NOT JUST RIGHT EXPERIENCES AS IRONIC RESULT OF PERSEVERATIVE CHECKING

Eva Anna Maria van Dis and Marcel A. van den Hout

Abstract

Objective: Patients with obsessive-compulsive disorder (OCD) typically report to have "not just right experiences" (NJRE). Up till now it is unclear which behavioral OCD features may give rise to NJREs. We used an induced checking paradigm to experimentally study whether perseverative checking elicits NJREs.

Method: Two experiments separately tested this hypothesis among n = 48 (Experiment I) and n = 55 (Experiment II) healthy participants. We used a virtual checking task, in which participants either checked gas stoves or light bulbs. All participants started and ended with a trial in which they checked a gas stove (i.e., pre- and post-test). In between, the experimental group repeatedly checked the gas stove (i.e., relevant checking), while the control group repeatedly checked light bulbs (i.e., irrelevant checking). At pre- and post-test, all participants answered questions about the corresponding gas stove checking trial (i.e., memory confidence, vividness and details) and rated their level of NJRE.

Results: In line with previous research, both experiments showed that relevant checking (as opposed to irrelevant checking) resulted in reduced memory confidence and less vivid and detailed recollections of the last checking trial. Most importantly, both studies found a medium effect for increased NJREs after relevant checking compared to irrelevant checking. Since not all results reached statistical significance in the individual studies, we combined the findings in a meta-analysis that clearly confirmed our hypotheses.

Conclusions: Data of Experiment I and II strongly suggest that repeated checking results in NJRE.

Eva Anna Maria van Dis and Marcel A. van den Hout
Utrecht University, the Netherlands

Corresponding author
E-mail: e.a.m.vandis@uu.nl

"Not just right experiences" (NJREs) refer to the uncomfortable sensations of things not being just right (Coles et al. 2003). Examples are that while one's hands have been washed they somehow do not *feel* clean or the act of closing a door feels uncompleted. Individuals diagnosed with Obsessive-Compulsive Disorder (OCD) typically report having such sensations of incompleteness. Several studies show that OCD patients report more NJREs compared to healthy controls (e.g., Ghisi et al. 2010, Ecker et al. 2014) and OCD symptoms seem to be positively related to the frequency of reported NRJEs (Ravid et al. 2014, Sica et al. 2012, Moretz and McKay 2009). These studies are correlational and therefore silent about causality.

Still, some studies did explore causality issues by examining consequences of experimentally induced NJREs. For example, Coles et al. (2005) attempted to elicit NJREs by bringing participants into an office that included eight experimental stimuli (e.g., a dishevelled bookshelf, a dirty sink, and a chair missing one armrest). Participants payed their attention to all these stimuli and rated for each stimulus (1) their discomfort, (2) whether they wanted to change or move things, and (3) whether they thought something bad might happen. This paradigm has been used in several studies and showed that induced NJREs breed subsequent OCD-like phenomena such as increased washing duration (Cougle et al. 2011) and the urge to check a gas stove (Cougle et al. 2013). Interestingly, the latter effects were only found in students with elevated checking symptoms and dispositional incompleteness scores, which suggests that the NJRE induction especially affects a specific subset of the population.

Even though NJREs may stimulate OCD-like responses, it does not rule out that, the other way around, OCD symptoms fuel NJREs. To our knowledge, only Mancini et al. (2008) addressed this hypothesis and tested in two experiments whether feelings of guilt – which are often prominent in OCD – evoke NJREs. In the first experiment, participants were classified as low or high in trait guilt. All participants received an affect manipulation in which they were asked to describe in detail either a neutral or a guilt-related autobiographical life event. Next, participants were instructed to rearrange 15 cards in the order that satisfied them mostly. Results showed that only high trait-guilt individuals in the guilt induction group reported higher state NJRE after the ordering task. This finding was replicated in a second experiment, in which Mancini et al. (2008; Exp. 2) also included a third group of subjects that was asked to write about an autobiographical event in which they felt they were victims of wrong-doing (i.e., victim induction group). Again, the high trait-guilt group that received the guilt induction reported higher NJRE compared to both the neutral and victim groups. This suggests, then, that NJREs not only serve to promote other OCD symptoms, but that, reciprocally OCD features (here: guilt) in its turn might exuberate NJREs. However, since guilt is a prominent *mental* feature of OCD, we are not yet sure whether this relationship to NJREs can

be generalized to *behavioral* characteristics of OCD.

What behavioral OCD phenomena could possibly lead to NJREs? Cutting through a range of clinical manifestations of the disorder, repeated checking is often recognized as most notably symptom (Tallis 1995). Repeated checking has some curious consequences. First, although repeated checking may be an effort to decrease doubt (e.g., whether the TV is switched off), it ironically seems to increase uncertainty (Rachman 2002). This hypothesis is supported by a long series of experimental studies using an "induced checking paradigm" in which individuals repeatedly checked a virtual gas stove (e.g., van den Hout and Kindt 2003 a-b, van den Hout and Kindt 2004, Radomsky et al. 2006, Boschen and Vuksanovic 2007, Giele et al. 2013, Dek et al. 2015). After repeated gas stove checking, individuals typically reported to be less confident about their performance and generally tended to rate their memory of the last task as less vivid and detailed. Hence, these studies revealed that repeated checking ironically increases uncertainty about memory.

Second, repeated checking could also result in feelings of dissociation. One study, for example, showed that OCD patients reported more dissociative feelings during moments of compulsive perseveration, compared to non-perseverative moments (e.g., "during the last minutes there were moments when things seemed unreal, as if I was dreaming") (Giele et al. 2016). This finding corresponds with several studies that demonstrated a positive relation between perseveration and dissociative experiences (e.g., van den Hout et al. 2008, van den Hout et al. 2009). These studies focused rather on perception (instead of memory) and found that participants who repeatedly stared at gas stoves (even in relatively short intervals of staring) experienced uncertainty about their perception and reported dissociative feelings. Reed (1985) reported on dissociative ambivalence in OCD patients and reported some quotes of patients ("I know that I have done it, but the memory is not clear somehow", or "I remember doing it in a way, but it's all fuzzy" or "It's as though the memory is there, but it's not *definite* enough"). Van den Hout and Kindt (2004; Exp. 6) used these quotes in an induced checking paradigm and asked participants to endorse the above-mentioned quotes after repeated checking. In the relevant checking condition, participants recognized themselves more in the quotes compared to participants in the control condition. Hence, it seems that the experiential nature of the dissociative feelings in patients is to some extent comparable to the reported feelings in the induced checking paradigm.

Both processes of increased uncertainty and dissociative *ambivalence* appears comparable to the ambivalence inherent to NJREs: The *feeling* of things being "not just right" is precisely peculiar, because the individual simultaneously realizes that the feeling is at odds with some objective state of affairs (e.g., a door *is* closed after you closed it). Since repeated checking gives rise to memory uncertainty and dissociative feelings, we expect that perseverative checking also provokes NJREs. We tested this hypothesis in two experiments in which we used an induced checking paradigm similar to the one developed by van den Hout and Kindt (2003a-b).

Experiment I

Method

Participants

The study sample included forty-eight students recruited at Utrecht University in January 2015 (*M* age = 22, *SD* = 3.7; 60% female). Participants received course credit or a small financial compensation for their participation. All participants provided written informed consent prior to their participation.

Procedure and experimental task

Participants were tested in dimly lit and sound-attenuated cubicles and started with a virtual checking task developed by van den Hout and Kindt (2003a-b). The checking task consisted of 22 trials in which participants either checked a six-burner stove or a set of six light bulbs.

Each trial started with a 4-s presentation of a schematic diagram that randomly indicated which three gas rings or lights had to be turned on. Next, the burner stove or light bulbs were presented and participants were asked to turn on the corresponding gas rings or lights by moving the mouse cursor. Then, participants were told to turn off the gas rings/lights. Each trial ended with a task to check with the mouse cursor whether the gas rings/lights were really turned off.

Before the experiment started, participants practiced with two trials in which they were asked to turn on and off six gas rings and six light bulbs, to become familiar with the checking task. The practice trials ended with false feedback information in which participants were told that only three out of six gas rings/lights bulbs were properly turned off. Participants received no feedback during the experiment. The actual experiment started with a burner stove trial and all participants completed a questionnaire about this first checking trial (see below). After this pre-test, subjects were either randomly allocated to the relevant checking group, or to the irrelevant checking group. Participants in the relevant checking group continued with a series of 20 burner stove trials, while participants in the irrelevant checking group completed 20 light bulbs trials. The experiment ended with a burner stove trial similar to the pre-test. Again, all participants filled in the questionnaire about this final checking trial (i.e., post-test).

Measurements

Memory accuracy

Participants received a schematic representation of the six gas rings and were asked to indicate which three rings they were instructed to check in the last checking trial (cf. van den Hout and Kindt 2003a-b).

Meta-memory assessment

Memory confidence was assessed by the question how certain participants were of the correctness of the accuracy question. Participants indicated their answers on a visual analogue scale (VAS) ranging from 0

(absolutely not confident) to 100 (absolutely confident). Memory vividness and detail of the last checking trial were separately rated on two VASs (0 = not vivid, 100 = extremely vivid; 0 = not detailed, 100 = extremely detailed).

NJRE scale

'Not just right experience' was assessed with a self-constructed scale that was tailored to the virtual checking task. We composed the following items that were answered on a VAS ranging from 0 (completely not applicable to me) to 100 (completely applicable to me):

1. It feels like operating the knobs did not go well the last time
2. While doing the task the last time, it felt like I had to do it in a certain way without knowing why
3. Thinking of the last time turning the knobs, it feels like my memory does not exactly fit what happened
4. Somehow it feels like the last knob task is not completely finished

Reliability analysis of the scale at post-test, confirmed a good internal consistency of the NRJE scale (Cronbrach's α = .72). However, item 2 had a relatively small item-rest correlation (r = .20) compared to other item-rest correlations (all above .47). Given that the alpha increased to .82 after deleting item 2, we decided to omit this item from the scale.

Results

We replaced 7 outliers with $M \pm 2.5$ SD. Results of memory accuracy could not be reported due to a technical flaw.

Meta-memory assessment

We used three separate 2 × 2 mixed model ANOVAs with memory confidence, vividness and detail as dependent variables to assess effects of Condition (between group factor: relevant vs. irrelevant checking) and Time (within group factor: pre vs. post) and the Time × Condition interaction.

As can be seen from **figure 1** (left upper panel), confidence in memory differed over Time, $F(1, 46)$ =

17.26, $p < .001$, $\eta_p^2 = .27$, but no statistical effects were found for Condition, $F(1, 46) < 1$, $p = .525$, $\eta_p^2 = .01$, and the interaction between Time × Condition, $F(1, 46) < 1$, $p = .443$, $\eta_p^2 = .01$.

Memory vividness differed statistically over Time, $F(1, 46) = 22.53$, $p < .001$, $\eta_p^2 = .33$, and Condition, $F(1, 46) = 5.59$, $p = .023$, $\eta_p^2 = .11$ (see **figure 1**, right upper panel). Importantly, the interaction effect between Time × Condition reached significance, $F(1, 46) = 18.55$, $p < .001$, $\eta_p^2 = .29$. Simple main effects analyses, using paired samples t-tests, showed that vividness ratings decreased over time in the relevant checking group, $t(23) = 5.33$, $p < .001$, but not in the irrelevant checking group, $t(23) < 1$, $p = .681$.

Ratings for memory detail differed statistically over Time, $F(1, 46) = 11.03$, $p = .002$, $\eta_p^2 = .19$, but not for Condition, $F(1, 46) < 1$, $p = .426$, $\eta_p^2 = .01$ (**figure 1**, left lower panel). Again, we found the crucial interaction between Time × Condition, $F(1, 46) = 19.48$, $p < .001$, $\eta_p^2 = .30$. Decomposing this interaction effect with paired samples t-tests revealed that detail ratings only decreased for the relevant checkers, $t(23) = 5.03$, $p < .001$, but not for the irrelevant checkers, $t(23) < 1$, $p = .402$.

NJRE

A mixed model ANOVA was used to examine the effects of Condition and Time on reported NJRE. Results are depicted in **figure 1** (right lower panel). There was no effect of Time, $F(1, 46) = 1.10$, $p = .299$, $\eta_p^2 = .02$, nor an effect of Condition, $F(1, 46) < 1$, $p = .929$, $\eta_p^2 = .00$, on self-reported NJRE. Contrary to our expectation, the crucial Time × Condition interaction did not reach significance, $F(1, 46) = 2.77$, $p = .103$, $\eta_p^2 = .06$.

Discussion of Experiment I and introduction to Experiment II

The findings with regards to detail and vividness replicated earlier observations (e.g., van den Hout and Kindt 2003a-b, van den Hout and Kindt 2004, Radomsky et al. 2006, Boschen and Vuksanovic 2007, Giele et al. 2013, Dek et al. 2015). However, while in earlier studies the decline in detail and vividness of memory in the relevant checking group was attended

Figure 1. *Memory confidence, vividness, detail and NJRE before and after relevant/irrelevant checking in Experiment I*
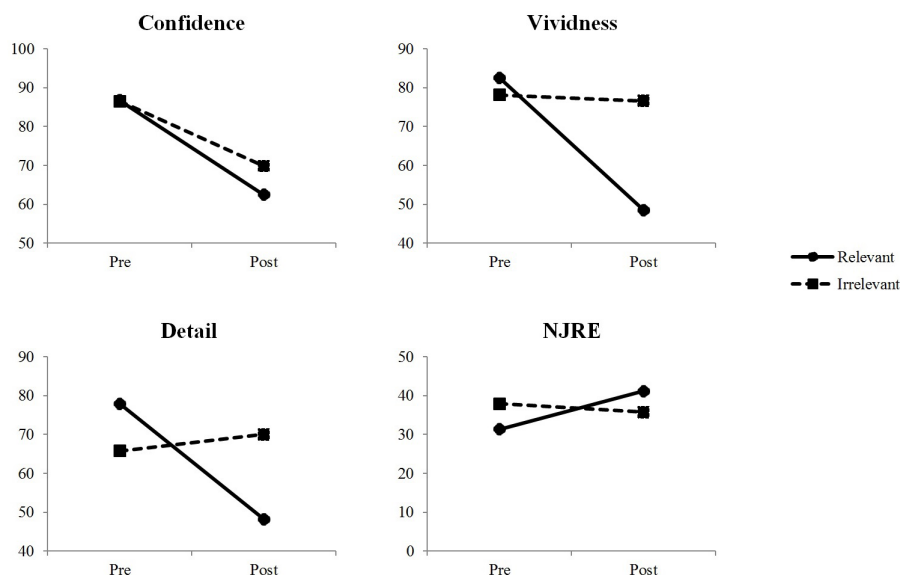
by a comparable drop in memory confidence, in the present study the conditions did not differ. That is, memory confidence decreased to a comparable degree after relevant and after irrelevant checking (see **figure 1**, left upper panel). We have no plausible explanation for this anomaly except that it represents a chance finding.

Contrary to our expectation, there was no support for the main hypothesis that repeated checking leads to NJRE. Still, although the crucial interaction effect did not reach statistical significance, there was a trend with an effect size at medium level (Cohen 1988). Given that effect sizes are much more informative than *p*-values (e.g., Cumming 2014), we did not want to rashly reject our main hypothesis and hence we decided to carry out a replication study. In general, replication studies are crucial to sound science (Cumming 2014, Open Science Collaboration 2015) and this certainly holds true in the present case. The aims of the replication study were to a) evaluate the plausibility that the null-results on memory confidence reflect a chance finding, b) re-examine possible effects of relevant checking on NJREs, and c) slightly adapt the methodology by assessment of objective memory accuracy and replacing an item of the NJRE measure.

Experiment II

Methods

Participants

Sixty-one students aged between 18 and 30 years were recruited at Utrecht University in May 2015 and participated in exchange for course credit or small monetary reward. Six participants indicated having participated in a similar experiment and were excluded from further analyses to ensure they did not participate in Experiment I. The final sample size included fifty-five participants (M age = 21, SD = 2.2; 71% females).

Procedure and experimental task

The procedure of this experiment was identical to Experiment I, with a few modifications described below. In order to get information on response accuracy, the experimental paradigm was reprogrammed in Matlab (version r2014b) which logged actual responses.

Measurements

Accuracy

We logged all responses to verify which gas rings were actually turned on by participants, conform instructions (i.e., objective accuracy).

Meta-memory assessment

At pre- and post-test, participants rated memory confidence, vividness and detail on a VAS (ranging from 0 to 100).

NJRE

We used the same NJRE scale as in Experiment I, but we omitted item 2 because of a negative item-rest correlation. As a replacement for this item we included another item that was based on the Revised NJRE Questionnaire (NJRE Q-R; Coles et al. 2003) which has been used in research by Mancini and colleagues (2008) and that was overlooked when designing Experiment I: *During the*

last checking trial, I had the unpleasant sensation that I was not doing it exactly as I should or the way I would have liked to do it. All questions were answered on a VAS ranging from 0 (completely not applicable to me) to 100 (completely applicable to me). The scale had a good internal consistency (Cronbach's α = .91).

Results

Prior to analyses 10 outliers were changed to $M \pm 2.5\ SD$.

Accuracy

In total, only 7 mistakes were made on the pre- and post-test. In the relevant checking group, no one made a mistake at pre-test, but four participants had inaccurate responses at post-test. In the irrelevant group, only three participants made a mistake at pre-test and no one at post-test.

Meta-memory assessment

Three separate 2×2 mixed model ANOVAs tested the effects of Condition (relevant vs. irrelevant checking) and Time (pre vs. post) on memory confidence, vividness and detail. Confidence in memory differed over Time (pre vs. post), $F(1, 53) = 4.22$, $p = .045$, $\eta_p^2 = .07$, and tended to differ for Condition (relevant vs. irrelevant checking), $F(1, 53) = 2.89$, $p = .095$, $\eta_p^2 = .05$ (see **figure 2**, left upper panel). The crucial interaction effect between Time × Condition was marginally significant, $F(1, 53) = 3.34$, $p = .073$, $\eta_p^2 = .06$. Participants in the relevant checking group experienced a reduction in memory confidence, $t(27) = 2.20$, $p = .037$, whilst participants in the irrelevant checking group did not differ in memory confidence over time, $t(26) < 1$, $p = .800$.

Memory vividness ratings differed for both Time, $F(1, 53) = 1.69$, $p = .200$, $\eta_p^2 = .03$, and Condition, $F(1, 53) = 2.57$, $p = .115$, $\eta_p^2 = .05$ (**figure 2**, right upper panel). Again, the Time × Condition interaction reached significance, $F(1, 53) = 5.83$, $p = .019$, $\eta_p^2 = .10$, and revealed that memory vividness dropped from pre to post in the relevant checking group, $t(27) = 2.19$, $p = .037$, but not in the irrelevant checking group, $t(26) = 1.09$, $p = .288$.

A similar pattern emerged for detail (**Figure 2**, left lower panel), where statistical differences were revealed for both Time, $F(1, 53) < 1$, $p = .797$, $\eta_p^2 = .00$, and Condition, $F(1, 53) = 4.84$, $p = .032$, $\eta_p^2 = .08$. The Time × Condition interaction showed that memory detail did not change over time in the relevant checking group, $t(27) = 1.58$, $p = .126$, but instead increased over time for the irrelevant checkers, $t(26) = 2.47$, $p = .020$; $F(1, 53) = 6.30$, $p = .015$, $\eta_p^2 = .11$.

NJRE

Results are depicted in **figure 2** (right lower panel). Self-reported NJRE differed over Time, $F(1, 53) = 3.74$, $p = .059$, $\eta_p^2 = .06$ and for Condition, $F(1, 53) = 3.83$, $p = .056$, $\eta_p^2 = .07$; the effects being marginally significant. The crucial Time × Condition interaction was marginally significant too, $F(1, 53) = 3.16$, $p = .081$, $\eta_p^2 = .06$. Follow-up analyses revealed that relevant checkers reported higher NJRE over time, $t(27) = 2.18$, $p = .038$, while irrelevant checkers reported similar NJRE levels over time, $t(26) < 1$, $p = .880$.

Eva Anna Maria van Dis and Marcel A. van den Hout

Figure 2. *Memory confidence, vividness, detail and NJRE before and after relevant/irrelevant checking in Experiment II*
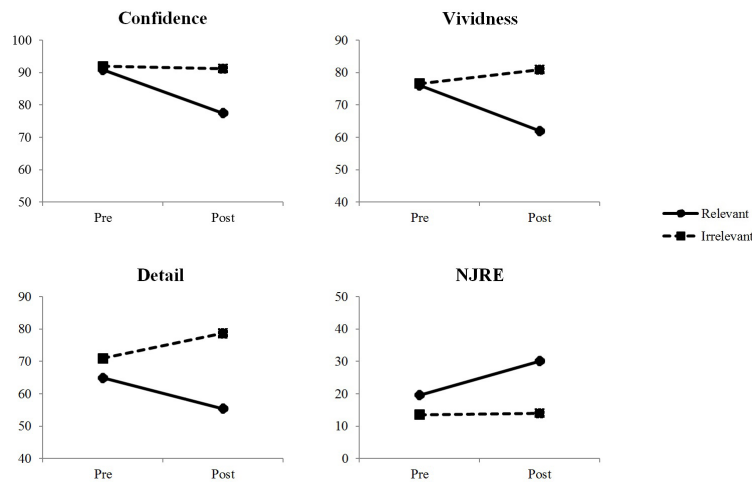
Discussion

Similar to Experiment I, the results showed once more that repeated checking affected ratings of memory vividness and detail. However, as to memory confidence and NJREs the crucial Time × Condition effects were only marginally statistically significant. The fact that, on the latter measures, interaction effects were only marginally significant in both experiments, may suggest that H0 cannot be rejected. Still, especially when considering the medium effect sizes, this rejection may induce a Type II error.

Note that Experiment I and II had next to identical designs, which perfectly allows for pooling the results from both experiments. Given that meta-analyses have, obviously, more power than individual studies and therefore provide better estimates about population parameters (Borenstein et al. 2009), we combined the results in a meta-analysis to see what the two datasets *together*, tell about statistical significances and effect sizes of the observed patterns.

Meta-analytic synthesis of Experiment I and II

Combining the results of Experiment I and II, we performed a meta-analysis on the meta-memory data (confidence, vividness and detail) and NJREs using Exploratory Software for Confidence Intervals (ESCI) developed by Cumming (2012a). Mean effect sizes were calculated using a random effects model. Since Cohen's *d* tends to overestimate the effect size in small samples, we used Hedges' *g* to reduce positive bias (Borenstein et al. 2009).

Figure 3 displays the forest plot of the standardized effect sizes and their confidence intervals with Time × Condition interaction effects on memory confidence, vividness, detail and NJRE respectively. As can be seen from **figure 3**, repeated checking seems to have had a relatively small impact on confidence in memory. The mean effect size for confidence was small to medium; $g = 0.37$, 95% CI [-0.02, 0.75]. On average, repeated checking had strong effects on memory vividness ($g = 0.93$, 95% CI [0.35, 1.54]) and memory detail ($g = 0.96$,

95% CI [0.37, 1.55]). Conform the main hypothesis, the data show a medium effect of repeated checking on NJRE ($g = 0.48$, 95% CI [0.10, 0.87]). Hence the combined results of Experiment I and II demonstrated that individuals who repeatedly checked the gas stove were on average more likely to experience NJREs compared to those who did not repeatedly check.

General discussion

In line with previous research, the current experiments demonstrated once more that repeated checking leads to memory distrust. Data from Experiment I and II showed that participants who repeatedly checked a gas stove reported less vivid and detailed memories of the last check, compared to participants who engaged in an irrelevant checking task. In addition, repeated checking slightly reduced memory confidence. The latter effect was somewhat small, even in the meta-analysis, which is rather surprising given the robustness of this effect in earlier studies (e.g. van den Hout and Kindt 2003a-b, Radomsky et al. 2006, Dek et al. 2015).

The primary aim of the experiments was to test whether repeated checking provokes NJREs. Experiment I yielded a trend in the predicted direction, but it was not statistically significant. In a replication study (Experiment II) we found again a trend that repeated checking resulted in NJREs. Since meta-analyses provide better parameter estimates than individual studies (Borenstein et al. 2009), we conducted a meta-analysis on the findings of both experiments. Even though meta-analyses are often conducted to provide large-scale quantitative reviews of literature, it is important to recognize that small scale meta-analyses (that include only two or three similar studies) already increase the precision of an estimated effect (Cumming 2012b). This overall analysis showed that repeated checking indeed leads to increased NJREs with a medium effect; a smaller effect than that observed for memory vividness and detail, but larger than for confidence in memory (see **figure 3**). While Mancini et al. (2008) were the first to provide evidence for the hypothesis that mental OCD features, such as

Figure 3. *Interaction effects between Condition (relevant vs. irrelevant checking) and Time (pre vs. post) on memory confidence, vividness, detail and NJRE in Experiment I and II*
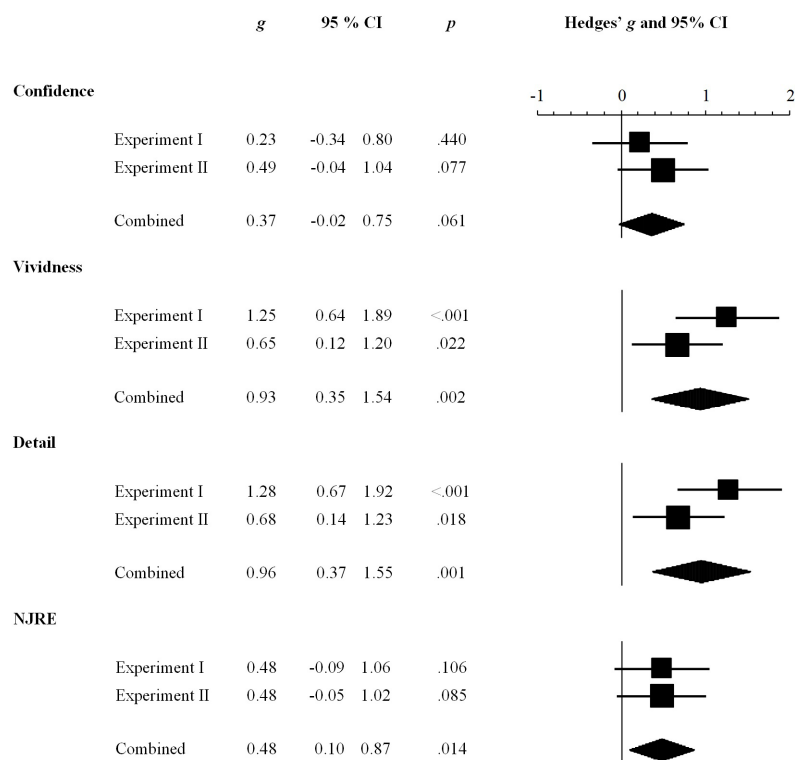
guilt, fuel NJREs; we extended this knowledge by providing new evidence for the hypothesis that also behavioral OCD features, such as repeated checking, may exuberate NJREs. This may be taken to suggest that NJREs serve to maintain repeated checking. Cougle et al. (2013) already showed in an experiment how induced NJREs increased the urge to check, and we showed that perseverative checking reciprocally evoked NJREs. This suggests that OCD patients may enter a vicious circle in which they experience NJREs and start with checking, which subsequently gives rise to more NJREs. Clearly, since all experimental studies on NJREs thus far we only conducted among healthy participants, more research among patients is needed to test whether we can truly speak of such a vicious circle and/or whether other OCD features (third variable) may be involved.

Even though the current experiments showed that repeated checking stimulated NJREs, there is a discrepancy between the current findings and experiences reported by OCD patients. For instance, in a recent diary study OCD patients reported that they achieved a state of certainty or "feeling right" half of the time after compulsive periods (Bucarelli and Purdon 2015). In addition, another study found that patients use "feeling right" as stop rule in compulsive washing (Wahl et al. 2008). Thus, patients report they sometimes "feel right" after perseveration, whilst we found that perseveration stimulated "not just right" feelings. How can these (seemingly) conflicting observations be reconciled? First, the lab-procedure may have failed to reproduce the experiences that occur during clinical checking. However, we attempted to stay close to the NJRE literature and we added in Experiment II the item used by Mancini and colleagues

(2008) which was based on the work of Coles et al. (2003). Moreover, if our paradigm fell short, one may expect that NJREs would not at all be affected by 'experimental checking'. Second, OCD patients may have "feeling right experiences" after checking due to the fact that they reached an artificial stopping criterion. Note that checking more than once (e.g., whether the door is closed) does not give more information and that repetition of checking has no "natural terminus" (Rachman 2002). Patients typically develop strategies to replace a natural ending of behavior by artificial and arbitrary stopping criteria (e.g., closing the door seven times). Reaching the number seven may then give relief – a "feeling right experience". Thus, possibly, patients misattribute this "feeling right experience" to the checking itself (e.g., checking results in a "feeling right experience") instead of attributing it to the artificial terminus (e.g., reaching the number seven).

Experimental checking induces a somewhat dissociative uncertainty about memory (van den Hout and Kindt 2003b, van den Hout and Kindt 2004). During clinical checking too, OCD patients commonly report dissociative experiences (Giele et al. 2016). Future research may focus on the conceptual and experiential nature of NJREs (e.g., to what extent can NJREs be distinguished from related OCD symptoms such as uncertainty of intolerance or dissociative uncertainty). Note that the current studies completely relied on self-report measures and we do not know the divergent validity of the NJRE scale we used (i.e., we cannot rule out whether the NJRE scale also measured other OCD features such as dissociation). Another direction for future research might be to study NJREs among OCD patients and investigate how these feelings develop over time during checking bouts. We showed

in our experiments that repeated checking results in NJREs, while patients sometimes report a "feeling right experience" after compulsive periods. As earlier mentioned, we speculate that OCD patients possibly misattribute a right feeling to reaching an artificial terminus. If so, the time pattern of NJREs during clinical checking could then be: 1) a relatively mild NJRE before checking sets in, 2) an immediate increase in NJRE once checking begins and that remains stable over the course of the checking bout, and 3) disappearance of NJRE once the artificial terminus is reached. Obviously, this inverted-U curve is speculative and future research could test whether this hypothesis holds.

In sum, we tested whether NJREs may be caused or exaggerated by repeated checking. The data of Experiment I and II strongly suggest that this is the case.

Acknowledgements

References

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009). *Introduction to meta-analysis.* John Wiley & Sons, Ltd, West Sussex.

Boschen MJ, Vuksanovic D (2007). Deteriorating memory confidence, responsibility perceptions and repeated checking: Comparisons in OCD and control samples. *Behaviour Research and Therapy* 45, 2098-2109.

Bucarelli B, Purdon C (2015). A diary study of the phenomenology and persistence of compulsions. *Journal of Behavior Therapy and Experimental Psychiatry* 49, 209-215.

Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge Academic, New York, NY.

Coles ME, Frost RO, Heimberg RG, Rheaume J (2003). Not just right experiences: Perfectionism obsessive-compulsive features and general psychopathology. *Behaviour Research and Therapy* 41, 681-700.

Coles ME, Heimberg RG, Frost RO, Steketee G. (2005). Not just right experiences and obsessive-compulsive features: Experimental and self-monitoring perspectives. *Behaviour Research and Therapy* 43, 153-167.

Cougle JR, Goetz AR, Fitch KE, Hawkins KA (2011). Termination of washing compulsions: A problem of internal reference criteria or 'not just right' experience? *Journal of Anxiety Disorders* 25, 801-805.

Cougle JR, Fitch KE, Jacobson S, Lee HJ (2013). A multi-method examination of the role of incompleteness in compulsive checking. *Journal of Anxiety Disorders* 27, 231-239.

Cumming G (2012a). Exploratory Software for Confidence Intervals (ESCI) [Software]. Retrieved from http://www.thenewstatistics.com

Cumming G (2012b). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* Routledge Taylor & Francis Group, New York.

Cumming G (2014). The new statistics: Why and how. *Psychological Science* 25, 7-29.

Dek ECP, van den Hout MA, Engelhard IM, Giele CL, Cath DC (2015). Perseveration causes automatization of checking behavior in obsessive-compulsive disorder. *Behaviour Research and Therapy* 71, 1-9.

Ecker W, Kupfer J, Gönner S (2014). Incompleteness and harm avoidance in OCD, anxiety and depressive disorders, and

non-clinical controls. *Journal of Obsessive-Compulsive and Related Disorders* 3, 46–51.

Ghisi M, Chiri LR, Marchetti I, Sanavio E, Sica C (2010). In search of specificity: "Not just right experiences" and obsessive–compulsive symptoms in non-clinical and clinical Italian individuals. *Journal of Anxiety Disorders* 24, 879–886.

Giele CL, van den Hout MA, Engelhard IM, Dek ECP, Hoogers EEP, De Wit K (2013). Ironic effects of compulsive perseveration. *Memory* 21, 417-422.

Giele CL, van den Hout MA, Engelhard IM, Dek ECP (2014). Paradoxical effects of compulsive perseveration: Sentence repetition causes semantic uncertainty. *Journal of Obsessive-Compulsive and Related Disorders* 3, 35-38.

Giele CL, van den Hout MA, Engelhard IM, Dek ECP, Toffolo MBJ, Cath DC (2016). Perseveration induces dissociative uncertainty in obsessive–compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry* 52, 1-10.

Grissom RJ, Kim JJ (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). Routledge, New York, NY.

Lazarov A, Dar R, Liberman N, Oded Y (2012). Obsessive-compulsive tendencies and undermined confidence are related to reliance on proxies for internal states in a false feedback paradigm. *Journal of Behavior Therapy and Experimental Psychiatry* 43, 556-564.

Mancini F, Gangemi A, Perdighe C, Marini C (2008). Not just right experience: Is it influenced by feelings of guilt? *Journal of Behavior Therapy and Experimental Psychiatry* 39, 162-176.

Moretz MW, McKay D (2009). The role of perfectionism in obsessive–compulsive symptoms: "Not just right" experiences and checking compulsions. *Journal of Anxiety Disorders* 23, 640-644.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349. doi:10.1126/science.aac4716

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

Radomsky AS, Gilchrist PT, Dussault D (2006). Repeated checking really does cause memory distrust. *Behaviour Research and Therapy* 44, 305-316.

Ravid A, Franklin ME, Khanna M, Storch EA, Coles ME (2014). "Not just right experiences" in adolescents: Phenomenology and associated characteristics. *Child Psychiatry & Human Development* 45, 193-200.

Reed GF (1985). *Obsessional experience and compulsive behavior: A cognitive-structural approach.* Academic Press, Orlando.

Sica C, Caudek C, Chiri LR, Ghisi M, Marchetti I (2012). "Not just right experiences" predict obsessive–compulsive symptoms in non-clinical Italian individuals: A one-year longitudinal study. *Journal of Obsessive-Compulsive and Related Disorders* 1, 159-167.

Tallis F (1995). *Obsessive Compulsive Disorder: A cognitive and neuropsychological perspective.* Wiley, Chichester.

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

Van den Hout MA, Kindt M (2003b). Repeated checking causes memory distrust. *Behaviour Research and Therapy* 41, 301-316.

Van den Hout MA, Kindt M (2004). Obsessive compulsive disorder and the paradoxical effects of perseverative behaviour on experienced uncertainty. *Journal of Behavior Therapy and Experimental Psychiatry* 35, 165-181.

Van den Hout MA, Engelhard IM, De Boer C, Du Bois A, Dek ECP (2008). Perseverative and compulsive-like staring causes uncertainty about perception. *Behaviour Research*

and Therapy 46, 1300-1304.
Van den Hout MA, Engelhard IM, Smeets M, Dek ECP, Turksma K, Saric R (2009). *Behaviour Research and Therapy* 47, 535-539.

Wahl K, Salkovskis PM, Cotter I (2008). I wash until it feels right: The phenomenology of stopping criteria in obsessive–compulsive washing. *Anxiety Disorders* 22, 143-161.

Appendix

Table 1. *Means and Standard Deviation of Subjective Ratings for Relevant and Irrelevant Checking Groups (Experiment I)*

	Relevant checking (n = 24)		**Irrelevant checking (n = 24)**	
	Pre	*Post*	*Pre*	*Post*
Meta-memory				
Confidence	86.78 (18.97)	62.46 (32.07)	86.51 (22.76)	69.83 (27.28)
Vividness	82.58 (17.71)	48.54 (26.39)	78.17 (20.38)	76.63 (21.89)
Detail	77.96 (19.08)	48.17 (30.08)	65.79 (25.50)	70.00 (23.00)
NJRE	31.38 (19.71)	41.21 (26.24)	37.95 (23.67)	35.72 (27.25)

Table 2. *Means and Standard Deviation of Subjective Ratings for Relevant and Irrelevant Checking Groups (Experiment II)*

	Relevant checking (n = 28)		**Irrelevant checking (n = 27)**	
	Pre	*Post*	*Pre*	*Post*
Meta-memory				
Confidence	90.91 (17.92)	77.47 (28.90)	91.97 (15.44)	91.19 (17.08)
Vividness	76.10 (28.36)	61.96 (30.76)	76.63 (23.45)	80.88 (22.53)
Detail	64.93 (33.77)	55.39 (32.32)	70.93 (23.28)	78.68 (18.39)
NJRE	19.58 (22.04)	30.13 (30.85)	13.51 (17.31)	13.95 (21.77)