Impaired executive functioning in subclinical compulsive checking with ecologically valid stimuli in a working memory task

Ben Harkin, Hannah Rutherford and Klaus Kessler*

Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK

We previously showed that working memory (WM) performance of subclinical checkers can be affected if they are presented with irrelevant but misleading information during the retention period (Harkin and Kessler, 2009, 2011). The present study differed from our previous research in the three crucial aspects. Firstly, we employed ecologically valid stimuli in form of electrical kitchen appliances on a kitchen countertop in order to address previous criticism of our research with letters in locations as these may not have tapped into the primary concerns of checkers. Secondly, we tested whether these ecological stimuli would allow us to employ a simpler (un-blocked) design while obtaining similarly robust results. Thirdly, in Experiment 2 we improved the measure of confidence as a metacognitive variable by using a quantitative scale (0–100), which indeed revealed more robust effects that were quantitatively related to accuracy of performance. The task in the present study was to memorize four appliances, including their states (on/off), and their locations on the kitchen countertop. Memory accuracy was tested for the states of appliances in Experiment 1, and for their locations in Experiment 2. Intermediate probes were identical in both experiments and were administered during retention on 66.7% of the trials with 50% resolvable and 50% irresolvable/misleading probes. Experiment 1 revealed the efficacy of the employed stimuli by revealing a general impairment of high- compared to low checkers, which confirmed the ecological validity of our stimuli. In Experiment 2 we observed the expected, more differentiated pattern: High checkers were not generally affected in their WM performance (i.e., no general capacity issue); instead they showed a particular impairment in the misleading distractor-probe condition. Also, high checkers' confidence ratings were indicative of a general impairment in metacognitive functioning. We discuss how specific executive dysfunction and general metacognitive impairment may affect memory traces in the short- and in the long-term.

Keywords: OCD, compulsive checking, working memory, executive control, ecological validity

INTRODUCTION

In this study we extended our previous research on working memory (WM) performance in subclinical checkers (Harkin and Kessler, 2009, 2011) by using stimuli that are more concordant with clinical symptomatology (Moritz and von Muehlenen, 2008). The rationale being that while we previously reported robust and replicable effects using letters in locations, a central criticism was that these stimuli do not directly relate to checking compulsions in clinical obsessive–compulsive disorder (OCD). This paper therefore provides a necessary methodological step forward by employing electrical kitchen appliances that are more relevant to checkers primary concerns (Rachman, 2002; Thordarson et al., 2004). Checking compulsions are most commonly observed in OCD with 50–80% of patients reporting this subtype (Henderson and Pollard, 1988; Rasmussen and Eisen, 1988; Antony et al., 1998). The motivation of pathological checking appears to reflect their distrust in a previous action and/or thought and so they check and recheck to compensate for their own perceived shortcomings (Rachman and Shafran, 1998). However, their attempts to overcome distrust paradoxically increases it, as repetitive checking has been observed to reduce memory accuracy (Radomsky and Alcolado, 2010), memory for details and confidence (van den Hout and Kindt, 2003).

Thus, it appears that memory is tainted by the process of checking, but as Rachman (2002) proposed, any cognitive theory must account for the specific mechanisms underlying memory impairment. The simplest interpretation suggests that basic capacity issues are responsible. However, a body of research indicated that this is not the case, with general capacity intact (e.g., Ciesielski et al., 2007; Henseler et al., 2008) and impairment only observed as task difficulty increases (e.g., Purcell et al., 1998a,b; e.g., van der Wee et al., 2003, 2007) or if a task places a high emphasis upon efficient executive control (Savage et al., 1999, 2000; Harkin and Kessler, in press). This suggests that rather than being attributable to a general mnestic deficit, memory impairment is secondary to executive dysfunction (Greisberg and McKay, 2003; van der Wee et al., 2003, 2007; Penades et al., 2005; Olley et al., 2007; Omori et al., 2007; Bannon et al., 2008).

In explanation, Harkin and Kessler (2009) originally proposed Baddeley's (2000) model of WM as a unifying framework to account for the domain and/or episode specific nature of memory impairment in OCD (e.g., "Did I turn the iron off?"). Baddeley's (1986) original model included the central executive for higher order control and two slave systems for the temporary maintenance of visuospatial (i.e., visuospatial sketchpad) and phonemic (i.e., phonological loop) information. While this model explained a range of experimental data it failed to account for the manner in which the cognitive system bound information from multimodal sources and accurately maintained them over a delay (see Baddeley, 2000 for review). As a solution Baddeley (2000) introduced the episodic buffer which explained the binding of multimodal information into temporally integrated representations. By multimodal we refer to aspects of a complex visual scene that are processed in different cortical streams, specifically: the ventral and dorsal processing streams for object and location representations, respectively. Therefore, if accurate task performance is dependent upon accurate object ("What") and location ("Where") information then this will rely upon the maintenance of accurate object–location conjunctions in what Baddeley had termed the "episodic buffer" (see Olson et al., 2006; Keizer et al., 2008). For example, Prabhakaran et al. (2000) reported that separately maintaining verbal (letters) and location information activated more posterior and disparate brain regions (*letters*: left inferior/Broca and left inferior parietal and temporal vs. *locations*: right frontal and bilateral superior parietal) compared to when letter–location information was bound (e.g., right prefrontal cortex). Also in a study of patients with medial temporal lobe amnesia, memory was impaired for accurate conjunction WM but was normal for independent objects and locations (Olson et al., 2006). Therefore, while the neuropsychological basis of binding is debatable, it is evident that objects and locations are processed independently (likely ventral vs. dorsal, respectively) but when required in a multimodal format (letter–location conjunctions) rely on a mediating system, akin to the episodic buffer.

Indeed, while there has been some dispute regarding the exact mechanism underlying multimodal binding in the episodic buffer, researchers tend to agree that attentional effort is required for its generation and maintenance (Wheeler and Treisman, 2002; Delvenne and Bruyer, 2006; Fougnie and Marois, 2009; Hyun et al., 2009). Therefore, memory impairment occurs if distraction is sufficient to interfere with attention-dependent bindings. Considering the established sensitivity of bindings to interference (see Wheeler and Treisman, 2002; Kessler and Kiefer, 2005; Delvenne and Bruyer, 2006; Fougnie and Marois, 2009; Hyun et al., 2009) we proposed that if a task taps into the executive deficits of OCD/checkers, then this will interfere with attention allocated to bindings, thus, impairing memory as a consequence (for a review, see Harkin and Kessler, in press). Accordingly, in our previous experiments (Harkin and Kessler, 2009) we had set out to: (1) *engage* the episodic buffer by using a memory task where accuracy was dependent upon the veridical maintenance of letters to locations, and then (2) *hamper* episodic buffer functionality by presenting information that was relevant to the executive impairments of high but not low checkers during the WM retention interval. Specifically, for the latter we probed an item that had *not* been presented during encoding: "Where was letter 'K'?" – while there had been no letter "K" in the

encoding set. We found that only high checkers' performance on the actual WM test (probe 2) was impaired in the context of a misleading probe (probe 1). Considering that an intermediate probe is irrelevant to the performance of the memory test, we conclude that checkers are more distracted by a misleading probe as it is not part of the encoded set. Checkers either cannot suppress the distractor itself, or cannot suppress the urge to check triggered by the misleading distractor. A process which we suggest is perhaps driven by impairment in inhibitory functioning specific to the checking but not the washing subtype (Omori et al., 2007). However, in our second series of experiments we observed that checkers' suffered similar memory impairments for resolvable and misleading spatial probes (see Experiment 2; Harkin and Kessler, 2011). Thus, while there is a delicate balance between resolvability and general distraction, in either case it appears that checkers' poorer memory is due to an executive deficit of inhibitory functioning which impairs attention-dependent bindings within the episodic buffer.
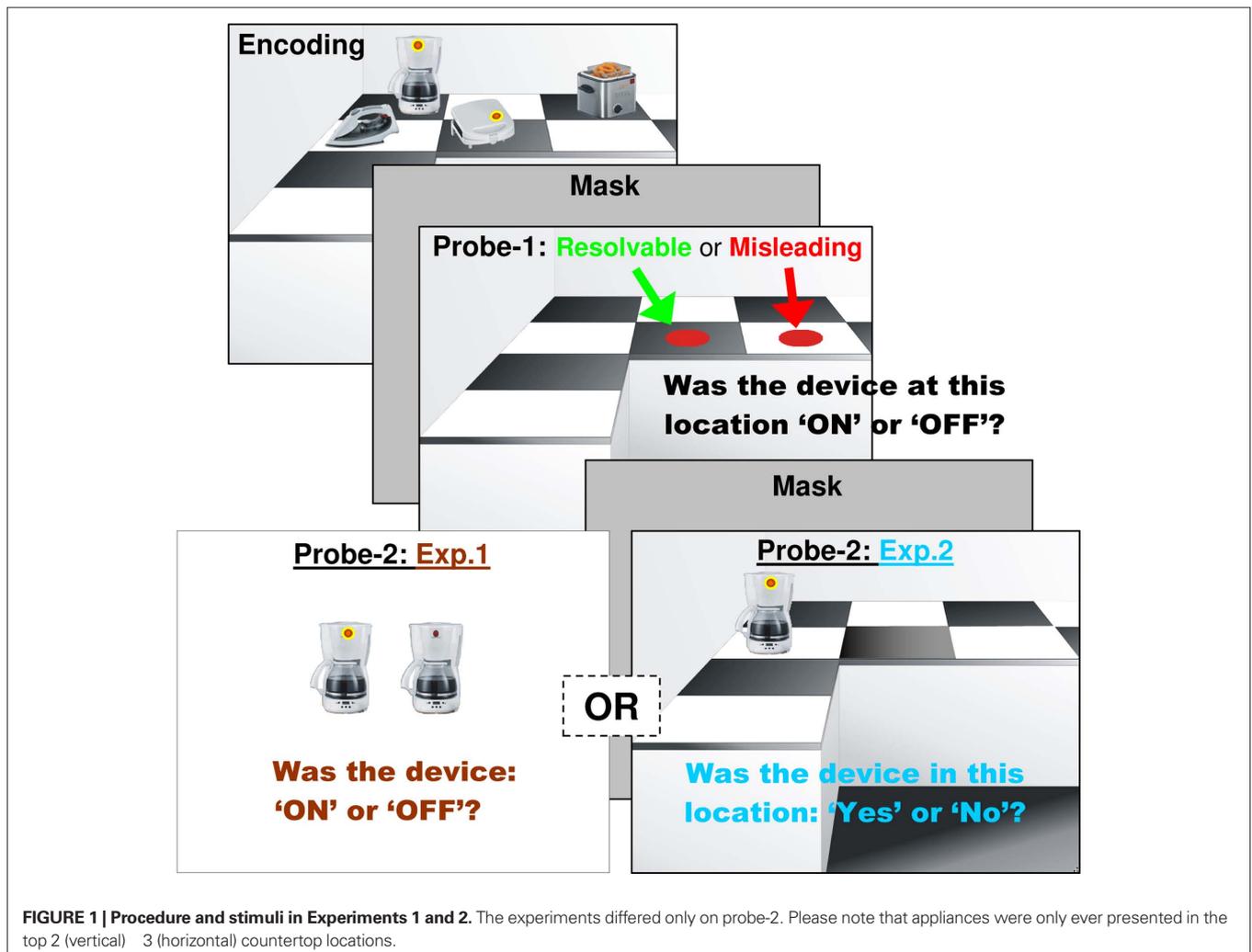
This explanation is supported by Omori et al. (2007) who reported in a clinical OCD patient group, that only for clinical checkers (not washers) were deficits in inhibition associated with poor episodic memory. Not only does this highlight the central role of inhibition in checking and OCD generally (see Chamberlain et al., 2005) but also suggests that this dysfunctional aspect of executive control may exist on a continuum between subclinical to clinical checking. For example, subclinical checkers have shown similar deficits to those observed in clinical OCD, i.e., the Wisconsin Card Sorting Task (Gershuny and Sher, 1995) and the Wechsler Memory Scale (Sher et al., 1984). Further, this same group have shown memory deficits for everyday activities (Sher et al., 1984), prospective memory impairments (Cuttler and Graf, 2007, 2008, 2009), and were poorer at distinguishing real from imagined events (Rubenstein et al., 1993). This has lead some researchers to suggest that a subclinical analog is a valid means of understanding a variety of features relevant to clinical OCD, especially as they are free from confounds such as medication, clinical state, or co-morbidity (Mataix-Cols et al., 1997, 1999). Indeed, considering this alongside the commonality of checking in OCD (50–80%; Henderson and Pollard, 1988; Rasmussen and Eisen, 1988; Antony et al., 1998) and the population generally (15%; Stein et al., 1997), subclinical checkers may provide a "purer" means for determining the specific impact of executive deficits upon WM functioning.

The notion that executive dysfunction and memory impairment is specific – i.e., only occurs when a memory task calls upon a dysfunctional component of the executive – is further supported in the literature. For example, van der Wee et al. (2003) used a spatial variant of the *n*-back WM task with four levels of load. It was only at the highest load level (3-back) that patients with OCD significantly differed from controls with errors of 48 vs. 25%, respectively. They argued that OCD patients may over-scrutinize their performance or have a deficit in supervisory (i.e., executive) processes, as opposed to deficits in maintenance or manipulation, which suggests that general capacity limitations are not responsible for the results. The stability of executive-memory impairment at higher levels of task complexity have been observed across a range of WM tasks, for example, the spatial WM task (Purcell et al., 1998a,b), paired association learning (Morein-Zamir et al., 2010), and the corsi block tapping task (Zielinski et al., 1991; Zitterl et al., 2001;

Moritz et al., 2003; Boldrini et al., 2005). In these instances OCD memory impairments are not attributable to capacity *per se* (i.e., intact at lower load levels; see also Ciesielski et al., 2007; Henseler et al., 2008) but rather represent a failure of executive functioning to match increasing task demands in terms of strategic resource organization. To reiterate, these findings indicate that successful WM performance is dependent upon sustained and correctly allocated attention (i.e., executive control), which locates poor memory in our studies and in others' research within the domain of executive dysfunction as opposed to impairments of basic WM capacity. We are careful here to note that executive functioning is intertwined with basic WM storage at all levels of task requirements and/or difficulty. However, in terms of the locus of impairment we agree with view that memory impairment observed here and in others' research is secondary to executive dysfunction (Greisberg and McKay, 2003).

Thus, considering these points alongside the methodological limitations of our previous research (using letters) we presented four electrical kitchen appliances located in six possible locations, of which two were "ON" (electrical light was bright red) and two were "OFF" (electrical light was dark red). The primary memory task (probe-2) required the participants to recall if an appliance had been "ON" or "OFF" (Experiment 1) or if an appliance was correctly located (Experiment 2) as shown in **Figure 1**. In both experiments, we used an intermediate spatial-location probe similar to Experiment 2 of Harkin and Kessler (2011), where it had produced stable group effects (i.e., low standard deviations) and substantial memory impairments in high compared to low checkers. This intermediate probe was presented at a location where an appliance had either been present (resolvable) or at a location that had been completely empty (misleading), participants had to indicate if the appliance at that location had been "ON" or "OFF". An additional yet critical development of our methodology related to trial-type ratio. In our previous experiments we presented two blocks, one with predominantly misleading trials (66%) and a counterbalanced block of resolvable trials as a result we could not exclude the influence that this had upon checkers' WM performance. Therefore, we currently used an equal trial-ratio (33% resolvable, 33% misleading, 33% no-probe-1) which allowed us to develop a clearer understanding of the specific effect(s) of trial-type and/or group on memory performance (probe-2). We predict that using such stimuli and probing the spatial location of threatening

FIGURE 1 | Procedure and stimuli in Experiments 1 and 2. The experiments differed only on probe-2. Please note that appliances were only ever presented in the top 2 (vertical)   3 (horizontal) countertop locations.

aspects of them may potentially enhance executive dysfunction, impair attention-dependent bindings (i.e., Experiment 1: state to appliance or Experiment 2: appliance to location) and perhaps produce novel memory and metacognitive impairments compared to our previous work.

EXPERIMENT 1
METHODS
Stimuli in Experiment 1 (and 2)
With the predominance of checking in OCD we employed ecologically valid stimuli that were concordant with this symptomatology. For example, the Vancouver Obsessional–Compulsive Inventory (VOCI; Thordarson et al., 2004), and the checking subscale specifically, ask respondents to indicate if they repeatedly check and recheck things like "switches, faucets, appliances, and doors" and "that the stove is turned off" (Thordarson et al., 2004). Additionally, Rachman (2002) highlighted the specific nature of perseveration: "Yes, I remember that I did check the stove but I cannot remember if I checked it satisfactorily. Was the switch fully turned off? I cannot remember if it is safe" (p. 631). Accordingly, we used images of electric kitchen appliances (fryer, iron, kettle, toaster, coffee machine, hob, microwave, sandwich maker) as encoding set stimuli and then asked two specific memory questions with respect to these stimuli in Experiment 1 (Appliance "ON/OFF" State) and Experiment 2 (Appliance Location).

Participants
A total of 40 participants (mean 20.8 years: 12 males, 28 females) from the University of Glasgow gave written informed consents. British Psychological Society ethical requirements were met, including that of participant debriefing. The 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: 20 low (mean: 0.5, SD: 0.61) and high (mean: 13.85, SD: 4.12) "checkers." Further, no statistical differences between the low and high groups were revealed in gender distribution ($p = 0.72$) or age ($p = 0.27$).

Procedure
Participants sat 60 cm from a computer screen with their head on a chin rest. At the beginning of each trial a fixation cross was presented for 2000 ms. A kitchen countertop was then presented for 6000 ms with four electrical kitchen appliances presented randomly in six possible locations as shown in **Figure 1**. Two of these appliances were "ON" as indicated by a red light and two were shown to be "OFF" with no accompanying light. After this a mask was presented for 1000 ms, this was to reduce the influence that possible image retention may have played in subsequent retrieval (i.e., distinct appliances and/or their "ON" states), thus isolating disturbances in later memory-probe performance to those of WM. After this a probe-1 question asked if a device at a specific location was either "ON" or "OFF." As in our previous research (Experiment 2; Harkin and Kessler, 2011) this probe was presented (3000 ms) at a location where there *had been* (resolvable) or *had not been* (misleading) a device in the original encoding set. Participants were asked to indicate if the device at this location (resolvable or misleading) was either "ON" (left

index finger of right hand) of "OFF" (middle index finger of right hand). This probe previously produced stable group effects (i.e., low standard deviations) and substantial memory impairments in high compared to low checkers. Additionally, using such an intermediate probe was motivated by our recent findings that an explicit, yet task irrelevant "ON" cue interfered with normal inhibitory functioning (i.e., inhibition of return (IOR); Posner and Cohen, 1984) of a high scoring subclinical OCD/checking group but not a low scoring group (Experiment 2). Thus, for high checkers drawing attention to the functional and threatening aspects of electrical appliances and probing empty locations may resonate with the established executive impairments of high checkers in inhibiting irrelevant thoughts and/or stimuli (Savage et al., 2000; Olley et al., 2007; Omori et al., 2007). Baseline trials were also included; these presented an empty kitchen countertop (i.e., no-probe-1) designed to measure WM under ideal conditions. A mask was again presented (1000 ms) before the actual memory task. In Experiment 1, probe-2 simply presented a single electrical appliance at the center of the screen, the participant had to indicate if they recalled it as being "ON" (right index finger) or "OFF" (right middle index finger) with respect to the original encoded set. Finally, participants were asked to indicate their confidence in their probe-2 decision as indicated simply by a "Confident" (right index finger) or "Not Confident" (right middle index finger) response.

There were 156 trials in total, 12 of which (at the beginning) were practice including resolvable and no-probe-1 trials only. The main experiment was then done in two blocks (with 5 min rest period between), each comprising 24 resolvable, 24 misleading, and 24 no-probe-1 trials presented in random order. Importantly, we employed an equal ratio of trial type in the current experiments: 33% resolvable, 33% misleading, 33% no-probe-1, while in our previous studies we had employed at least one block with 66% misleading trials (and a counterbalanced block of predominantly resolvable probe-1 trials, cf. Harkin and Kessler, 2009, 2011). We did this to remove the influence of trial-type ratio which had to be counterbalanced across two blocks in our previous experimental designs. This allowed us to develop a clearer understanding of the specific effect(s) of trial-type and/or group on memory performance (probe-2). For example, in our original experiment (Harkin and Kessler, 2009) it is possible that high checkers' poor performance on misleading trials was driven by the novelty/surprise caused by an unfamiliar trial type.

Design
A two (*Group*: low vs. high checkers) by three (*Probe-1*: resolvable, misleading, no-probe-1) by two (*Probe-2 State*: ON, OFF) mixed design was employed with *group* as the between- and *probe-1* and *probe-2 state* as the within-subjects factors.

RESULTS AND DISCUSSION
MANOVAs for a 2 3 2 design were carried out for reaction times, accuracy and confidence on probe-2 responses due to violations of the sphericity assumption (Mauchley's tests).

Probe-2 Response Latencies
The MANOVA (2 3 2) for probe-2 latencies revealed a main effect of group, high checkers (1898.4 ms) were significantly slower in responding than the low group [1573.4 ms; $F(1,38) = 10.65$, $p = 0.047$,

$\eta_p^2 = 0.099$]. A main effect for trial type [$F(2,76) = 5.59$, $p = 0.005$, $\eta_p^2 = 0.128$], reflected slower RTs overall for misleading trials compared to resolvable [$F(1,38) = 9.32$, $p = 0.004$] and no-probe-1 trials [$F(1,38) = 9.20$, $p = 0.004$]. This suggests that for all participants making a probe-2 location decision is particularly sensitive to a misleading intermediate probe: encouraging participants to examine the state of an appliance at a location where there is none slows subsequent location based responding. A significant main effect for probe-2 state [$F(1,38) = 24.7$, $p < 0.001$, $\eta_p^2 = 0.393$] revealed that all participants were slower in responding to an appliance that was "OFF" (1847.6 ms) compared to "ON" (1624.2 ms) in the encoded set.

Probe-2 Accuracy
The MANOVA (2   3   2) for probe-2 accuracy revealed a main effect of group [$F(1,38) = 4.27$, $p = 0.046$, $\eta_p^2 = 0.101$], with high checkers (87.3%) significantly less accurate than the low group (94.1%). Importantly, as high checkers were significantly slower in making their responses, we can rule out a speed-accuracy trade-off as an explanation for their poorer accuracy. A main effect for trial type [$F(2,76) = 4.08$, $p = 0.046$, $\eta_p^2 = 0.100$], reflected no-probe-1 trials were more accurate than resolvable [$F(1,38) = 5.93$, $p = 0.02$, $\eta_p^2 = 0.097$] or misleading trials [$F(1,38) = 6.70$, $p < 0.05$]. Therefore, for all participants an intermediate probe resulted in poorer probe-2 state accuracy compared to trials with no intermediate probe.

Confidence Responses
The MANOVA (2   3   2) for confidence responses concentrated upon the total "not-confident" responses of each participant in each condition. A main effect for trial type [$F(2,76) = 7.99$, $p = 0.002$, $\eta_p^2 = 0.153$] reflected lower confidence for all participants for misleading trials compared to resolvable [$F(1,38) = 4.60$, $p = 0.038$] and no-probe-1 trials [$F(1,38) = 10.27$, $p = 0.002$]. Also a main effect of probe-2 state [$F(1,38) = 26.68$, $p < 0.001$, $\eta_p^2 = 0.412$] indicated that all participants had less confidence for an electrical appliance that had been "OFF" than "ON." No effects involving group reached significance.

To sum up, we found a general accuracy deficit for high checkers that could reflect general capacity issues. However, based on our previous research (Harkin and Kessler, 2009, 2011) and research reported by others (Ciesielski et al., 2007; Henseler et al., 2008), we did not believe this to be the case. In contrast, we hypothesized that the employed probe-2 may have focused checkers' attention too strongly on the threatening aspect of the stimuli (electric on/off status), hence introducing a generally higher level of interference during encoding, maintenance, and/or retrieval fuelled by anxiety. Hence, we devised a second experiment that differed from Experiment 1 regarding the feature dimension of the memory test (probe-2). Instead of probing the state of an appliance (on vs. off) we probed its location (correct vs. incorrect). We expected a more differentiated pattern across conditions with a special role for misleading trials.

EXPERIMENT 2
METHODS
Participants
A total of 40 participants (mean: 21.85 years, 13 males and 27 females) from the University of Glasgow gave written informed consents. British Psychological Society ethical requirements were met, including that of participant debriefing. As before, the checking subscale was used to obtain two groups: 20 low (mean: 0.0, SD: 0.0) and high (mean: 13.75, SD: 6.16) "checkers." Further, no statistical differences between the low and high groups were revealed in gender distribution ($p = 0.31$) or age ($p = 0.58$).

Procedure
Experiment 2 was identical to Experiment 1 with two exceptions. (1) *Probe-2*: We presented an electrical appliance either at the correct (50%) or incorrect (50%) location with respect to the encoding set and asked participants to indicate if it was correctly or incorrectly located (see **Figure 1**). (2) *Confidence*: We asked participants to indicate their confidence on a sliding scale from 0 (no confidence at all) to 100 (complete confidence). We expected this scale to be more sensitive in detecting between-group differences in meta-cognition than the binary response option employed in Experiment 1.

There were 156 practice trials in total, 12 of which were practice trials including resolvable and no-probe-1 trials only. The main experiment was then done in two blocks (with 5 min rest period between), each comprising of 24 resolvable, 24 misleading, and 24 no-probe-1 trials presented in random order. As in Experiment 1 an equal ratio of misleading, resolvable and no-probe-1 trials were used.

Design
A two (*Group*: low vs. high checkers) by three (*Probe-1*: resolvable, misleading, no-probe-1) by two (*Probe-2 Location*: Correct, Incorrect) mixed design was employed with *group* as the between- and *probe-1* and *probe-2 location* as the within-subjects factors.

RESULTS AND DISCUSSION
MANOVAs for a 2   3   2 design were carried out for reaction times, accuracy, and confidence on probe-2 responses due to violations of the sphericity assumption (Mauchley's tests).

Probe-2 Response Latencies
A main effect of trial type [$F(2,76) = 4.01$, $p = 0.022$, $\eta_p^2 = 0.095$] reflected the linear increase in RTs across resolvable (1847.4 ms), misleading (1943.9 ms), and no-probe-1 trials (2019.9 ms). We suggest that the presence of an intermediate probe (resolvable or misleading) may focus the attention of checkers to responding which primes them to subsequent responding, leading to faster responding in these conditions compared to when no intermediate probe (i.e., no response priming) is presented. This pattern was previously observed in our original experiments, which when considered in relation to the different probe-1 RTs of Experiment 1 (Misleading > Resolvable = No-Probe-1) indicates that the relationship between probe-1 and the specificity of probe-2 is sufficient to influence RTs. A main effect of probe-2 location [$F(1,38) = 39.31$, $p < 0.001$, $\eta_p^2 = 0.508$] showed that participants responded slower to an appliance that was correctly located (2067.8 ms) with respect to the encoded set compared to one that was incorrectly located (1806.4 ms).

Probe-2 Accuracy
The main effect of probe-2 location reached significance [$F(1,38) = 42.86$, $p < 0.001$, $\eta_p^2 = 0.530$], which reflected poorer accuracy for correctly (79.4%) compared to incorrectly located

appliances (94.5%). When considered alongside the RT main effect for probe-2 this suggests that correctly located appliances are more difficult to resolve which is reflected in slower RTs. In explanation, an incorrect location can be disproved by at least two partial representations such as remembering which object actually had been in the probe location or by remembering the correct location of the probe object. This is not the case for correct probes where this particular object–location binding has to be received veridically. Group trial type was significant [$F(1,38) = 3.42, p = 0.038, \eta_p^2 = 0.082$]. Analysis of the simple main effects for group at each level of trial-type revealed a significant group difference (low = 90.5% vs. high = 82.3%) for misleading trials [$F(1,38) = 7.52, p = 0.009$; see **Figure 2**], whereas, for resolvable and no-probe-1 trials no statistically significant group difference was observed ($p = 0.084$ and 0.366, respectively). We further analyzed the simple main effects within each group to determine the locus of between-condition performance differences. For the low group, no differences were reported between resolvable, misleading, or no-probe-1 trials (i.e., all $p > 0.3$). On the other hand, for high checkers, responses were less accurate for misleading trials than no-probe-1 trials [$F(1,38) = 5.99, p = 0.019$], but responses for resolvable and no-probe-1 trials were similarly accurate ($p = 0.361$). Overall this suggest that the significant interaction between group trial is due to the special role of misleading trials within the high checkers as well as with respect to group differences.
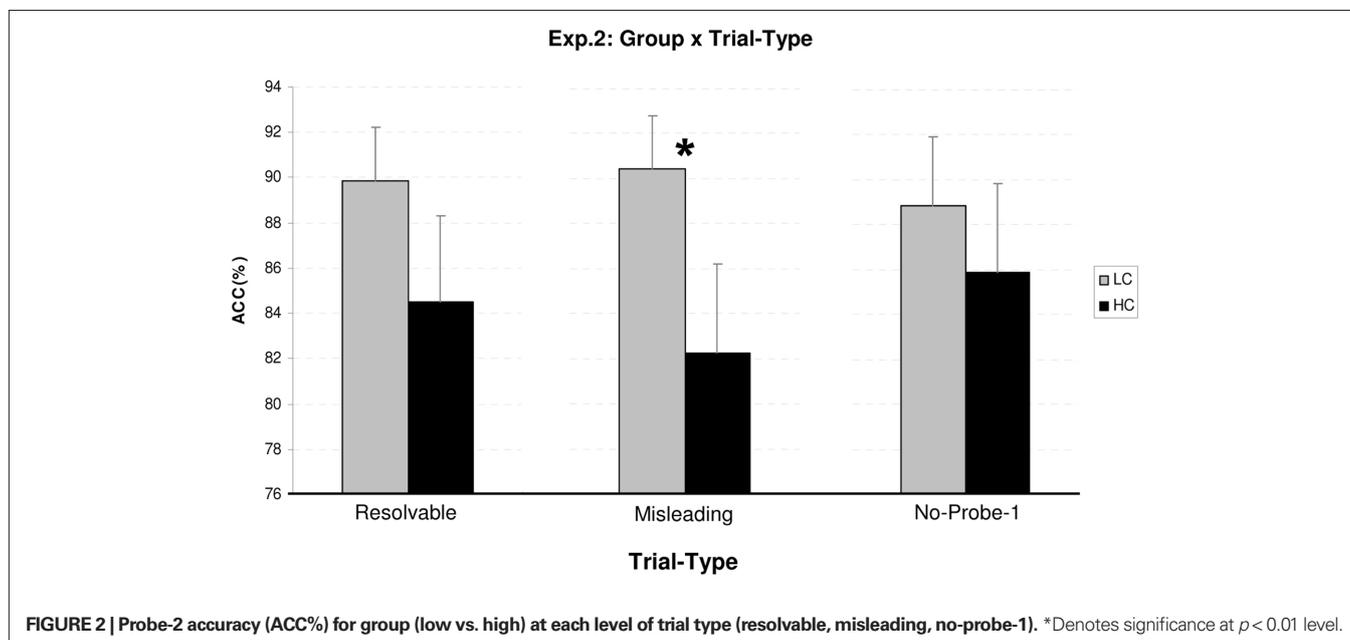
Confidence Ratings

A main effect of group revealed [$F(1,38) = 5.30, p = 0.027, \eta_p^2 = 0.122$] that high checkers (70.67) had poorer confidence overall compared to low checkers (80.71). Trial-type reached significance [$F(2,76) = 5.87, p = 0.004, \eta_p^2 = 0.134$], which was driven by poorer confidence on misleading trials compared to resolvable [$F(1,38) = 4.67, p = 0.037$] and no-probe-1 trials [$F(1,38) = 8.15,$

$p = 0.007$]. This suggests that a misleading intermediate probe was sufficient to reduce confidence in all participants. Probe-2 location reached significance [$F(1,38) = 20.51, p < 0.001, \eta_p^2 = 0.351$] and reflected less confidence for correctly compared to incorrectly located appliances. Poorer confidence for a correctly located appliance reflected the poorer accuracy that all participants had in this condition. The group probe-2 location interaction approached significance [$F(1,38) = 3.65, p = 0.064, \eta_p^2 = 0.087$], with group differences observed for incorrectly [$F(1,38) = 8.31, p = 0.006$] but not correctly located appliances [$F(1,38) = 2.23, p = 0.144$]. Thus, the low checkers mirrored the general trend of the probe-2 location main effect (i.e., poorer performance for correct than incorrect), whereas the high group had poorer confidence across both conditions.

Correlations between accuracy and confidence were conducted for each group and both groups showed significant relationships (low group: $r = 0.56, n = 20, p = 0.01$; high group: $r = 0.71, n = 20, p = 0.000$) indicating that for all participants confidence mirrors accuracy. In a further analysis we subtracted confidence scores from accuracy scores for each participant in each condition, which produced what we termed a discrepancy score. A discrepancy score of 0 indicates that accuracy and confidence mirror each other, whereas an increasing discrepancy score indicates that confidence is numerically less than preceding accuracy. We were primarily interested in group differences in discrepancy scores across trial-types, as this could indicate conditions, where confidence and accuracy might only diverge in high checkers, revealing a metacognitive deficit.

In a MANOVA analysis of the discrepancy scores the interaction between group trial-type reached significance [$F(1,38) = 3.14, p = 0.049, \eta_p^2 = 0.076$]. Analysis of the simple main effects for group at each level of trial-type revealed a significant group difference (LC = 6.42 vs. HC = 14.18) for no-probe-1 trials [$F(1,38) = 5.42, p = 0.025$] but not

FIGURE 2 | Probe-2 accuracy (ACC%) for group (low vs. high) at each level of trial type (resolvable, misleading, no-probe-1). *Denotes significance at $p < 0.01$ level.

for resolvable [$F(1,38) = 0.60$, $p = 0.442$] or misleading trials [$F(1,38) = 0.76$, $p = 0.389$]. This indicates that low and high checkers confidence-accuracy discrepancy is similarly inflated in trials when there is an intermediate probe: accuracy is greater than confidence. However, in no-probe-1 trials low checkers accuracy–confidence is more concordant (6.42) compared to high checkers whose discrepancy score (14.18) is similar to that observed in resolvable (12.47) and misleading trials (14.04). We interpret that high checkers suffer a task independent impairment in their metacognitive functioning which is expressed here as less confidence in their accuracy on no-probe-1 trials.

DISCUSSION

The present experiments used electrical kitchen appliances that were concordant with the symptomatology of those afflicted with obsessive–compulsive checking (Rachman, 2002; Thordarson et al., 2004). We did this in an attempt to address a primary criticism of our previous research (Harkin and Kessler, 2009, 2011) that letters in locations do not resonate with the primary concerns of checkers. We predicted that for high checkers using episodically rich stimuli and questioning a threatening aspect of them (i.e., "ON/OFF" state of probe-1) would provide a greater challenge to the attention-dependent bindings required for accurate memory recall. In a separate study on IOR with the same kitchen appliances as employed here, we indeed confirmed that drawing attention directly to the functionality of these electric appliances turned IOR into positive priming in high- but not in low scoring OCD/checkers in a subclinical sample. In short, this means that high scorers' attention perseveres on a threatening stimulus once it was drawn to it, underpinning the ecological validity of our stimuli.

We observed that group effects differed between experiments, a finding we attribute to employing ecologically valid stimuli and probing different features of the memory in Experiment 1 (electric state on/off) compared to Experiment 2 (location). Experiment 1 supported our claim that our stimuli were compatible with OCD/checking symptomatology by revealing a main group effect in reaction times and accuracy. However, reaction times and accuracy data also indicated that the particular manipulations in Experiment 1 may have resulted in a degree of interference in all participants. Specifically, probe-2 reaction times were slower after a misleading intermediate probe, suggesting that this experiment encouraged all participants to access the "ON/OFF" states of the appliances which then slowed subsequent responding to a state-based probe-2 question. Memory decisions regarding appliances' "ON" or "OFF" states (probe-2) were significantly slower for all participants in misleading compared to resolvable or no-probe-1 trials. Memory accuracy was significantly poorer after resolvable and misleading trials compared to no-probe-1 trials. So for all participants continually focusing on ON/OFF states appears to have come at the cost to their performance. Together, the strengths of these general effects could have been sufficient to obscure group effects but this proved not to be the case: High checkers were generally slower and poorer at recalling the state of an electric appliance compared to low checkers.

As we did not include an independent cognitive index of WM functioning, high checkers' poorer accuracy overall (compared to low scoring checkers) could be interpreted as impaired WM capacity. However, we argue against this for a number of reasons (for a review see Harkin and Kessler, in press). Firstly, if checkers have a general WM capacity impairment then this would have influenced our previous results (Harkin and Kessler, 2009, 2011). A general impairment would negatively affect WM performance irrespective of the content of the encoded set, i.e., similar no-probe-1 impairment for letters and electrical appliances. Secondly, if checkers suffered from basic capacity impairment, then memory would *not* be influenced by the specificity of the probe-2 question, whereby they would necessarily have impaired appliance-location (Experiment 2) memory in the no-probe-1 condition. Thirdly, there is a convergence of evidence showing that basic WM capacity is intact (Ciesielski et al., 2007; Henseler et al., 2008) with impairment only observed at high load levels when tasks stress dysfunctional components of executive control in OCD patients (Zielinski et al., 1991; Purcell et al., 1998a,b; Zitterl et al., 2001; Moritz et al., 2003; van der Wee et al., 2003; Boldrini et al., 2005; Morein-Zamir et al., 2010). Finally, considering that in simple memory tasks subclinical checkers have outperformed OCD patients (Tuna et al., 2005) and controls (Irak and Flament, 2009), it is unlikely that our group of subclinical checkers had anomalous capacity issues. Rather, it is likely that they have executive impairments analogous to those observed in clinical OCD (Mataix-Cols et al., 1997, 1999; Omori et al., 2007), which interferes with efficient state–appliance–location bindings during encoding and/or maintenance. This is in agreement with the perspective that memory impairments in OCD are secondary to executive dysfunction (Greisberg and McKay, 2003) and it is further in agreement with the metacognitive deficit revealed in Experiment 2.

The differences between low and high checkers were somewhat more subtle, yet even more revealing in Experiment 2: (1) Performance of high checkers was significantly affected on misleading trials compared to baseline (no-probe 1 trials). (2) The misleading condition revealed the strongest group difference with the best performance for low- and the worst performance for high checkers across all trial conditions. (3) In contrast to Experiment 1, high checkers' performance on no-probe-1 trials did not significantly differ from the performance of low checkers. Finally, there was a statistical trend for a group difference on the resolvable trials that was reminiscent of the significant differences we had observed before with a spatial probe and abstract stimuli (letters in locations, Experiment 2 in Harkin and Kessler, 2011). There, a spatial probe had been generally distracting for high checkers. Here however, when the stimuli were relevant to checkers' symptoms (electric appliances with switches) a misleading probe provides additional impairment to that caused by an intermediate spatial probe resulting in the main, statistically reliable difference. This is corroborated by the significant interaction between group and trial type and further detailed analysis which revealed that high checkers performed significantly worse on misleading compared to baseline trials while performance on resolvable compared to baseline trials did not significantly differ. In contrast, the performance of low checkers did not significantly differ for any trial-type comparison.

In explanation, based on Experiment 1 we argue that checkers' attention is generally biased toward the threatening aspects of the appliances. In Experiment 2 this is moderated by the emphasis on spatial locations of probe-2, but may still provide high checkers with a slight advantage in accessing the state of an appliance at a resolvable compared to a misleading location during probe-1. This may explain why the group difference for resolvable trials did not reach significance while it did for misleading trials. We argue that our explanation in terms of attention biased by the threatening aspects of the stimuli may be particularly true when locations are being challenged during probe 2 (cf. Experiment 2) rather than if stimuli identities are challenged. This proposal is based on the results of Harkin and Kessler (2009, 2011) that showed that high checkers exhibited memory impairments when questioned about the location of a certain stimulus, but not when questioned about the identity of a stimulus at a certain location. That is, maintaining the correct location of an appliance in WM depends more strongly on sustained attention than maintaining the identity of the appliance. Indeed, identity representations may be harder to disrupt than location representations because the identity of a stimulus is based on concepts stored in long-term memory (LTM), whereas the location of a stimulus is arbitrary and specific to the experimental context. In contrast to our previous studies, however, we employed an equal ratio of misleading, resolvable, and no-probe-1 trials throughout our two experiments (in contrast to counterbalanced ratios across two blocks in Harkin and Kessler, 2009, 2011) which further underpins the robustness of our findings with ecologically valid scenarios.

Finally, we suggest that high checkers' intact no-probe-1 performance in Experiment 2, in contrast to generally impaired performance in Experiment 1, is due to task differences regarding the memory probe (probe 2). Specifically, Experiment 1 required the accurate recall of the appliances' "ON/OFF" status while Experiment 2 probed the correct location of an appliance. As this no-probe-1 impairment was neither previously reported (Harkin and Kessler, 2009, 2011) nor was it observed in Experiment 2, the locus of the difference must be specific to the probe-2 task in Experiment 1 where attention was again focused on the threatening aspects of the stimuli (electric on/off status). We propose that this may have in turn affected the encoding, maintenance and/or retrieval of multimodal bindings in Experiment 1 in form of interference fuelled by anxiety. In fact we regard the group main effect in Experiment 1 as confirmation of the ecological validity of our stimuli.

While group differences in confidence were not observed in Experiment 1, Experiment 2 revealed a group main effect for a lack of confidence in high scorers. This highlights that a continuous confidence scale (Experiment 2) is not only more sensitive for detecting group effects but it also lends itself to a wider range of statistical analyses compared to the binary forced-choice (Experiment 1). The main effect in Experiment 2 indicates that high checkers have a global (trial-type independent) impairment in confidence compared to low checkers. That is, although correlations where high between accuracy and confidence for both groups, only high checkers showed a significant discrepancy for the no-probe 1 trials. This dissociation between performance and confidence in the baseline condition in particular, suggests a metacognitive deficit in form of impaired performance monitoring that is present in high- but absent in low checkers.

CONCLUSIONS

The current findings confirm that checkers' memory impairments are secondary to executive dysfunction, especially when ecologically valid stimuli are employed. The different accuracy patterns of high compared to low checkers between Experiments 1 and 2 allow us to make the following conclusions. In Experiment 1, we observed a novel finding with high checkers showing a robust impairment in their ability to accurately recall the state ("ON" or "OFF") of an electrical appliance. A group effect which was surprisingly not influenced by trial type (resolvable, misleading, no-probe). While superficially this appears to indicate a general impairment in WM capacity, we have highlighted a number of reasons why this is an unsatisfactory explanation. We conclude that this novel, general impairment is rather specific to the memory task (probe-2) in Experiment 1 that biased subclinical checkers toward the threatening electric on/off status of the appliances, which in turn generally interfered with multimodal bindings in the episodic buffer. In contrast, Experiment 2 revealed the expected, more differentiated pattern with a special status for misleading trials: Performance of high checkers was significantly affected on misleading trials compared to baseline trials and the strongest group difference was observed in the misleading condition.

In Experiment 2 we successfully employed a continuous confidence scale that allowed us to calculate discrepancy scores between accuracy and confidence for each participant in each condition. The main result was that while there overall strong correlations between accuracy and confidence in both groups, only the high checkers revealed a significant discrepancy in the baseline condition. Although they reached their highest performance levels in this condition, their confidence did not improve, which we interpret as supporting a metacognitive deficit that is absent in low checkers. The importance of memory and metacognitive impairments in OCD is corroborated by reports that poor memory and checking influences the severity of obsessional thinking (Purcell et al., 1998a; Park et al., 2006).

LIMITATIONS

The following limitations of our study have to be considered. Firstly, using a subclinical group always raises the issue of their relevance as an analog to a clinical group. We agree, however, with Mataix-Cols et al. (1997, 1999) that subclinical OCD groups are a valid means of determining which cognitive factors play a role in clinically defined OCD, particularly considering their reduced medication and potential for co-morbidities. We therefore expect that the pattern observed here with subclinical checkers could be more pronounced using clinical OCD patients, yet, also more variable. Secondly, despite the claim that a subclinical group provides a "purer" indication of the cognitive impairments specific to this subtype; we did not control for anxiety or depression nor did we provide an independent cognitive index of WM functioning and so cannot exclude possible group differences. Thirdly, subjects were not explicitly matched for education; however, they were selected from an undergraduate population, thus, ensuring a homogenous educational background for all participants, which is yet another advantage of a subclinical sample. Fourthly, we did not counterbalance the

keys for the forced-choice confidence responses in Experiment 1 and so cannot determine if a lateralization bias influenced participants' responding and possibly masking existing group differences.

REFERENCES

Antony, M. M., Downie, F., and Swinson, R. P. (1998). "Diagnostic issues and epidemiology in obsessive compulsive disorder," in *Obsessive–Compulsive Disorder: Theory, Research, and Treatment*, eds R. P. Swinson, M. M. Antony, S. Rachman, and M. A. Richter (New York: Guilford Press), 3–32.

Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.*, 4, 417–423.

Bannon, S., Gonsalvez, C. J., and Croft, R. J. (2008). Processing impairments in OCD: it is more than inhibition! *Behav. Res. Ther.* 46, 689–700.

Boldrini, M., Del Pace, L., Placidi, G. P., Keilp, J., Ellis, S. P., Signori, S., Placidi, G. F., and Cappa, S. F. (2005). Selective cognitive deficits in obsessive–compulsive disorder compared to panic disorder with agoraphobia. *Acta Psychiatr. Scand.* 111, 150–158.

Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W., and Sahakian, J. (2005). The neuropsychology of obsessive compulsive disorder: the importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neurosci. Biobehav. Rev.* 29, 399–419.

Ciesielski, K. T., Hamalainen, M. S., Geller, D. A., Wilhelm, S., Goldsmith, T. E., and Ahlfors, S. P. (2007). Dissociation between MEG alpha modulation and performance accuracy on visual working memory task in obsessive compulsive disorder. *Hum. Brain Mapp.* 28, 1401–1414.

Cuttler, C., and Graf, P. (2007). Subclinical compulsive checkers' prospective memory is impaired. *J. Anxiety Disord.* 21, 338–352.

Cuttler, C., and Graf, P. (2008). Subclinical checking compulsions are related to impaired prospective memory independently of depression, anxiety and distractibility. *J. Anxiety Disord.* 22, 642–654.

Cuttler, C., and Graf, P. (2009). Checking-in on the memory deficit and meta-memory deficit theories of compulsive checking. *Clin. Psychol. Rev.* 29, 393–409.

Delvenne, J. F., and Bruyer, R. (2006). A configural effect in visual short-term memory for features from different parts of an object. *Q. J. Exp. Psychol.* 59, 1567–1580.

Fougnie, D., and Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Vis. Cogn.* 17, 48–66.

Gershuny, B. S., and Sher, K. J. (1995). Compulsive checking and anxiety in a nonclinical sample – differences in cognition, behavior, personality, and affect. *J. Psychopathol. Behav. Assess.* 17, 19–38.

Greisberg, S., and McKay, D. (2003). Neuropsychology of obsessive–compulsive disorder: a review and treatment implications. *Clin. Psychol. Rev.* 23, 95–117.

Harkin, B., and Kessler, K. (2009). How checking breeds doubt: reduced performance in a simple working memory task. *Behav. Res. Ther.* 47, 504–512.

Harkin, B., and Kessler, K. (2011). How checking as a cognitive style influences working memory performance. *Appl. Cogn. Psychol.* 25, 219–228.

Harkin, B., and Kessler, K. (in press). The role of working memory in compulsive checking and OCD: a systematic classification of 64 experimental findings from 49 studies. *Clin. Psychol. Rev.*

Henderson, J. G., and Pollard, C. A. (1988). Types of obsessive compulsive disorder in a community sample. *J. Clin. Psychol.* 44, 747–752.

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

Hyun, J. S., Woodman, G. F., and Luck, S. J. (2009). The role of attention in the binding of surface features to locations. *Vis. Cogn.* 17, 10–24.

Irak, M., and Flament, M. F. (2009). Attention in sub-clinical obsessive–compulsive checkers. *J. Anxiety Disord.* 23, 320–326.

Keizer, A. W., Colzato, L. S., and Hommel, B. (2008). Integrating faces, houses, motion, and action: spontaneous binding across ventral and dorsal processing streams. *Acta Psychol. (Amst)* 127, 177–185.

Kessler, K., and Kiefer, M. (2005). Disturbing visual working memory: electrophysiological evidence for a role of the prefrontal cortex in recovery from interference. *Cereb. Cortex* 15, 1075–1087.

Mataix-Cols, D., Junque, C., Sanchez-Turet, M., Vallejo, J., Verger, K., and Barrios, M. (1999). Neuropsychological functioning in a subclinical obsessive–compulsive sample. *Biol. Psychiatry* 45, 898–904.

Mataix-Cols, D., Junqué, C., Vallejo, J., Sànchez-Turet, M., Verger, K., and Barrios, M. (1997). Hemispheric functional imbalance in a sub-clinical obsessive–compulsive sample assessed by the Continuous Performance Test, Identical Pairs version. *Psychiatry Res* 72, 115–126.

Morein-Zamir, S., Craig, K. J., Ersche, K. D., Abbott, S., Muller, U., Fineberg, N. A., Bullmore, E. T., Sahakian, B. J., and Robbins, T. W. (2010). Impaired visuospatial associative memory and attention in obsessive compulsive disorder but no evidence for differential dopaminergic modulation. *Psychopharmacology (Berl.)* 212, 357–367.

Moritz, S., Kloss, M., Jahn, H., Schick, M., and Hand, I. (2003). Impact of comorbid depressive symptoms on nonverbal memory and visuospatial performance in obsessive–compulsive disorder. *Cogn. Neuropsychiatry* 8, 261–272.

Moritz, S., and von Muehlenen, A. (2008). Investigation of an attentional bias for fear-related material in obsessive–compulsive checkers. *Depress. Anxiety* 25, 225–229.

Olley, A., Malhi, G., and Sachdev, P. (2007). Memory and executive functioning in obsessive–compulsive disorder: a selective review. *J. Affect. Disord.* 104, 15–23.

Olson, I. R., Page, K., Moore, K. S., Chatterjee, A., and Verfaellie, M. (2006). Working memory for conjunctions relies on the medial temporal lobe. *J. Neurosci.* 26, 4596–4601.

Omori, I. M., Murata, Y., Yamanishi, T., Nakaaki, S., Akechi, T., Mikuni, M., and Furukawa, T. A. (2007). The differential impact of executive attention dysfunction on episodic memory in obsessive–compulsive disorder patients with checking symptoms vs. those with washing symptoms. *J. Psychiatr. Res.* 41, 776–784.

Park, H. S., Shin, Y. W., Ha, T. H., Shin, M. S., Kim, Y. Y., Lee, Y. H., and Kwon, J. S. (2006). Effect of cognitive training focusing on organizational strategies in patients with obsessive–compulsive disorder. *Psychiatry Clin. Neurosci.* 60, 718–726.

Penades, R., Catalan, R., Andres, S., Salamero, M., and Gasto, C. (2005). Executive function and nonverbal memory in obsessive–compulsive disorder. *Psychiatry Res* 133, 81–90.

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and Performance x: Control of Visual Processing*, eds H. Bouma and D. G. Bouwhuis (Hillsdale, NJ: Erlbaum), 531–556.

Prabhakaran, V., Narayanan, K., Zhao, Z., and Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3, 85–90.

Purcell, R., Maruff, P., Kyrios, M., and Pantelis, C. (1998a). Cognitive deficits in obsessive–compulsive disorder on tests of frontal-striatal function. *Biol. Psychiatry* 43, 348–357.

Purcell, R., Maruff, P., Kyrios, M., and Pantelis, C. (1998b). Neuropsychological deficits in obsessive–compulsive disorder: a comparison with unipolar depression, panic disorder, and normal controls. *Arch. Gen. Psychiatry* 55, 415–423.

Rachman, S. (2002). A cognitive theory of compulsive checking. *Behav. Res. Ther.* 40, 625–639.

Rachman, S., and Shafran, R. (1998). "Cognitive behavioural features of obsessive–compulsive disorder," in *Obsessive–Compulsive Disorder: Theory, Research and Treatment*, eds R. Swinson, M. Antony, S. Rachman, and M. Richter (New York: Guilford Publications), 51–78.

Radomsky, A. S., and Alcolado, G. M. (2010). Don't even think about checking: mental checking causes memory distrust. *J. Behav. Ther. Exp. Psychiatry.* 41, 345–351.

Rasmussen, S. A., and Eisen, J. L. (1988). Clinical and epidemiologic findings of significance to neuropharmacologic trials in OCD. *Psychopharmacol. Bull.* 24, 466–470.

Rubenstein, C. S., Peynircioglu, Z. F., Chambless, D. L., and Pigott, T. A. (1993). Memory in sub-clinical obsessive–compulsive checkers. *Behav. Res. Ther.* 31, 759–765.

Savage, C. R., Baer, L., Keuthen, N. J., Brown, H. D., Rauch, S. L., and Jenike, M. A. (1999). Organizational strategies mediate nonverbal memory impairment in obsessive–compulsive disorder. *Biol. Psychiatry* 45, 905–916.

Savage, C. R., Deckersbach, T., Wilhelm, S., Rauch, S. L., Baer, L., Reid, T., and Jenike, M. A. (2000). Strategic processing and episodic memory impairment in obsessive compulsive

disorder. *Neuropsychology* 14, 141–151.

Sher, K. J., Mann, B., and Frost, R. O. (1984). Cognitive dysfunction in compulsive checkers – further explorations. *Behav. Res. Ther.* 22, 493–502.

Stein, M. B., Forde, D. R., Anderson, G., and Walker, J. R. (1997). Obsessive–compulsive disorder in the community: an epidemiologic survey with clinical reappraisal. *Am. J. Psychiatry* 154, 1120–1126.

Thordarson, D. S., Radomsky, A. S., Rachman, S., Shafran, R., Sawchuk, C. N., and Ralph Hakstian, A. (2004). The Vancouver Obsessional Compulsive Inventory (VOCI). *Behav. Res. Ther.* 42, 1289–1314.

Tuna, S., Tekcan, A. I., and Topcuoglu, V. (2005). Memory and metamemory in obsessive–compulsive disorder. *Behav. Res. Ther.* 43, 15–27.

van den Hout, M., and Kindt, M. (2003). Phenomenological validity of an OCD-memory model and the remember/know distinction. *Behav. Res. Ther.* 41, 369–378.

van der Wee, N. J., Ramsey, N. F., Jansma, J. M., Denys, D. A., van Megen, H. J., Westenberg, H. M., and Kahn, R. S. (2003). Spatial working memory deficits in obsessive compulsive disorder are associated with excessive engagement of the medial frontal cortex. *Neuroimage* 20, 2271–2280.

van der Wee, N. J., Ramsey, N. F., van Megen, H. J., Denys, D., Westenberg, H. G., and Kahn, R. S. (2007). Spatial working memory in obsessive–compulsive disorder improves with clinical response: a functional MRI study. *Eur. Neuropsychopharmacol.* 17, 16–23.

Wheeler, M. E., and Treisman, A. M. (2002). Binding in short-term visual memory. *J. Exp. Psychol. Gen.* 131, 48–64.

Zielinski, C. M., Taylor, M. A., and Juzwin, K. R. (1991). Neuropsychological deficits in obsessive–compulsive disorder. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 4, 110–126.

Zitterl, W., Urban, C., Linzmayer, L., Aigner, M., Demal, U., Semler, B., and Zitterl-Eglseer, K. (2001). Memory deficits in patients with DSM-IV obsessive–compulsive disorder. *Psychopathology* 34, 113–117.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.