (no probe-1). Correspondingly, the misleading block (180 trials) was made up of 40 resolvable trials, 120 misleading trials, and again 20 baseline trials. The sequence of these blocks was counterbalanced across participants in order to avoid order effects.

Results

MANOVAs for a 2 × 2 × 3 design were carried out for reaction times, accuracy and confidence on probe-2 responses due to violations of the sphericity assumption (Mauchley's tests). As our theoretical predictions focused on the effect of checking induced by resolvable vs. misleading probe-1 trials we also conducted 2 × 2 × 2 ANOVAs removing the no-probe-1 trials. The datasets of two participants were not used in further analysis as accuracy was at chance levels in at least one condition. All other participants performed well above chance level in all conditions (>70% accuracy).

Probe-2 response latencies (RT)

The MANOVA (2 × 2 × 3) for probe-2 latencies revealed a main effect of trial type ($F(2,72) = 10.65$, $p < 0.001$, $\eta 2p = 0.378$) and the ANOVA for the reduced 2 × 2 × 2 design (without no-probe-1 trials) also revealed a main effect of trial type ($F(1,36) = 9.46$, $p < 0.004$, $\eta 2p = 0.208$). This indicates that the misleading trials were slower than the resolvable trials. The easy but most infrequent no-probe-1 trials were slowest (Table 1).

Table 1. - Reaction times for all participants for resolvable, misleading and no probe-1 trial types.

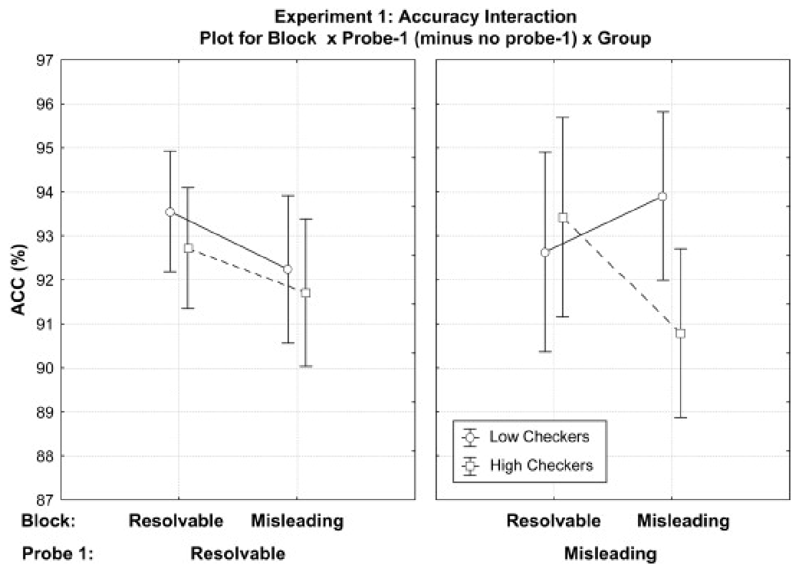
Trial type	Resolvable	Misleading	No probe-1
RT (ms)	1782.021	1896.743	1982.715

Probe-2 response accuracy (AC)

The MANOVA (2 × 2 × 3) for probe-2 accuracy revealed a main effect of block ($F(1,36) = 5.64$, $p < 0.03$, $\eta^2_p = 0.135$) which was indicative of generally less accuracy in the misleading block than the resolvable block. The main effect of trial type was also significant ($F(2,35) = 3.53$, $p < 0.04$, $\eta^2_p = 0.168$) and indicated the greater accuracy in the no-probe-1 trials relative to the resolvable and misleading trials. Importantly, the simple effect for no-probe-1 trials (baseline) revealed no significant difference between high- and low-scorers ($F(1,36)$, $p < 0.5$, $p > 0.48$) indicating that WM capacity was comparable between groups.

Next, we removed the no-probe-1 trials to focus on the more relevant resolvable and misleading trials. The ANOVA for the reduced 2 × 2 × 2 design revealed a significant 3-way interaction (see Fig. 2) for group × block × trial type ($F(1,36) = 4.35$, $p < 0.05$, $\eta^2_p = 0.108$). To clarify which conditions generate this complex interaction we split the analysis into two more simple 2-way ANOVAs of group × block, for resolvable and misleading trials separately (see left and right plots in Fig. 2). Only the interaction for misleading probe-1 trials (right plot in Fig. 2) reached significance ($F(1,36) = 5.98$, $p < 0.02$, $\eta^2_p = 0.142$), suggesting that accuracy (on the subsequent probe-2) for misleading trials differed significantly between blocks and between checking groups. Most interestingly, this difference appears to exist only within the misleading block (low checkers = 93.90 vs. high = 90.79) supporting our hypotheses that high checkers are less able to learn from the experimental context even if it emphasizes the irrelevance of the misleading probe-1s.

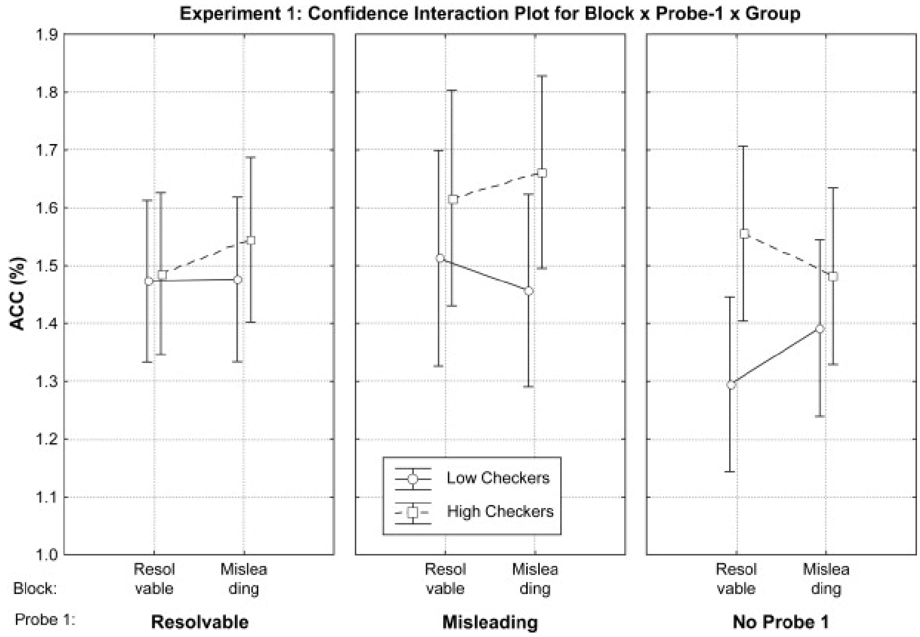
Fig. 2. - Experiment 1: group (low vs. high checkers) × block (resolvable vs. misleading) × trial type (resolvable vs. misleading) interaction plot for probe-2 accuracy. Vertical bars denote standard errors.



Probe-2 response confidence ratings (CR)

The MANOVA (2 × 2 × 3) for confidence in probe-2 responses revealed a significant main effect of trial type ($F(2,35) = 8.67$, $p < 0.001$, $\eta^2_p = 0.331$) with misleading probe-1s inducing the least confidence in subsequent probe-2 responses. In addition, however, there was a significant 3-way interaction of group × trial type × block ($F(2,35) = 4.16$, $p < 0.03$, $\eta^2_p = 0.192$). Fig. 3 shows that high checkers indeed show decreased confidence compared to low checkers, but that this difference is not consistent across trial types and blocks, i.e., a quite similar pattern is observed for the two groups for resolvable trials in both blocks (left in Fig. 3) while a more dissimilar pattern is revealed for misleading and no probe-1 trials (middle and right in Fig. 3).

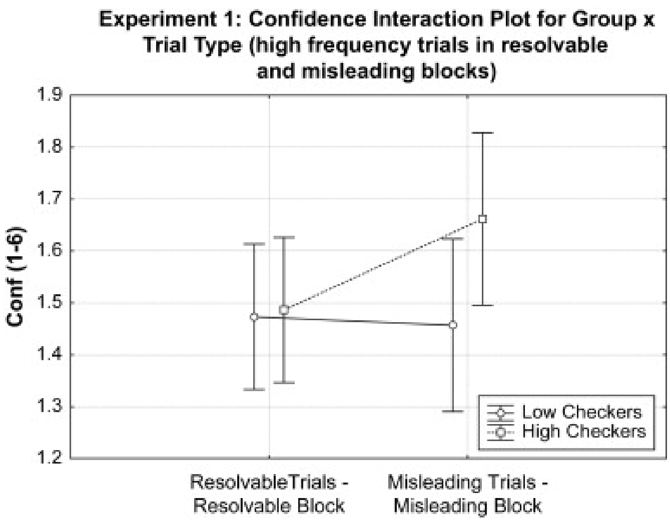
Fig. 3. - Experiment 1: group (low vs. high checkers) × block (resolvable vs. misleading) × trial type (resolvable, misleading, or no P1) interaction plot for confidence in probe-2 responses. The scale ranged from 1 (highly certain) to 6 (highly uncertain), i.e., lower values reflect higher confidence. Vertical bars denote standard errors.



The ANOVA for the reduced 2 × 2 × 2 design (without no-probe-1 trials) failed to reveal any significant results, suggesting that the MANOVA results were substantially driven by the difference between misleading and no-probe-1 trials (Fig. 3 middle and right graphs).

Finally we directly compared confidence between resolvable and misleading trials for equal frequencies to further understand the role of the misleading trials in the significant 3way interaction of the MANOVA. That is, we compared the 120 resolvable trials of the resolvable block to the 120 misleading trials of the misleading block. “Group” was included as a second factor. This 2 × 2 ANOVA revealed a significant group × trial type interaction ($F(1,36) = 8.56$, $p < 0.006$, $\eta 2p = 0.192$) that is shown in Fig. 4. This interaction further substantiates the difference in confidence ratings between groups observed for misleading trials within the misleading block (middle graph Fig. 3), where high checkers had less confidence than low checkers.

Fig. 4. - Experiment 1: group (low vs. high checkers) × trial type (resolvable trials in the resolvable block vs. misleading trials in the misleading block) interaction plot for confidence in probe-2 responses. To re-iterate, lower values reflect higher confidence. Vertical bars denote standard errors.



Discussion of Experiment 1

We found evidence in reaction time (RT) and confidence ratings (CR) data, suggesting that our manipulation was successful in inducing checking, although the WM task was very easy (all conditions revealed mean accuracies over 90%). This effect was most evident in RTs, where across blocks and groups probe-2 responses were performed faster after a resolvable probe-1 than after a misleading probe-1. This effect was supported by the confidence ratings, where misleading trials led to lower confidence than resolvable and no-probe-1 trials. Furthermore, high checkers had less confidence than low checkers on misleading trials compared to resolvable trials when the respective trial types where blocked together. Finally and most importantly we observed group differences in performance accuracy for misleading trials within the mostly misleading block as reflected in a significant interaction of group, trial type and block. This suggests that checkers cannot easily ignore a misleading cue even if the experimental context emphasizes the irrelevance of the cue. Since this is the result with the potentially strongest impact on our understanding of compulsive checking we wanted to ensure its reliability. In a replication study we presented the misleading block only and focused on the group differences for the misleading trials.

Experiment 2

Methods

Participants

40 Volunteers (mean age 23.88: 14 males, 25 females) participated in this second study and a median split of the VOCI scores was used again to obtain a group of high (mean = 8.65, SD = 3.70) and a group of low (mean = 1.05, SD = 1.18) scorers on the checking scale.

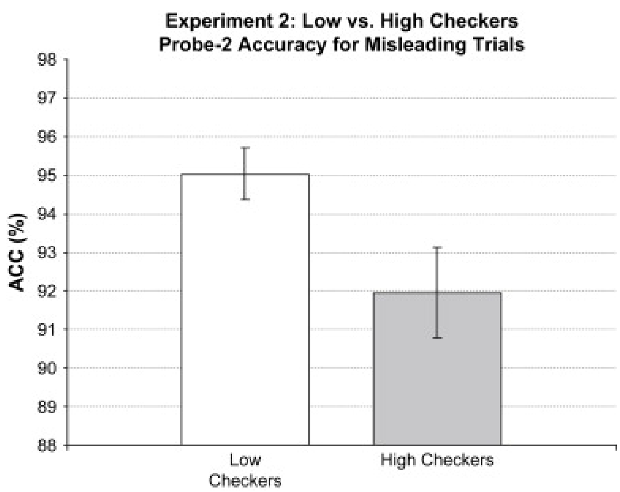
Stimuli and procedure

The same stimuli and procedure as in Experiment 1 were employed. The only change was that only the misleading block was presented (2/3 misleading trials).

Results and discussion

In order to test for a replication of the main finding of Experiment 1 we carried out hypothesis driven t-tests to compare probe-2 accuracy for high and low checkers. We expected high checkers to show again a lower performance for misleading probe-1 trials, which was supported by a significant t-test ($t(1,37) = 2.276$, $p < 0.029$) (Fig. 5). The t-tests for the resolvable and no-probe-1 trials did not reach significance (both $t(1,37) < 0.35$, $p > 0.56$), supporting the notion that there were no general differences in WM capacity. With respect to confidence ratings (CR) numerically observed group differences did not reach significance (resolvable, misleading, no-probe-1: $p > 0.28$).

Fig. 5. - Experiment 2: group comparison (low vs. high checkers) for probe-2 accuracy on misleading trials only. Vertical bars denote standard errors.

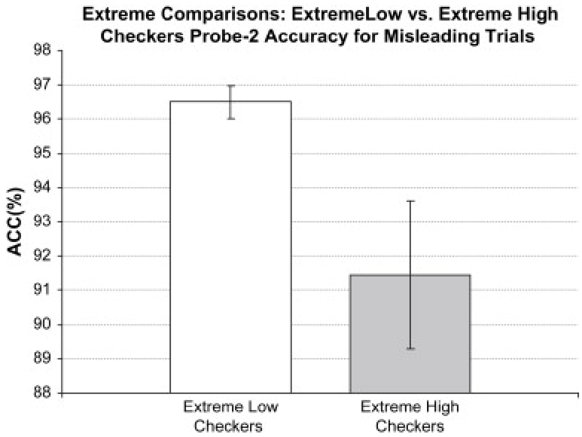


The accuracy outcome is a clear replication of the main finding in Experiment 1, allowing for a convincing conclusion that higher checking disposition is related to attenuated performance within the episodic part of working memory if misleading information is provided. However, in both experiments we used a median split to create the two checking groups (low vs. high). As a result, the high checking group in both experiments (Exp 1 = 9.53; Exp 2 = 8.65) scored below the clinical mean (15.6) of compulsive checkers on the checking subscale of the VOCI (Thordarson et al., 2004) making any conclusions only tentative with respect to the clinical population. Therefore, we conducted a “meta-comparison” where we compared the extremely high checkers (mean score 15.8) to the lowest scorers (0.5) across both experiments.

Extreme groups meta-comparison

We compared the data of the 10 participants with the highest scores on the checking subscale of the VOCI (mean = 15.8, SD = 2.57) to the data of the 10 lowest scoring participants (mean = 0.5, SD = 0.71) from both experiments. Only the data from the misleading block were employed for participants drawn from Experiment 1 (n = 6) to keep the data congruent to Experiment 2. Like in Experiment 2, hypothesis driven t-tests were conducted. We expected that extreme high checkers would differ from extreme low checkers only for misleading trials (see Fig. 6), a hypothesis supported by a significant t-test ($t(1,18) = 2.289$, $p < 0.034$). The t-tests for the resolvable ($t(1,18) = 0.141$, $p > 0.175$) and no-probe-1 ($t(1,18) = 0.33$, $p > 0.745$) trials did not reach significance, supporting the notion that the two extreme groups were comparable with respect to general WM capacity. Again, numerically observed group differences for confidence ratings did not reach significance (resolvable, misleading, no-probe-1: $p > 0.39$).

Fig. 6. - Extreme scorers meta-analysis: group comparison (extreme low vs. extreme high checkers) for probe-2 accuracy on misleading trials only. Vertical bars denote standard errors.



Discussion

The aim of the present study was to show that internal WM representations, i.e., cross-modal bindings within the EB, can be affected by unproductive checking. Deteriorated WM performance was expected to be more pronounced in participants with a high checking predisposition. We found evidence in reaction time (RT) and confidence ratings (CR) data of Experiment 1 suggesting that our manipulation was successful in inducing a certain amount of detrimental checking in all participants. This effect was most evident in the RTs of Experiment 1 (Table 1), where across groups responses on the actual memory test (probe-2) were performed faster after a resolvable probe-1 than after a misleading probe-1. This effect was supported by the confidence ratings, where misleading trials led to lower confidence than resolvable and no-probe-1 trials. Regarding our group hypothesis Experiment 1 revealed less confidence for high checkers than low checkers in misleading trials but not in resolvable trials (Fig. 4). However, in Experiment 2 and in the extreme groups meta-comparison only numerical differences were observed, possibly suggesting that the task was too easy to affect metacognitive judgments in a straightforward way. That is, group differences were only revealed as part of quite complex 3way and 2way interactions in Experiment 1, which was not possible with the reduced design in Experiment 2 and the meta-comparison.

Importantly and in agreement with previous findings (e.g., Ciesielski et al., 2007 and Henseler et al., 2008) we did not observe general differences in WM capacity between high and low checkers: performance on resolvable and no probe-1 trials was comparable in both experiments as well as in the extreme groups meta-comparison. This suggests that group differences are not a WM capacity issue per se – especially with low demands employed here and in other research (e.g. Van der Wee et al., 2003). According to our group hypothesis, we observed differences in performance accuracy in Experiment 1 when irrelevant but misleading probes were maximally concentrated: for misleading trials when these trials where highly frequent (misleading block). This crucial finding was replicated in Experiment 2 and was also revealed in the extreme group comparison across experiments underpinning the clinical relevance of our findings.

We conclude that our experimental manipulation resonated with personal checking dispositions and affected WM representations. It appears that low checkers have learned more readily to ignore irresolvable probes especially in an experimental context where such probes were highly frequent so their irrelevance became even more obvious (misleading block). In contrast, high checkers might have "checked yet another time" whether the probe letter "really" wasn't anywhere. That is, high scorers appeared to be less able to suppress the misleading probes and the associated intrusive thoughts. In turn, this might have initiated repeated scans through WM to compare the irresolvable probe with each letter-location binding over and over again. We propose that the competition between a strong, visually present letter-stimulus and the fragile letter-location bindings in the EB weakens these multimodal representations. Repeated checking due to insufficient suppression of misleading information might have therefore resulted in repeated competition and increasingly weaker bindings.

Fig. 7 shows an adaptation of Baddeley's (2000) model that condensates the interpretation of our findings into a general explanatory framework for compulsive checking. According to this framework compulsive checking could involve three components that together make up a vicious circle.

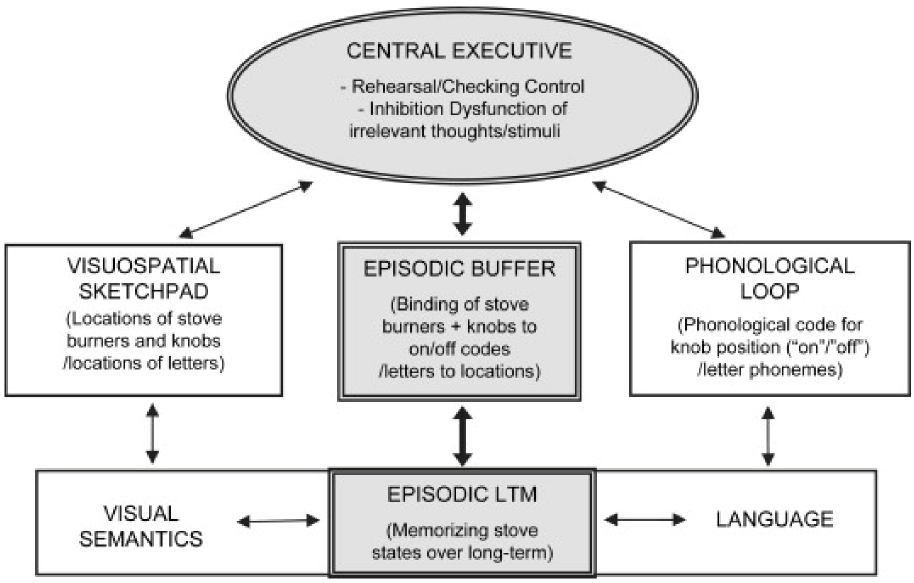
1. Executive dysfunction could result in a lack of suppression of misleading information, which is in strong agreement with the susceptibility to intrusive thoughts (Bannon et al., 2008 and Salkovskis, 1998) and the general executive dysfunction (Olley et al., 2007 and Omori et al., 2007) reported in clinical OCD samples. The misleading/intrusive information can be internally generated in form of intrusive thoughts ("I think I left a burner on") or can be externally provided in form of challenging questions ("Where was the letter?" or "Are you ABSOLUTELY SURE that you turned all burners off?"). This explains domain-specific working memory deficits because WM performance is only disrupted when the WM task requires a component of the central executive that is dysfunctional, e.g., does not suppress intrusive information. If there is no external challenge or the stimulus domain does not induce intrusive thoughts then no performance deficit will be observed. This point should be considered in all WM research that compares OCD performance to typical populations. In the case of our experiments we provided an external challenge and we observed the effects although the stimuli were not related to individual checking domains. Our considerations might also provide an explanation for the potential progression of subclinical checking towards clinical. Checking might be likely to create conditioned associations over time between external or internal challenges of performance (e.g. "Did you/I turn the stove off?") and intrusive thoughts (e.g. "I think I left a burner on") mediated by anxiety (e.g. MacLeod & Mathews, 1991). This would in turn incrementally increase the likelihood of detrimental checking in this specific domain and could lead to self-reinforcement of intrusive thoughts (e.g. Hartston & Swerdlow, 1999).

2. The lack of suppression of misleading information in turn might trigger repeated checking of the EB contents. In the case of our findings, the competition between a visually present probe letter and fragile letter-location bindings in the EB of WM weakens these bindings the more often this competition is repeated. With intrusive thoughts the challenge for the bindings is generated endogenously within the system itself and the more often the bindings are reactivated and their veridicality challenged the less reliable they will become (see Kessler & Kiefer, 2005, for a general account of endogenous vs. exogenous WM reactivation). Paradoxically, while high-scorers check to

improve their performance it actually undermines performance by reducing the accuracy of the WM representation. In that sense checking critically differs from mere rehearsal. Both processes imply reactivation of memory representations, yet, while rehearsal reactivates "without questioning", checking seems to imply a lack of confidence in the veridicality of the reactivated information that is detrimental without the original sensory information to check against, or even with competing new sensory information present.

3. The final component is the consolidation of EB representations into episodic LTM. If the EB representations are progressively weakened by checking then the consolidated representations in LTM will be affected as well, thus, further increasing the likelihood of subsequent checking in LTM that has been shown to decrease accuracy and confidence in episodic representations (van den Hout and Kindt, 2003a and van den Hout and Kindt, 2003b). Self-awareness of repeated loss of accuracy and confidence in memories may finally increase the likelihood and the strength of misleading intrusive thoughts which would then be harder to ignore (e.g. Hartston & Swerdlow, 1999). High checkers might therefore end up in a vicious circle of checking at various stages of memory that does not improve but further deteriorates memory traces. The notion proposed here slightly shifts the explanatory focus from retrieval to consolidation, which has direct clinical relevance. We suggest that a combined WM and LTM explanation might provide a comprehensive etiological starting point for the qualitatively different experience that individuals with checking disorders appear to exhibit in their pathological desire to check and their dissatisfaction with it after it has been executed.

Fig. 7. - Adaptation from Baddeley (2000) to explain the data of the present study. The grey parts of the WM framework highlight the components and their interactions which we propose to be involved in compulsive checking. Further explanations in the text.



The clinical potential of our findings consists in shifting the focus for understanding compulsive checking from retrieval of episodic traces to consolidation and retention in working memory. Spatial WM performance has been previously shown to improve with clinical response in OCD patients indicating that WM performance benefits from therapeutic intervention (Van der Wee et al., 2007). Therefore, our studies may inform current therapies like Attention Training (ATT; Wells, 2000) and Cognitive-Behaviour Therapy (Clark, 2005; see Abramowitz, 2006 for a review) of additional cognitive targets, i.e., the central executive and the episodic buffer. ATT, for example, aims to enhance executive control over attention and cognitive processes through selective attention, attention switching and divided attention exercises. ATT treats spontaneously occurring intrusive thoughts as "noise" that does not require attention but suppression. ATT also attempts to externalise the attentional focus detaching it from the pathological self-referent mode that is observed in checking. Indeed, it has been found that high checkers' memory performance is improved when attentional focus is shifted away from the actual memory task (Ashbaugh & Radomsky, 2007). This suggests that contrary to the checkers' intuition, a relaxing, non-checking attentional focus actually improves memory performance particularly when combined with reduced attention to intrusive thoughts. Klein and Boals (2001) reported that the expressive writing of a negative life event reduced the frequency of intrusive and avoidant thinking, which in turn was associated with an increase in WM capacity indicating that intrusive thoughts may interfere with central executive control over WM. In support of such approaches, our findings emphasise the need for treating the executive deficits in suppressing irrelevant/misleading information and letting episodic (working) memory processes "take their course".

The WM processes we have modulated in our experiments are very transient and rather implicit (i.e. automatic) to obsessional-compulsive thinking (e.g. Harkin and Mayes, 2008 and Henseler et al., 2008). Compared to the subsequent more obvious episodic memory symptoms these transient and implicit processing deficits might be easily overseen for therapeutic interventions. Clinical response may therefore benefit from a process of "guided discovery" (Wells, 2000) where the patient is made aware of such implicit factors. Our results also suggest that mistrust in the veridicality of WM representations might trigger detrimental checking and deterioration of episodic memories in general. Besides training on selective suppression of intrusive thoughts/cues OCD therapy could therefore benefit from enhancing the trust in the veridicality of memories disconfirming the belief that checking is advantageous to memory performance while highlighting its detrimental effect. The ultimate goal would be to achieve a positive long-term effect on trust. This could possibly be achieved for instance by using "external memory devices" like taking/recording notes that would support memories when compulsively checked. In fact research has revealed that high checkers are likely to take more notes as they seem to anticipate memory deficits (Cuttler & Graf, 2007). It would be essential to understand in more detail the exact role of such note-taking: is it merely a futile reflection of anxiety and a lack of confidence or is it a strategy that ensures a stable level of performance for coping with the demands of daily living. In the latter case systematic note-taking could help to increase the trust in memory veridicality and achieve long-term changes. Our research suggests that memories are already detrimentally questioned at the level of WM, so specific training with short retention intervals could be beneficial and easy to implement.

The limitations of our research at this stage are fourfold. Firstly, the use of sub-clinical samples of high checkers limits the conclusions that can be made with respect to clinical populations. However, we argue that the results of our extreme groups meta-comparison (clinically scoring vs. lowest scoring participants across Expts. 1 and 2) substantiate the clinical implications of our findings. Furthermore, the result that even within the typical population checking tendencies impact on WM performance is of importance. Future research, however, should ensure the validity of our claims by using a larger clinically scoring or a clinically diagnosed sample. Also, if the observed WM performance is specific to pathological checking, then it should differ from performance associated with other obsessive-compulsive sub-types (e.g., hoarding, contamination, cf. Abramowitz, McKay, & Taylor, 2005) and other disorders, i.e., generalized anxiety disorder, social phobia and depression.

Secondly, we did not explicitly manipulate checking per se but hypothesized that high checkers are likely to check the content of WM more often relative to low checkers if irrelevant but misleading information is provided. It could however be that checkers are simply more distracted by an irrelevant probe, which reduces the attentional resources required for rehearsing the encoded information (see Cuttler & Graf, 2007, for a similar notion). This does not necessarily imply enhanced checking behaviour per se. Nevertheless, the detriment in memory performance of high checkers was observed for misleading trials only, while resolvable and no-probe-1 trials were comparable. This suggests that it was the misleading content and not the mere presence of a distracting probe that attenuated WM performance. Yet this does not fully rule out the possibility of distraction and future research should directly manipulate checking within a WM paradigm, which unfortunately is not trivial without overly affecting the primary WM task.

Thirdly, the use of letters and locations has limited ecological validity with respect to checkers' idiographic believe systems and anxieties. The use of ecologically valid stimuli within a WM paradigm might reveal even stronger effects than the ones reported here – especially if the high-checking group was drawn from a clinical population. This could also shed light on the implications of anxiety associated with specific checking domains (e.g. MacLeod & Mathews, 1991).

Finally, the basic WM task employed here was very easy. Stronger group differences regarding the impact of misleading information might be revealed with a harder task (cf. Van der Wee et al., 2003).

References

J. Abramowitz, D. McKay, S. Taylor
Special series subtypes of obsessive-compulsive disorder: introduction
Behavior Therapy, 36 (2005), pp. 367‑369

J.S. Abramowitz
The psychological treatment of obsessive-compulsive disorder
Canadian Journal of Psychiatry, 51 (2006), pp. 407‑416

M.M. Antony, F. Downue, R.P. Swinson
Diagnostic issues and epidemiology in obsessive-compulsive disorder
R.P. Swinson, M.A. Antony, S. Rachman, M.A. Richter (Eds.), Obsessive compulsive disorder: Theory, research, and treatment, Guilford Press, New York (1998), pp. 3‑32

A.R. Ashbaugh, A.S. Radomsky
Attentional focus during repeated checking influences memory but not metamemory
Cognitive Therapy and Research, 31 (2007), pp. 291‑306

B.J. Baars
The conscious access hypothesis: origins and recent evidence
Trends in Cognitive Sciences, 6 (2002), pp. 47‑52

A.D. Baddeley
Working memory
Clarendon Press, Oxford, England (1986)

A.D. Baddeley
The episodic buffer: a new component in working memory?
Trends in Cognitive Sciences, 4 (2000), pp. 417‑423

A.D. Baddeley
Working memory: looking back and looking forward
Nature Reviews Neuroscience, 4 (2003), pp. 829‑839

S.G. Ball, L. Baer, M.W. Otto
Symptom subtypes of obsessive-compulsive disorder in behavioral treatment studies: a quantitative review
Behaviour Research and Therapy, 44 (1996), pp. 995‑1006

S. Bannon, C.J. Gonsalvez, R.J. Croft
Processing impairments in OCD: it is more than inhibition!
Behaviour Research and Therapy, 48 (2008), pp. 689‑700

K.T. Ciesielski, M.S. Hamalainen, D.A. Geller, S. Wilheml, T.E. Goldsmith, S.P. Ahlfors
Dissociation between MEG alpha modulation and performance accuracy on visual working memory task in obsessive compulsive disorder
Human Brain Mapping, 28 (2007), pp. 1401‑1414

D.A. Clark
Focus on cognition in cognitive behavior therapy for OCD: Is it really necessary?
Cognitive Behavior Therapy, 34 (3) (2005), pp. 131‑139

M.E. Coles, A.S. Radomsky, B. Horng
Exploring the boundaries of memory distrust from repeated checking: increasing external validity and examining thresholds
Behaviour Research and Therapy, 44 (2006), pp. 995-1006

C. Cuttler, P. Graf
Sub-clinical compulsive checkers' prospective memory is impaired
Journal of Anxiety Disorders, 21 (2007), pp. 228-352

S. Dehaene, L. Naccache
Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework
Cognition, 79 (2001), pp. 1-37

B.L. Harkin, G. Mayes
Implicit awareness of ambiguity: a role in the development of obsessive-compulsive disorder
Behaviour Research and Therapy, 48 (7) (2008), pp. 861-869

H.J. Hartston, N.R. Swerdlow
Visuospatial priming and stroop performance in patients with obsessive-compulsive disorder
Neuropsychology, 13 (1999), pp. 447-457

I. Henseler, O. Gruber, S. Kraft, C. Krick, W. Reith, P. Falkai
Compensatory hyperactivations as markers of latent working memory dysfunctions in patients with obsessive-compulsive disorder: an fMRI study
Journal of Psychiatry and Neuroscience, 33 (3) (2008), pp. 209-215

M.A. van den Hout, M. Kindt
Phenomenological validity of an OCD-memory model and the remember/know distinction
Behaviour Research and Therapy, 41 (2003), pp. 369-378

M.A. van den Hout, M. Kindt
Repeated checking causes memory distrust
Behaviour Research and Therapy, 41 (2003), pp. 301-316

M.A. van den Hout, M. Kindt
Obsessive-compulsive disorder and the paradoxical effects of preservative behaviour on experienced uncertainty
Behaviour Research and Therapy, 35 (2004), pp. 165-181

K. Kessler, M. Kiefer
Disturbing visual working memory: electrophysiological evidence for a role of prefrontal cortex in recovery from interference
Cerebral Cortex, 15 (2005), pp. 1075-1087

K. Klein, A. Boals
Expressive writing can increase working memory capacity
Journal of Experimental Psychology: General, 130 (2001), pp. 520-533

P.A. MacDonald, M.M. Antony, C.M. Macleod, M.A. Richter
Memory and confidence in memory judgments among individuals with obsessive compulsive disorder and non-clinical controls
Behaviour Research and Therapy, 36 (6) (1997), pp. 497-505

C. MacLeod, A. Mathews
Biased cognitive operations in anxiety: accessibility of information or assignment of processing priorities
Behaviour Research and Therapy, 29 (1991), pp. 599-610

T.O. Nelson, L. Narens
Metamemory: a theoretical framework and new findings
in: G. Bower (Ed.), The psychology of learning and motivation, Vol. 26Academic Press, New York (1994)

A. Olley, G. Malhi, P. Sachdev
Memory and executive functioning in obsessive-compulsive disorder: a selective review
Journal of Affective Disorders, 104 (2007), pp. 15-23

I.M. Omori, Y. Murata, T. Yamanishi, S. Nakaaki, T. Akechi, M. Mikuni et al.
The differential impact of executive attention dysfunction on episodic memory in obsessive-compulsive disorder patients with checking symptoms vs. those with washing symptoms
Journal of Psychiatric Research, 41 (2007), pp. 776‑784

S. Rachman
A cognitive theory of compulsive checking
Behaviour Research and Therapy, 40 (2002), pp. 625‑639

Rachman, Shafran
Cognitive and behavioural features of obsessive-compulsive disorder
R.P. Swinson, M.M. Anthony, S. Rachman, M.A. Richeter (Eds.), Obsessive compulsive disorder: Theory, research and treatment, Guilford Press, New York (1998)

A.S. Radomsky, P.T. Gilchrist, D. Dussault
Repeated checking really does cause memory distrust
Behaviour Research and Therapy, 44 (2006), pp. 305‑316

P.M. Salkovskis
Cognitive-behavioral approach to understanding obsessional thinking
British Journal of Psychiatry, 173 (35S) (1998), pp. 53‑63

C.R. Savage, T. Deckersbach, S. Wilhelm, S.L. Rauch, L. Baer, T. Reid et al.
Strategic processing and episodic memory impairment in obsessive-compulsive disorder
Neuropsychology, 14 (1) (2000), pp. 141‑151

A.P. Shimamura
Toward a cognitive neuroscience of metacognition
Consciousness and Cognition, 9 (2000), pp. 313‑323

S. Sternberg
High-speed scanning in human memory
Science, 153 (1969), pp. 652‑654

F. Tallis
The neuropsychology of obsessive-compulsive disorder: a review and consideration of clinical implications
Journal of Clinical Psychology, 36 (1997), pp. 3‑20

D.S. Thordarson, A.S. Radomsky, S. Rachman, R. Shafran, C.N. Sawchuk, A.R. Hakstian
The Vancouver obsessional compulsive inventory (VOCI)
Behaviour Research and Therapy, 42 (11) (2004), pp. 1289‑1314

N.J.A. Van der Wee, N.F. Ramsey, J.M. Jansma, D.A. Denys, H.J.G.M. van Megen, H.M.G. Westenberg et al.
Spatial working memory deficits in obsessive compulsive disorder are associated with excessive engagement of the medial frontal cortex
NeuroImage, 20 (2003), pp. 2271‑2280

N.J.A. Van der Wee, N.F. Ramsey, H.J.G.M. van Megen, D. Denys, G.M. Westenberg, R.S. Kahn
Spatial working memory in obsessive-compulsive disorder improves with clinical response: a functional MRI study
European Neuropsychopharmacology, 17 (2007), pp. 16‑23

A. Wells
Emotional disorders and metacognition: Innovative cognitive therapy
Wiley, Chichester (2000)