Relevance vs. Valence: Directed forgetting in sub-clinical checkers and washers

Metehan Iraķ1
Dicle Çapan

1. Introduction

Obsessive-compulsive disorder (OCD) is a psychiatric disorder characterized by recurrent obsessions and compulsions that are perceived as irrational, and cause significant impairment and distress. Difficulty inhibiting irrelevant information (e.g. obsessive thoughts) is a key factor of OCD (Cohen, Lachenmeyer & Springer, 2003). In general OCD patients’ attention span is mostly allocated to threat (or anxiety) relevant stimuli related to their concerns and symptoms, therefore these individuals have limited capacity to selectively attend to relevant information while simultaneously ignoring irrelevant information (van den Heuvel, Veltman, Groenewegen et al., 2005).

There is little evidence for global impairment of attention in OCD patients, with no differences from healthy controls found for reaction time alertness, speed of information processing, attention span, or sustained attention (Kuelz, Hohagen, & Voderholzer, 2004). In adults, selective attention biases have been reported in OCD and possibly co-morbid conditions, such as anxiety and affective disorders. Several studies found these biases to be directly related to threat-relevant stimuli (e.g., Diniz et al., 2004; Hartston & Swerdlow, 1999; Summerfeldt & Endler, 1998; Tavares, Drevets, & Sahakian, 2003). Foa and McNally’s (1986) study addressing this issue used a dichotic listening task to demonstrate that threat-relevant stimuli were perceived more easily than neutral stimuli. Thus, there is evidence for an abnormal processing bias towards OCD-relevant stimuli on paradigms such as dot-probe and directed forgetting (Diniz et al., 2004; Moritz et al., 2004), and using a modified Stroop task, Foa et al. (1993) found that washers were more attentive to contamination words than controls, and in general OCD patients showed more interference than controls for threat-related words.

1 E-Mail: metehan.irak@eas.bahcesehir.edu.tr
Aspects of the behavior seen in people with OCD (e.g., checking) are suggestive of memory problems (e.g., encoding). However, studies on reality monitoring and memory for self-actions in OCD patients have failed to find evidence of memory impairment (e.g., Constans, Foa, Franklin, & Mathews, 1995; Hermans, Martens, De Cort, Pieters, & Eelen, 2003). In addition, previous results on the relationship between compulsive symptoms and memory problems have been inconsistent. A few investigators have reported that OCD patients showed impaired memory compared to normal controls (Savage et al., 2000; Tallis, Pratt, & Jamani, 1999; Tuna, Tekcan, & Topcuoglu, 2005; Zitterl et al., 2001). In a study by Singh et al. (2003), OCD subjects showed low performance than controls on verbal and visuospatial working memory tasks. On the other hand, other studies found that OCD patients showed better recall performance than controls for threat-relevant information (e.g., Constans, Foa, Franklin, & Mathews, 1995; Radomsky & Rachman, 1999; Radomsky, Rachman, & Hammond, 2001). In one study, OCD patients were similar to controls in their recall of neutral objects, but showed higher performance at recalling dangerous stimuli (Tolin et al., 2001), and in another OCD checkers were impaired in the recall of their own actions, but less in recall of other information (Ecker & Engelkamp, 1995). In Wood and colleagues’ (2002) meta-analysis on checking and memory, checkers were impaired on verbal free recall, verbal cued recall, and recall of actions, but not on recognition; this meta-analysis has been criticized because data were collapsed across both OCD- and sub-clinical participants.

It remains unclear whether emotional valence (personal emotional significance) or personal relevance (related with symptom) is the key factor for memory bias in obsessive-compulsive disorder. Thus, the first goal of the present study was to provide additional data in terms of memory bias in sub-clinical checkers and sub-clinical washers in comparison to control subjects. A directed forgetting procedure was used in the study because it has been frequently used in studies with sub-clinical groups and as well as patients with anxiety (e.g., McNally, Metzger, Lasko, Clancy, & Pitman, 1998; Tolin, Hamlin, & Foa, 2002; Wilhelm, McNally, Baer, & Florin, 1996). The directed forgetting task provides an especially appropriate test of memory bias. In general, there are two common methods of directed forgetting, namely; the item method and the list method. In the present study, the item method was used. In the item method, each word (or item) is followed by an instruction either to remember (R) or to forget (F) that word. Related literature indicated that the directed forgetting is related with selective rehearsal process. Participants generally remember more R words than F words. Because as expected, R words receive more rehearsal than F words, which should not be rehearsed once the participants are informed that these items are not going to be presented in the next steps (Johnson, 1994; MacLeod, 1998). Thus, item method directed forgetting might provide valuable information as to how sub-clinical washers and checkers, compared to control
participants, process information that needs to be ignored after the instruction. Studies on cognitive biases generally measured memory performance for negative information or stimulus. However, in the current study participants were presented with neutral words along with symptom-relevant (washing and checking) words. It is first expected that both sub-clinic washers and checkers would show a directed forgetting effect, specifically they would remember more F words compared to controls, and that this effect would occur only for their symptom-related words (not for neutral words). For instance, it is expected that sub-clinical washers would remember more washing F items compared to sub-clinical checkers and controls, and vice versa. We also predicted that both personal relevance and valence would contribute to impaired directed forgetting in sub-clinical washers and checkers, compared to controls.

2. Method

2.1 Participants

Volunteer university students (n=242) completed the Maudsley Obsessive Compulsive Inventory (MOCI; Erol & Savasır, 1988; Hodgson & Rachman, 1977; Sternberger & Burns, 1990). Following previous research (e.g. MacDonald, Antony, Macleod, & Richter, 1997; Rubenstein, Peynircioglu, Chambliss, & Pigott 1993), those who received a score of 4 or more on the checking subscale were classified as checkers, those who received a score of 4 or more on the washing subscale were classified as washers, and those with checking and washing scores of 0 or 1 were classified as controls. A total of 55 subjects between the ages of 18 and 25 participated in the study; 13 checkers (4 males, 9 females), 27 washers (11 males, 16 females), and 15 controls (8 males, 7 females). All participants denied past diagnosis or treatment of any neurological or psychiatric disorders.

3. Materials

3.1 Maudsley Obsessional-Compulsive Inventory (MOCI)

MOCI (Hodgson & Rachman, 1977) consists of 30 Yes–No statements measuring the severity of OCD symptoms on four subscales (cleaning, checking, doubting, and slowness). The adaptation study of MOCI in Turkish population was completed by Erol and Savasır (1988). The Turkish version of the MOCI has adequate reliability and validity values. Internal consistency analysis indicated that the Turkish form of MOCI’s Cronbach’s Alpha values were 0.81. These values ranged from 0.31 to 0.71 for subscales.
3.2 State-Trait Anxiety Inventory (STAI)

STAI (Spielberger, Gorsuch, Luschene, Vagg, & Jacobs, 1983) consists of 40 Likert-type items measuring an individual’s anxiety at that time and in general. Each item is rated on a 4-point Likert scale; almost never (1) to almost always (4). STAI scores range from 20 to 80 points for each subscale. Higher scores indicate more anxiety. The Turkish version has been shown to be reliable and valid (Öner & LeCompte, 1983).

3.3 Beck Depression Inventory (BDI)

BDI (Beck et al., 1961) is a self-report measure assessing the level of depressive symptomatology. It is composed of 21 questions or items, each with four possible responses. Each response is assigned a score ranging from zero to three, indicating the severity of the symptom. The Turkish version has been shown to be reliable and valid (Hisli, 1988).

3.4 The word list

A list of 54 words, containing an equal number of (18) washing, checking, and neutral words, was constructed. The list was constructed on the basis of the earlier work (e.g., Irak & Flament, 2009; Radomsky & Rachman, 1999). There were an additional four neutral buffer words (two in the beginning and two at the end of the study list) to minimize primacy and recency effects in recall.

4. Procedure

All participants were tested individually in a special testing room and each participant was seated in front of a computer screen. Each administration period consisted of 3 steps namely, learning (study) session, a free recall test, and a recognition (Yes/No question) test. Each participant informed that this would be a memory task where they will be presented with a list of words and then asked to remember as many words from the study list as possible. They were also told that each word will be presented for 2 s on the screen followed by a specific instruction regarding that word; an “R” would indicate that they would need to pay attention and remember that word, and an “F” meant that that word could be ignored, because they would not be asked to remember those words. Each word was presented for 2 s, followed by the instruction to remember or to forget, which remained on the screen for 3 s. Participants saw a list of 58 words consisting of 54 target words and 4 buffer words. Half of the target words were followed by an instruction to remember and the other half by an instruction to forget. The buffer words were always
followed by an R instruction, and were not included in the statistical analyses. Across participants, each word was followed by an instruction to remember or to forget equally often.

After the learning session, participants were given a blank A4 paper and asked to remember (recall) as many of the presented words as possible regardless of the instructions (R vs. F) related with the words. Participants were given 15 min for the free recall task. After the free recall test, a recognition test (Yes/No question) was given. They were presented with a list of words containing the studied words along with an equal number of unstudied (distractor) words. The distractor words also consisted of checking, washing, and neutral words (18 each, 54 in total). Subjects were asked to indicate, for each word, whether they saw it in the previous session or not. There was no time limit during the recognition test. Each experimental session took approximately 25-30 min per participants.

5. Design

A 3x3x2 mixed factorial design was carried out with the group (checkers, washers, and controls) as a between-subjects variable, and word type (checking, washing, and neutral) and instruction (remember vs. forget) as within-subjects variables.

6. Results

6.1 Background variables

As shown on Table 1, One-Way ANOVA showed that there was no significant difference between the sub-clinical checkers, washers, and controls groups in terms of age $F(2,54) = 1.30, p = .281$, education level $F(2,53) = .949, p = .394$, BDI total score $F(2,54) = 2.71, p = .076$, and State Anxiety $F(2,54) = .80, p = .473$. However; significant differences were observed in terms of Trait Anxiety $F(2,54) = 3.74, p < .05$ and MOCI total score $F(2,54) = 32.40, p < .001$. Post-hoc Tukey analysis showed that Trait Anxiety total score for checkers ($M = 49.38, SD = 3.59$) was significantly higher than controls ($M = 44.27, SD = 4.64$). For MOCI total score, the differences between washers ($M = 11.89, SD = 4.08$), checkers ($M = 15.92, SD = 3.01$), and controls ($M = 5.87, SD = 1.77$) were significant.

| Table 1. Means and standard deviations of sub-clinical washer, sub-clinical checker, and control groups on background variables. |
|-----------------|-----------------|-----------------|
| Washers         | Checkers        | Control         |
| Age             | Education       | BDI             |
| Trait Anxiety   | MOCI            |                 |
Relevance vs. Valence: Directed forgetting in sub-clinical checkers and washers

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>13.4</td>
<td>1.00</td>
<td>12.69</td>
<td>0.95</td>
<td>12.67</td>
<td>0.90</td>
</tr>
<tr>
<td>Age</td>
<td>20.1</td>
<td>1.47</td>
<td>20.08</td>
<td>1.32</td>
<td>20.33</td>
<td>1.50</td>
</tr>
<tr>
<td>BDI Total</td>
<td>8.26</td>
<td>4.51</td>
<td>13.23</td>
<td>9.45</td>
<td>9.00</td>
<td>6.31</td>
</tr>
<tr>
<td>State Anxiety</td>
<td>41.0</td>
<td>3.33</td>
<td>42.38</td>
<td>4.68</td>
<td>40.27</td>
<td>6.26</td>
</tr>
<tr>
<td>Trait Anxiety</td>
<td>47.7</td>
<td>5.65</td>
<td>49.38</td>
<td>3.59</td>
<td>44.27</td>
<td>4.64</td>
</tr>
<tr>
<td>MOCI Total</td>
<td>11.9</td>
<td>4.08</td>
<td>15.92</td>
<td>3.01</td>
<td>5.87</td>
<td>1.77</td>
</tr>
</tbody>
</table>

6.2 Recall performance

Paired sample t-tests indicated that there was significant directed forgetting effect for all three groups. All subjects recalled more R words than F words. Mean and standard deviations for recall performance were presented at Table 2 and for recognition performance were presented at Table 3. In order to further investigate the effect of group on the magnitude of the directed forgetting effect, separate 3 (group) x 3 (word type) mixed design ANOVAs for recall of R and F words were carried out. For recall performance of F words, the word type main effect was significant, $F(2,104) = 11.33, p < .01, \eta^2 = .18$. Pairwise comparison indicated that subjects recalled more washing ($M = 3.53$) and checking ($M = 3.16$) items than neutral words ($M = 2.18$). On the other hand, neither group main effect nor interaction effect was significant ($p > .421$). For recall performance of R words, the word type main effect was significant, $F(2,104) = 3.56, p < .05, \eta^2 = .06$. Pairwise comparison indicated that subjects recalled more washing ($M = 1.52$) and checking ($M = 1.42$) items than neutral items ($M = .96$). On the other hand, neither group main effect nor interaction effect was significant ($p > .649$). Consequently, results on recall performance showed significant directed forgetting effect for three groups. Sub-clinical subjects recalled more F words than the control subjects.

The bias towards symptom-related items may be linked to anxiety rather than the high washing or checking score per se. Even though we observed significant differences between checkers and controls, in terms of trait anxiety score only, to address the role of anxiety as a confounding variable, Pearson correlation between anxiety scores and recall of Rand F words was carried out. Results were not significant ($p \geq .074$). So, at least for the current study, level of anxiety does not seem to be related to recall performance of the participants.

7. Recognition
Hit rates (percentage of ‘yes’ responses to studied items) are not adequate measure of recognition. For this reason sensitivity and response criterion (bias) was calculated.

7.1 Hit rates

There was a significant directed forgetting effect for three groups. Paired sample t-test results indicated that, hit rates for R words (.34 vs .23 for washers, .31 vs .22 for checkers, and .32 vs .24 for controls) are significantly higher than F words ($t_{wahers} = 6.34$, $df = 26$, $p < .001$; $t_{checkers} = 6.34$, $df = 26$, $p < .001$; $t_{controls} = 3.21$, $df = 14$, $p < .01$). Moreover, the effect was substantially bigger for washers (.11) compared to checkers (.09) and controls (.08).

Two 3 (groups) x 3 (word type) mixed design ANOVAs were carried out on hit rates for R and F words, separately. The group main effect and also the word type and group interaction effect for R words, was not significant ($p = .363$ and $p = .494$, respectively). On the other hand word type main effect was significant, $F(2,104) = 16.53$, $p < .001$, $\eta^2 = .24$. Pairwise comparisons with Bonferroni corrections showed that the hit rates for checking items (.74) was significantly higher than washing items (.59), and neutral items (.63). For F words, the word type main effect was significant, $F(2,104) = 19.15$, $p < .001$, $\eta^2 = .27$. Pairwise comparisons with Bonferroni corrections showed that the hit rates for checking items (.57) was significantly higher than washing items (.38), and neutral items (.44). On the other hand, the group main effect and also the word type and group interaction effect for F words, was not significant ($p = .794$ and $p = .324$, respectively).

7.2 False alarms

Two 3 (groups) x 3 (word type) mixed design ANOVAs were carried out on false alarm rates. Results showed that the word type main effect was significant, $F(2,104) = 5.72$, $p < .01$, $\eta^2 = .10$. Pairwise comparisons with Bonferroni corrections showed that false alarm rates to neutral items (1.37) were significantly higher than washing items (1.18) and checking items (.73). On the other hand neither the group main effect ($p = .053$) nor the interaction effect ($p = .057$) was significant.

The sensitivity in recognition performance was also calculated. Sensitivity in recognition refers to how well participants can differentiate between studied (old) and non-studied (new) items. The $A'$ was carried out as a measure of sensitivity, because this measure does not depend on the assumptions of normality and homogeneity of variance. The $A'$ can range between 0 and 1, with .5 indicating chance level performance (Pollack
Results indicated that all three groups—namely sub-clinical checkers, sub-clinical washers, and controls—showed discrimination above chance (.92, .90, and .88, respectively). On the other hand, a 3 (group) x 3 (word type) mixed design ANOVA indicated that neither main effects nor the interaction effect were significant ($p \geq .202$).

### Table 2. Means (and standard deviations) for recall performance according to group status, word type, and instruction.

<table>
<thead>
<tr>
<th></th>
<th>Washers</th>
<th>Checkers</th>
<th>Control</th>
<th>Washers</th>
<th>Checkers</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Remember</td>
<td>Forget</td>
<td></td>
<td>Remember</td>
<td>Forget</td>
<td></td>
</tr>
<tr>
<td>Checking</td>
<td>1.69 (.95)</td>
<td>2.44 (1.89)</td>
<td>2.4 (1.64)</td>
<td>0.92 (1.09)</td>
<td>1.02 (1.32)</td>
<td>0.96 (1.07)</td>
</tr>
<tr>
<td>Washing</td>
<td>3.85 (1.0)</td>
<td>3.67 (1.52)</td>
<td>3.07 (1.39)</td>
<td>2.07 (1.62)</td>
<td>1.56 (1.26)</td>
<td>0.923 (1.12)</td>
</tr>
<tr>
<td>Neutral</td>
<td>3.48 (1.5)</td>
<td>3.08 (1.5)</td>
<td>2.93 (1.87)</td>
<td>1.37 (1.69)</td>
<td>1.69 (1.44)</td>
<td>1.2 (1.61)</td>
</tr>
</tbody>
</table>

### Table 3. Means (and standard deviations) for recognition performance according to group status, word type, and instruction.

<table>
<thead>
<tr>
<th></th>
<th>Washers</th>
<th>Checkers</th>
<th>Control</th>
<th>Washers</th>
<th>Checkers</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Remember</td>
<td>Forget</td>
<td></td>
<td>Remember</td>
<td>Forget</td>
<td></td>
</tr>
<tr>
<td>Checkin g</td>
<td>5.6 (1.18)</td>
<td>4.92 (1.19)</td>
<td>5.44 (1.91)</td>
<td>3.67 (2.77)</td>
<td>3.70 (1.88)</td>
<td>2.85 (2.03)</td>
</tr>
<tr>
<td>Washing</td>
<td>6.96 (1.02)</td>
<td>6.6 (1.96)</td>
<td>6.33 (1.63)</td>
<td>5.41 (1.85)</td>
<td>4.69 (1.93)</td>
<td>5.2 (2.43)</td>
</tr>
<tr>
<td>Neutral</td>
<td>6.15 (1.92)</td>
<td>5.46 (1.45)</td>
<td>5.4 (1.76)</td>
<td>3.81 (1.92)</td>
<td>4.23 (1.92)</td>
<td>3.87 (2.36)</td>
</tr>
</tbody>
</table>

### 8. Discussion

Our results indicated standard directed forgetting effect in sub-clinical washers, sub-clinical checkers, and control participants; the three groups recalled more of the R items than F items. The first important finding of the present study was the difference between R and F words which was significantly larger for the washers and checkers compared to controls. A second important finding was that sub-clinical participants showed a significant memory bias for symptom-related words. In detail, the sub-clinical washers and checkers recalled more of the symptom-relevant items they were asked to forget than the control group. For instance, checkers showed increased memory bias for checking items than washers, similarly washers showed increased memory bias for washing items than checkers. The third important finding was participants showed similar
levels of bias in recognition memory for the F items. It was found that hit rates for R words were significantly higher than F words. On the other hand, results indicated that false alarm rate for neutral items was significantly higher than washing and checking items.

The memory bias in this study concurs with previous work regarding OCD. The present study showed that sub-clinical washers and checkers had difficulty forgetting symptom–related items which they were briefly exposed to. Importance of this result is that this bias was observed for the items that participants were asked to forget but not for items they were asked to remember. Specifically, sub-clinical checkers showed greater memory bias for checking items compared to washing items and vice versa. This memory bias is especially consisted with the characteristics of the item-method procedure. There is common agreement in the related literature that selective rehearsal processing leads to difference between R and F items (MacLeod, 1998; Tekcan, Tas, Topcuoglu, & Yucel, 2008). In other words, R items are remembered better than F items, because when R instruction is given participants continue to rehearse the R items, whereas participants do not further process F items when the F instruction is given. Our results indicated that the sub-clinical washers and checkers continued to rehearse the F items, especially their own symptom-relevant items. So, it can be concluded that sub-clinical washers and checkers were unable to stop the rehearsal process they initiated when they first saw a symptom-related item. This conclusion also concurs previous studies on sub-clinical checkers (e.g. Irak & Flament, 2009). Irak and Flament (2009) examined attention bias under different attention paradigms. It was found that, even though instructions and tasks were different in each attention paradigm namely, focused, divided and passive attention, the checkers showed similar attention biases in all paradigms. Authors concluded that obsessive-compulsive checkers showed an attention bias and this bias is independent from the type of attention. The authors also suggest that threat-relevant stimuli (relevance) in the environment of obsessive-compulsive individuals are always foremost or dominant, since, in OC checkers, the attention process ignores other factors such as conditions, instructions, or tasks.

In another study, Tolin, Hamlin, and Foa (2002) used a directed forgetting paradigm. No significant difference was found between OCD patients, anxious patients, and non-anxious controls in terms of recall performance. On the other hand, OCD patients showed greater impaired forgetting for OCD-relevant words than did non-anxious and anxious controls, during recognition test. The authors concluded that the valence of items did not emerge as a contributing factor to impaired forgetting. However, the results of the present study indicated that participants showed a memory bias for both symptom-related (either checking or washing) and neutral words, suggesting that symptom-relevance rather than the valence of the items might have been the important factor. As mentioned before,
sub-clinical checkers showed greater memory bias for checking items compared to washing items and vice versa. McNally (1995) hypothesized that when compared to positive items or information, OCD patients remember more negative information because they use selective encoding process for it. Studies on dichotic listening (e.g., Foa & McNally, 1986) and emotional Stroop task (e.g., Foa, Ilai, McCarthy, Shoyer, & Murdock, 1993; Lavy, van Oppen, & van den Hout, 1994) showed that obsession-relevant items disproportionately capture resources in people with OCD. Besides, Foa and Kozak’s (1986) cognitive network theories of anxiety disorders might better explain our results. They argued that cognitive fear networks contain both negative (e.g., threat) and also positive (e.g., escape) information. Thus, positive and negative components reveal similar biases in information processing system.

However this interpretation should be treated with caution because of limitations in the present study. First, this study was done in a community population, and not in a clinical sample. The second point is, even though it is an experimental study, low sample size may have rendered subtle differences on the results. Although there has been some documentation of OCD patients having specific deficits in memory and attention, the role of these basic cognitive processes in the phenomenology and treatment of OCD is still unclear. Our results provide additional information in terms of memory bias in sub-clinical individuals. Results of the present study concur with previous studies that have found symptom-relevance rather than the valence of items might have been the important factor in determining performance. The information processing approach may result in the development of improved or novel assessment procedures and treatments for OCD. It would be fruitful for the future studies to confirm these results using a clinical OCD group.

References


Relevance vs. Valence: Directed forgetting in sub-clinical checkers and washers


