

‘I wash until it feels right’ The phenomenology of stopping criteria in obsessive–compulsive washing

Karina Wahl^{a,*}, Paul M. Salkovskis^b, Imogen Cotter^c

^a University of Luebeck, Department of Psychiatry and Psychotherapy, Ratzeburger Allee 160, 23538 Luebeck, Germany

^b Institute of Psychiatry, Department of Psychology, P077, De Crespigny Park, Denmark Hill, London SE5 8AF, United Kingdom

^c Department of Clinical Psychology, Oxford Regional Training Course in Clinical Psychology, Oxfordshire Mental Healthcare Trust, Warneford Hospital, Warneford Lane, Headington, Oxford OX3 7JX, United Kingdom

Received 21 August 2006; received in revised form 16 February 2007; accepted 21 February 2007

Abstract

Recent elaborations of cognitive behavioral theory in OCD suggest that difficulties in deciding when to stop a compulsive action may be related to the use of counter-productive termination criteria by obsessional patients [Salkovskis, P. M. (1999). Understanding and treating obsessive–compulsive disorder. *Behaviour Research and Therapy*, 37, s29–s52]. Such criteria are characterized by their subjective nature, i.e. a primarily *internal* reference point (e.g. ‘just right’ feelings), and are conceptualized as the “top level” of a general strategy involving elevated evidence requirements. Thirty-eight obsessional washers, 41 obsessionals with other problems and 43 healthy controls were interviewed about and rated two situations varying in the degree of urgency to wash; they also washed their hands in a behavioral test. Washers reported using subjective criteria more frequently and rated them as more important for the termination of the washes than the other groups in questionnaire, interview and laboratory data. Both obsessional groups considered more criteria before stopping than the healthy controls, suggesting that using multiple criteria is a general strategy. The data are consistent with the predictions of the elaborated cognitive-behavioral model of OCD. They indicate that the use of subjective criteria and elevated evidence requirements is affected by the perceived significance of the situation in a similar way for obsessional and non-obsessional individuals.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Obsessive–compulsive disorder; Compulsions; Elevated evidence requirements; Stopping criteria

1. Introduction

Obsessive compulsive disorder (OCD) is a severe and persistent psychological problem with immense negative effects on the individual’s social and working life as well

as on their family (Bobes et al., 2001; Koran, 2000; Koran, Thienemann, & Davenport, 1996; Parkin, 1997). Patients suffer not only from marked anxiety and discomfort associated with their obsessional thoughts but also from the compulsive or neutralizing behavior that is performed in order to prevent the feared consequences from happening (Zaudig, 2006). Obsessional patients characteristically engage in repeated and/or prolonged episodes of compulsive behavior, feeling unable to stop repeating some action over and over again.

* Corresponding author. Tel.: +49 451 500 2468; fax: +49 451 500 2474.

E-mail address: karina.wahl@psychiatrie.uk-sh.de (K. Wahl).

Although other phenomena related to compulsions, such as the motivation to ritualize in the first place, or the anxiety reduction which follows it, have been extensively examined both theoretically and experimentally (Rachman & Hodgson, 1980; Rachman & Shafran, 1998; Salkovskis, Thorpe, Wahl, Wroe, & Forrester, 2003) factors influencing the termination of (and failure to terminate) compulsive activities have received comparatively little research attention. It has been proposed that the repetitiveness of some compulsive behavior, particularly checking, can be accounted for by poor memory for actions (Ecker & Engelkamp, 1995). However, when memory for OCD relevant stimuli is assessed this is found to be intact or even enhanced in OCD (e.g. Radomsky, Rachman, & Hammond, 2001). Data on memory in OCD suggests that the most consistent effects are low levels of *confidence* in memory for OCD related stimuli (Tolin et al., 2001).

The focus on the current study is on the difficulty people suffering from OCD have in stopping behaviors such as washing and checking once they have started what would otherwise be a normal activity. Rather than taking a memory perspective, we instead re-conceptualize the difficulty in stopping behavior as *a problem of decision making*, specifically the process of deciding that enough has been done. The idea that there are problems in decision making processes fit better with the basic phenomenology of OCD. The patient who has washed until their hands bleed does not report doubt that they have been washing or even how long they have washed, but instead report *being uncertain whether they have washed enough*.

This phenomenon can be understood by extending current cognitive theories of OCD, particularly those described by Freeston, Rheaume, and Ladouceur (1996), Rachman (1998) and Salkovskis (1999). The CB model is based on the idea that it is not the intrusive thoughts per se which leads to discomfort and compulsive actions, but the meaning that the person attaches to them. Only if an obsession is interpreted as indicating that one might be responsible for serious harm to oneself or others does it result in the range of reactions and responses characteristic of OCD. The CB-model (Salkovskis, 1999) proposes that a common response to fears of being responsible for harm is the use of potentially counter-productive “stop criteria”, where the person actively seeks to achieve a particular *subjective or emotional state* as a way of deciding that they have completed an activity. This includes the deliberate seeking of a particular mood state, a sense of satisfaction or completeness and “just right” feelings as a way of deciding that it is appropriate

to stop behavior. Such internal states are inherently more difficult for the person to evaluate than sensory input and it therefore takes longer to decide whether they have been met or not. The use of “just right” and other *subjective* states to decide on the termination of an action could be regarded as an example of the operation of “Elevated Evidence Requirements”, motivated by the perception of importance of the outcome of the “stop” decision (Wahl & Salkovskis, submitted for publication). The idea of “elevated evidence requirements” has been discussed previously as an important maintaining factor in chronic worry (Tallis, Eysenck, & Mathews, 1991).

The addition of the concept of EER to cognitive theories of OCD suggests that not only is the *quality* of the criteria that are used to stop an action different for individuals with and without OCD – obsessional patients emphasizing *subjective* criteria – but also the *quantity* of criteria that are taken into consideration. We suggest that obsessional patients are likely to consider *multiple* criteria before they can reach a decision, whereas individuals without any obsessional problems consider only few criteria. Only if they are satisfied that all criteria have been met can they make a decision about whether to stop.

An implication of the active use of elevated evidence requirements is that the decision making process is controlled in the sense that it requires mental effort, is deliberate and conscious (McNally, 1995). Decision making at the end of a non-compulsive wash, on the other hand, would be relatively automatic, i.e. involve little or no mental effort, is not necessarily conscious and not deliberately initiated or terminated. We suggest that the deployment of elevated evidence requirements in decision making is not necessarily a pathological strategy or a generalized cognitive style. Instead, the theory suggests that everybody can and does use elevated evidence requirements given the right circumstances. The extent to which people require more evidence before reaching a decision varies according to the *perceived personal importance* of that decision. For most people, the decision about which sock to put on first requires little consideration of the factors involved. However, deciding whether to take a new job would normally lead to seeking a range of objective information (salary, conditions, location and so on) *combined with* the general felt sense of whether this was the right thing to do or not. Thus, the perception of personal significance and importance of any particular decision is likely to determine the extent to which more evidence is actively sought in order to reach a decision. The more important a decision is, the more likely it is that the evidence sought will include subjective (“it feels right”) elements. These subjective elements

can override the objective elements (“It seemed like a good job, but it didn’t feel right for me”). This ties in with the proposed link between an inflated sense of responsibility for harm in people suffering from OCD (Salkovskis, 1985, 1999). The harm which they perceive might occur in OCD relevant situations will lead them to treat normally unimportant decisions as if they were dangerous leading not only to the tendency to engage in neutralizing, but also to a requirement for elevated levels of evidence in the decision that they had completed the task.

Our conceptualization of elevated evidence requirements is related to other concepts of internal states in OC patients that have recently been investigated by Coles et al., Richards and Davey et al. Coles et al. (2003) examined “not just right experiences (NJRE)” in large undergraduate samples. They found that self-rated NJRE were moderately correlated with OCD features, in particular checking and ordering, and maladaptive domains of perfectionism. Correlations were higher than correlations with depression, social anxiety, trait anxiety or worry. However, the study examined the existence of NJRE in an undergraduate sample and did not investigate its relationship to the termination of compulsive action in OCD patients. This is the aim of the current study. Richards (1995) found that obsessional patients reported relying on internal states such as feelings of rightness or completeness when deciding when to stop a compulsive action, whereas non-clinical controls reported relying on more objective or external criteria like the sound of the door lock. However, Richards investigated the importance of a number of reasons *relative to each other*, but not their *absolute* importance—which is the aim of the current study. Additional evidence for the importance of subjective criteria (such as mood) in the termination of compulsive actions comes from studies by Davey, Startup, Zara, MacDonald, and Field (2003) and MacDonald and Davey (2005). They found that perseveration in a checking task depended on the combination of two stopping rules (“Have I checked as many as possible?” versus “Do I feel like checking?”) with positive and negative mood inductions. In particular, perseveration was most significant when negative mood was combined with the rule “as many as possible”.

A recent account of problematic ‘stop mechanisms’ in OCD by Szechtman and Woody (2004) postulates that compulsive rituals are prolonged by the absence of a “feeling of knowing” which normally terminates a biologically based security motivational system. Similar to our model, the “feeling of knowing” refers to a state of internal origin and thus emphasizes the importance of

subjective criteria in the termination of a compulsive ritual. However, they claim that the “feeling of knowing” is entirely based on internal perceptions such as the performance of a certain (checking or washing) behavior and it is assumed to operate automatically. In contrast, we assume that the “feeling of rightness” is a part of a conscious, i.e. non-automatic, decision making process combining both external *and* internal cues.

The idea of elevated evidence requirements in OC patients supplements current cognitive theories which emphasize the importance of certain dysfunctional beliefs such as “intolerance of uncertainty” (Sookman & Pinard, 2002) *by specifying the way* in which OC patients deal with their uncertainty. A general belief about the necessity of being certain is not sufficient in order to explain why obsessional patients have difficulties in stopping their compulsive behavior. Our theory provides a plausible mechanism that links “intolerance of uncertainty” with the clinically observed difficulty in stopping a compulsive action by specifying the decision making criteria.

In the present study, we have sought to apply the extended cognitive theory to obsessional *washing* because hand washing is an almost universal behavior (rather than being confined to people suffering from OCD), allowing comparison between obsessional washers, obsessional patients who do not have a problem with washing and healthy controls. In washing, the use of elevated evidence requirements is likely to involve some combination of the person looking at their hands, feeling (in a tactile sense) whether they are still sticky, and seeking to feel right about it and not to have any doubts that their hands are clean. These criteria would be operated in a deliberate (strategic) way. In contrast, a “ordinary”, non-obsessive wash would be finished once the person sees that the hands are not dirty any more (if they had been visibly dirty before the wash), or after the passage of a brief period of time, and would be terminated with little deliberate effort (relatively automatically).

In order to evaluate the use of elevated evidence requirements as predicted by our theory, we conducted three connected studies: (1) a standardized and structured interview, (2) an inventory and (3) a behavioral test. Participants were obsessional washers, obsessional patients who did not have a washing problem and healthy controls. The use of an obsessional control group is more stringent than the typical comparison with anxious controls since non-washing obsessionals limit confounding variables that can complicate the interpretation of group differences, such as indecision, memory deficits and so on. Additionally, we investigated all groups in two different situations, one with high

perceived significance and one with low perceived significance. These two different situations allow us to assess the role of the importance of the situation in applying elevated evidence requirements.

The hypotheses being investigated were: firstly, for obsessional washers, internally referenced criteria are more important than for non-washing obsessionals or non-clinical controls. These group differences are expected to be more pronounced in the “most needed wash” compared to a “least needed wash”. It was expected that objective criteria would be at least as important for the washers than for the control groups. Secondly, it was predicted that the total number of criteria being taken into consideration would be higher for obsessional washers than for controls. Thirdly, the decision making process was expected to be more conscious and effortful for obsessional washers than controls. Finally, we predicted that the importance of the situation interacts with the use of elevated evidence requirements: the higher the perceived personal significance of the situation, the more likely the use of elevated evidence requirements.

2. Part 1: A semi-structured interview

A standardized interview technique was chosen since to date little is known about stopping criteria in OCD patients. The main objective of this part of the study was to generate ideas about stopping criteria in obsessional washing. Additionally, we categorized the qualitative data according to our own theory in order to see whether we could find indication of the use of elevated evidence requirements. It was predicted that washers report subjective or internally referenced criteria more often than controls.

2.1. Methods

2.1.1. Overview

Participants were interviewed about two different situations in which they had washed their hands; one situation when they really felt that they needed to wash their hands (for the washers, this was a compulsive wash) and one when they did not feel that they needed to wash, but did so anyway because it was a social custom or a habit. The main focus of the interview was on the termination of the wash.

2.1.2. Participants

Thirty-eight obsessional patients with predominantly washing problems (washer), 42 obsessional patients without washing problems (non-washing OCD)

and 43 participants without any known psychiatric problems (healthy controls) participated in the washing interview study. The same participants also participated in the washing inventory study and the behavioral test. All participants were diagnosed according to DSM-IV criteria, using the Structured Clinical Interview (SCID: First et al., 1995) administered by experienced clinicians or a trained research worker. Washers were obsessional patients who were either diagnosed as OCD based on the washing problem *alone* or whose *primary* obsessional problems were focused on compulsive hand washing and contamination fears (i.e. these being more time consuming, interfering more with the person's life and being more distressing relative to other obsessional problems the person may have had). Compulsive hand washing was defined as “repetitive hand washing a person feels driven to perform due to obsessional fears or rigid rules”. Contamination fears were defined as “fears of getting contaminated with dirt, germs, or other substances that the person considers dangerous by means of real or imagined contact with an object.” Obsessional patients were classified as non-washers if their primary obsessional problem was *not* washing (but checking, mental rituals, etc.). Obsessional patients with a rudimentary washing problem or slight contamination fears were not included in the study. The final non-washing OCD group consisted of 20 obsessional patients with predominantly checking problems and the remaining patients were suffering from predominantly obsessions and mental neutralizing, one of whom also had an obsessional ordering problem and one an additional hoarding compulsion.

Exclusion criteria were age below 16 or over 75 or current or past psychotic episode. Patients were recruited from an outpatient service at the Department of Psychiatry, Oxford, or at the Institute of Psychiatry, London. They participated in the study shortly before or at the early stages of their treatment (CBT for OCD). It was ensured that “stop criteria” had not been discussed in therapy. Seven more patients were recruited from a self-help group (‘Obsessive Action’). Healthy controls were recruited from a data base of individuals who had indicated an interest in taking part in research in Oxford. Participants were paid £5 or travel expenses.

Demographic variables, primary and secondary diagnosis and length of the obsessional problem are displayed in Table 1. The three groups did not differ in age, $F(2,119) = 0.443$, $p = 0.643$, gender ratio, $\chi^2(2) = 2.256$, $p = 0.324$, or marital status, $\chi^2(2) = 1.882$, $p = 0.390$. Occupational status was significantly different between groups, $\chi^2(2) = 11.132$, $p < 0.004$, indicating

Table 1
Demographic and diagnostic characteristics by group

	W (<i>n</i> = 38)	NW (<i>n</i> = 41)	HC (<i>n</i> = 43)
Gender			
Female:male	14:24	19:22	23:20
Age			
<i>M</i> (S.D.)	36.6 (11.8)	35.8 (11.2)	38.3 (14.6)
Duration of OCD in years ^a			
<i>M</i> (S.D.)	16.4 (12.4)	13.5 (10.9)	–
Years in education			
<i>M</i> (S.D.)	13.1 (2.6)	14.5 (2.9)	14.6 (2.3)
Occupation ^b			
Employed or full time education	62.2% (23)	62.5% (25)	90.7% (39)
Not employed ^c	37.8% (14)	37.5% (15)	9.3% (4)
Marital status			
With partner (married or with partner)	52.6% (20)	41.5% (17)	55.8% (24)
Without partner (single, divorced, widowed)	47.4% (18)	58.5% (24)	44.2% (19)
Primary DSM-IV diagnosis			
OCD	100% (38)	100% (41)	0% (0)
Secondary DSM-IV diagnosis			
Panic without agoraphobia	5.3% (2)	7.3% (3)	0% (0)
Panic with agoraphobia	2.6% (1)	0% (0)	0% (0)
Agoraphobia without panic	0% (0)	2.4% (1)	0% (0)
Social phobia	5.3% (2)	4.9% (2)	0% (0)
Generalized anxiety disorder	7.9% (3)	4.9% (2)	0% (0)
Specific phobia	0% (0)	2.4% (1)	0% (0)
Major depression	21.1% (8)	14.6% (6)	0% (0)
Alcohol abuse	5.3% (2)	0% (0)	0% (0)
No psychopharmacological drugs	50% (19)	34.2% (14)	97.7% (42)
Medication			
Antidepressant drugs	47.4% (18)	58.5% (24)	2.3% (1)
Other psychopharmacological drugs (including anxiolytics and benzodiazepines)	0% (0)	2.4% (1)	0% (0)

Note: Means are shown with standard deviations in parentheses. Percentages are shown with frequencies in parentheses. W, washer; NW, non-washing obsessionals; HC, healthy controls.

^a For two washers the duration of the obsessional problem is missing.

^b For one washer, and one non-washing OCD patient information about the occupational status is missing.

^c Not employed included unemployed, housewife/houseman or retired.^dFor one obsessional washer and two non-washing obsessionals information about medication is missing.

that only a small proportion of non-clinical controls did not work or study. Years spent in education were also significantly different between groups, $F(2,119) = 28.268$, $p = 0.018$. Post hoc Tukey HSD tests demonstrated that the washers spent less time in education than non-washing obsessionals, $p < 0.05$, and healthy controls, $p < 0.05$. The two control groups did not differ, $p = 0.992$. Eighteen washers and 15 non-washing obsessionals met diagnostic criteria for an additional axis-I disorder. Both clinical groups were comparable in the length of the obsessional problem, $t(75) = 1.073$, $p = 0.287$ and the proportion of patients that took medication, $\chi^2(2) = 2.564$, $p = 0.278$.

2.1.3. Material and laboratory equipment

The washing interview took place in the laboratory at the Department of Psychiatry in Oxford or at the Institute of Psychiatry in London. Experimenter and participant were seated opposite each other next to a table and an OPTIMUS CTR-108 tape recorder was used to audiotape the interview.

2.1.4. Measures

2.1.4.1. Standardized questionnaires. The standardized questionnaires used for this study were Beck Anxiety Inventory (BAI: Beck, Steer, & Garbin, 1988), Beck Depression Inventory (BDI: Beck, Ward,

Mendelsohn, Mock, & Erbaugh, 1961); The State-Trait Anxiety Inventory (STAI: Spielberger, 1983); Obsessive–compulsive Inventory (OCI: Foa, Kozak, Salkovskis, Coles, & Amir, 1998).

2.1.4.2. The washing interview. The washing interview started with a definition of a ‘most needed’ or compulsive wash as:

“A situation when you felt compelled to wash your hands. This means a situation when you felt driven to wash your hands. This can be in response to an intrusion (that is, doubts, ideas or thoughts like ‘I am contaminated’) or it may be in direct response to contamination. You might recognize the hand washing as senseless or excessive, but feel you have to do it anyway.”

If control participants said that they could not think of any such situation the interviewer asked them to think of a recent situation when they *most felt* they needed to wash their hands. The ‘non-compulsive wash’ was defined as

‘a particular situation in the last few days when you did not feel driven to wash, but did so anyway because it was the thing to do (e.g. social custom or habit)’.

If washers replied that they could not think of such a situation, they were asked to think of an occasion when they *least needed* to wash their hands. This was true for 11%.

Participants were asked to think of a particular situation when they had to wash their hands in a compulsive way during the last few days. To facilitate the recollection of the wash participants were given a few minutes to imagine the situation clearly and to think about where they were, what time of the day it was, what they were doing, if they were alone and how they were feeling at that time. They were subsequently asked to take a few minutes to imagine what they did when they washed their hands at that time and eventually to take the experimenter through the wash step by step. The importance of having a clear picture was emphasized. The experimenter wrote down the sequence of each distinct washing action and then repeated the sequence to ensure she had understood it correctly. Participants were encouraged to describe what was on their mind at that time, what seemed to them to be the worst thing that could have happened if they did not wash their hands, how long they thought the wash took, and what made them start. Finally, the interviewer asked how participants decided when to

stop washing. Each reason provided was rated on a 0 to 100 visual analogue scale (VAS) for how important it was in order to stop washing at that time (0 = not at all important; 100 = extremely important).

2.1.4.3. Psychometric properties of the interview. Topics of the interview were standardized, the order of the questions was pre-determined, and rules were set as to which and how many prompts were allowed in order to keep biased information processing to a minimum (the interview is available from the authors on request). Rules were agreed on as to what kind of prompts were to follow after different types of answers and which information would be accepted as the final answer. Thus, choices for the interviewers were kept to a minimum. For example, the first question for the “reasons for stopping” was: “How did you decide when to stop washing?” If participants provided a reason, interviewers were instructed to continue with “Which other factors influenced your decision to stop?” If, on the other hand, participants did not provide an answer or said “I don’t know”, a further question was “Which factors influenced your decision to stop?” Both questions were continued until participants gave the same answer twice. The second question for stopping reasons was “At that time, how important out of a 100 was each of these factors for you to make the decision to stop?” These measures should increase the reproducibility of the interview, its reliability and possibly its validity. A random sample of five written protocols was checked with the taped interviews to identify errors due to attention, memory or protocol biases on the interviewer’s side. There was a perfect agreement in content with the exception of repetitions which were not included in the written protocols.

2.1.5. Procedure

Participants were sent a questionnaire pack with the instruction to complete all measures the day before they participated in the study. On arrival, participants were trained in the use of the VAS used during the washing experiment. Part one of the semi-structured interview about a compulsive or most needed wash was administered first (see description of interview for details). Interviewers were two trained postgraduate level psychologists supervised by a senior clinical psychologist. Three pilot interviews from each interviewer were videotaped and discussed in supervision, any inconsistencies in the questions or prompts were discussed and the interviews amended if necessary. Participants were asked to complete the washing inventory (see Section 3, Part 2) referring to the

compulsive or most needed wash just discussed. Part two of the interview was administered and participants subsequently completed the washing inventory in relation to the least needed wash. Twenty-nine participants (10 washers, 17 non-washing OCD and 2 healthy controls) were asked to fill in the washing inventories again the next day and to return it by post. This was done in order to establish test–retest reliability.

2.1.6. Treatment of data

2.1.6.1. Qualitative analyses of interview data. The verbatim transcripts of the interviews served as the raw data. A content analysis of this data was conducted, focusing on the question ‘How did you decide when to stop washing?’ The data were reduced using a category system developed for this purpose; with similar reasons given by participants being allocated to a common category according to a manual (the category system is available from the authors on request). The initial category system was based on the nine categories of the washing inventory (see Section 3.1.4 below). Once all reasons under ‘How did you decide when to stop washing?’ had been coded according to the category system, and the inter rater reliability established, one rater read through the complete interviews again in order to see whether at some point during the interview any other reasons for stopping were mentioned or whether the meaning of any of the stopping criteria changed when the interview context was taken into account. If any new reasons emerged, they were categorized as before. The content and frequency of this final categorization were then compared between groups using quantitative methods.

To establish the inter-rater reliability of the category system a second rater, who was blind to diagnosis, rated all interviews using the final categories and coding rules. The inter-rater reliability was calculated using Cohen’s κ coefficient, which represents the standardized percentage agreement above that expected by chance (Wirtz & Caspar, 1994). The inter-rater agreement for the category system was high (Cohen’s $\kappa = 0.874$ for the compulsive wash; Cohen’s $\kappa = 0.848$ for the non-compulsive wash).

2.1.6.2. Categories used. Each response to the question: ‘How did you decide when to stop washing?’ was classified according to the category system that was developed for that purpose. Categories were grouped together depending on whether the decision was predominantly based on objective (i.e. perceptions of external changes) or subjective criteria (i.e. referring to

some kind of internal psychological state such as mood or an idiosyncratic rule) and on what type of subjective and objective criteria was employed, resulting in four large categories that were of theoretical interest for our study (see Appendix A for sub-categories):

External, perception: Comprising statements that were based on a sensory perception.

Internal, feeling/mood: Indication that a general feeling of rightness/completeness, satisfaction, cleanliness, mood or an epistemological sense of ‘just knowing’ was considered for the decision to stop.

Internal rules, memory: Predetermined rules or memory related cues determined the wash.

Internal, effort: Indication that a certain amount of effort had to be put into the wash.

The value of the larger categories was calculated by computing the sum of the importance ratings across sub-categories and subsequently dichotomizing the total variable ($0 = 0$; $>0 = 1$). The resulting category scores reflect the number of participants who used at least one of the criteria in this category.

2.1.6.3. Treatment of missing data. If single items were missing on the questionnaires (BDI, BAI, STAI-S, STAI-T, OCI), they were replaced by the mode of the scale items (BDI, BAI, STAI-S, STAI-T) or the subscale items (OCI). If there were two modes, the lowest one was taken. An upper limit of item replacements per scale or subscale was agreed on (one for all subscales other than OCI total where two were allowed).

2.2. Results

2.2.1. Overview

For the most needed wash, more washers than non-washing obsessionals or healthy controls spontaneously reported using their feelings when making the decision to stop washing. These group differences were also found for the least needed wash. Fewer washers than healthy controls reported using external perceptually based criteria in the most needed wash, but not in the least needed wash. Washers reported using more criteria overall than controls for both types of washes.

2.2.2. Measures of psychopathology

A description of the psychopathology measures per group and group comparison statistics are shown in Table 2. One-way ANOVAs with group (washers (W) versus non-washing obsessionals (NW) versus healthy

Table 2
Comparison of mean measures of psychopathology

	M (S.D.)			F(2,119)
	W	NW	HC	ANOVA group effect
BDI	18.92 (8.86) ^a	17.83 (9.22) ^a	6.47 (5.67) ^b	30.775 ^{**}
BAI	18.29 (9.94) ^a	18.71 (9.07) ^a	6.6.7 (5.08) ^b	28.895 ^{**}
STAI, trait	57.74 (8.95) ^a	56.41 (9.58) ^a	36.77 (9.65) ^b	64.819 ^{**}
STAI, state	48.16 (11.72) ^a	47.63 (12.02) ^a	31.19 (9.75) ^b	31.142 ^{**}
OCI total distress	79.26 (26.74) ^a	59.00 (26.44) ^b	9.67 (10.13) ^c	104.455 ^{**}
OCI washing distress	21.32 (7.93) ^a	4.10 (4.38) ^b	1.31 (2.02) ^c	166.696 ^{**}
OCI checking distress	17.08 (8.33) ^a	16.12 (8.86) ^a	1.40 (2.21) ^b	62.550 ^{**}
OCI neutralizing distress	7.74 (6.03) ^a	7.68 (5.76) ^a	1.02 (1.51) ^b	26.129 ^{**}
OCI ordering distress	8.53 (5.01) ^a	6.85 (6.59) ^a	1.69 (2.29) ^b	21.125 ^{**}
OCI obsessions distress	14.05 (8.10) ^a	15.27 (6.87) ^a	2.62 (4.05) ^b	21.125 ^{**}
OCI hoarding distress	4.11 (3.52) ^a	3.83 (3.77) ^a	1.29 (1.80) ^b	26.129 ^{**}
OCI doubting distress	6.89 (3.32) ^a	5.49 (3.66) ^a	0.59 (1.04) ^b	53.126 ^{**}

Means with different superscripts (a and b) differ significantly according to Tukey HSD post hoc tests, $p < 0.05$. BDI: Beck Depression Inventory; BAI: Beck Anxiety Inventory; STAI, trait/state: The State-Trait Anxiety Inventory, trait/state version; OCI: Obsessive–compulsive Inventory.

^{**} $p < 0.001$.

controls (HC)) as between-subjects factor and Tukey HSD tests as multiple comparisons demonstrated that washing obsessionals and other obsessionals were comparable in terms of depression (BDI), general anxiety (BAI and STAI trait) and state anxiety (STAI state) before the washing interview, whereas non-clinical controls had significantly lower scores on all of these measures. The washers were characterized by a higher overall level of obsessive compulsive symptomatology (OCI totals and distress), and higher scores on the OCI washing subscale compared to other obsessionals and non-clinical controls. The obsessional control group had higher OCI total scores and washing scores than the non-clinical controls. On all other OCI subscales (checking, neutralizing, ordering, obsessions, doubting and hoarding) the OCD groups were not significantly different from each other but had significantly higher scores than the non-clinical control group.

2.2.3. Frequency of self-reported reasons to stop washing

The frequency of responses falling into the four main categories described above (external perception, internal feeling, internal rules/memory and internal effort) were analyzed across groups using Chi-square. For the “most needed wash”, where these were significant at $\alpha \leq 0.01$, two further analyses were then undertaken: (a) the overall category was partitioned and (b) the individual items composing them were then analyzed using that same α , and where these were significant in turn, they were also partitioned by group. For the “least

needed wash”, the same analysis was conducted using less conservative α of ≤ 0.05 . The analysis for differences was therefore set at a relatively conservative level, whilst where there was no difference predicted, a more liberal α was used. Follow-up group comparisons were Bonferroni corrected.

2.2.3.1. Most needed wash. There was a significant association between group and external perception, internal feeling, internal rules/memory but not internal effort (see Table 3). For external criteria, fewer washers than healthy controls used at least one kind of perceptual information, $\chi(1) = 10.152$, $p < 0.003$. The difference between washers and non-washers did not reach significance, $\chi(1) = 2.937$, $p = 0.087$, nor did the difference between two control groups, $\chi(1) = 2.339$, $p = 0.126$. The only sub-component which reached significance was the use of visual cues, $\chi(2) = 10.900$, $p < 0.01$. Fewer washers than healthy controls reported using visual cues, $\chi(1) = 10.673$, $p < 0.003$. The difference between washers and non-washing obsessionals did not reach significance, $\chi(1) = 3.622$, $p = 0.057$, nor did the difference between two control groups, $\chi(1) = 2.494$, $p = 0.114$.

For internal feelings, more washers than non-washing obsessionals, $\chi(1) = 10.588$, $p < 0.003$, and healthy controls, $\chi(1) = 15.007$, $p < 0.003$ reported using any kind of internal feelings. The control groups did not differ from each other, $\chi(1) = 0.417$, $p = 0.518$. Only one individual component reached significance, ‘it feels right’, $\chi(2) = 15.397$, $p < 0.01$. When partitioned into groups, more washers than healthy controls used a

Table 3
Comparison of percentages for categories ‘reasons to stop washing’

	W	NW	HC	Total	χ^2 (2)
Most needed wash					
External, perception	14 (36.9) ^a	23 (56.1) ^{a,b}	31 (72.1) ^b	68 (55.7)	10.164 ^{**}
Internal, feeling	23 (60.5) ^a	10 (24.4) ^b	8 (18.6) ^b	41 (33.6)	18.240 ^{**}
Internal, rules/memory	4 (10.5) ^a	0 (0) ^a	0 (0) ^a	4 (3.3)	9.142 ^{**}
Internal, effort	5 (13.2)	4 (9.8)	3 (7.0)	12 (9.8)	0.870
Least needed wash					
External, perception	16 (42.1)	17 (41.5)	21 (48.8)	54 (44.3)	.567
Internal, feeling	14 (36.8) ^a	4 (9.8) ^b	5 (11.6) ^b	23 (18.9)	11.723 [*]
Internal, rules/memory	1 (2.6)	0 (0)	0 (0)	1 (0.8)	2.229
Internal, effort	3 (7.9)	1 (2.4)	1 (2.3)	5 (4.1)	2.024
External, perception	16 (42.1)	17 (41.5)	21 (48.8)	54 (44.3)	.567

Note: Frequencies and percentages in parentheses of participants within group who used *at least one* of the sub-categories that compose the larger categories. For clarity, only the larger categories are displayed. Detailed results are available from the authors on request.

Means with different superscripts (a and b) differ significantly using follow-up Chi-square analysis, $p < 0.02$.

* $p < 0.05$.

** $p < 0.01$.

feeling of rightness, $\chi(1) = 12.912$, $p < 0.003$. The difference between washers and non-washing obsessionals in ratings of ‘it felt right’ did not reach significance, $\chi(1) = 5.178$, $p = 0.023$, nor were the control groups different from each other, $\chi(1) = 3.263$, $p = 0.071$. For internal rules/memory, differences did not hold up when comparisons were partitioned by group.

2.2.3.2. Least needed wash. Of the four primary analyses, external perception, internal memory and internal effort did not reach the basic significance level of 0.05. However, internal feeling did (see Table 3). When partitioned into groups, more washers than non-washing obsessionals, $\chi(1) = 8.224$, $p < 0.016$, and more washers than healthy controls, $\chi(1) = 7.143$, $p < 0.016$, reported using any kind of feeling. The difference between the control groups was not significant, $\chi(1) = 0.077$, $p = 0.782$.

2.3. Discussion of semi-structured interview

The most important finding in our interview study is that we were successful in identifying internal reasons linked to feelings of ‘rightness’. Reasons for stopping a wash could be meaningfully grouped into more subjective (or internally referenced) and more objective (or externally referenced) categories. As expected, more washers than non-washing obsessionals or healthy controls reported using any type of internally referenced feelings in order to stop a compulsive/most needed wash. Looking at the individual ‘feeling criteria’ only a feeling of rightness remained significant between

groups. We found that washers reported using a feeling of rightness more often than healthy controls, however, inconsistent with our hypothesis, difference did not reach significance between the washers and the non-washing obsessionals.

Contrary to predictions, a feeling of rightness was also mentioned by more washers than controls in the *least needed wash*. This result points towards a *general strategy* involving the use of internal feelings. Alternatively, it could be an artifact of the data collection procedure, since it was agreed that – if washers could not remember a time when they washed in a non-compulsive way – they should take a recent situation when they *least* felt they needed to wash. However, only 11% of the participants chose such a situation which renders the alternative explanation unlikely.

In addition, there is some indication that pre-terminated rules and memory based cues play a considerable role in the termination of a compulsive wash such as a certain amount of time, number of times or having no doubts. Note, however, that these criteria were observed less often than internal feelings. Obsessional washers may have used these particular rules and memory cues more often than the other two groups, but these were ‘trumped’ by a feeling of rightness which was regarded as crucial and therefore dominated the recollection of termination of the wash. Future research should clarify the possible hierarchy of termination criteria. The finding of memory based considerations is also consistent with research that found a lack of confidence in obsessional memories (Ecker & Engelkamp, 1995; Radomsky et al., 2001; Zitterl et al., 2001).

Interviewers could not be kept blind to diagnosis, which may have biased results. However, several measures were taken to minimize interviewer effects (the interview was standardized including the use of prompts, the scripts were randomly checked with the taped interviews and the raters who categorized the data were not told of the participants' group status and hypothesis; inter-rater reliability was high).

Overall, the results of the free-responses in the semi-structured interview demonstrated that some kind of internal feeling is more important for obsessional washers than for the control groups in the termination of a wash. Our second study represents a focused evaluation of the specific hypotheses using rating scales to derive quantitative data.

3. Part 2: The washing inventory

The second part of the study was designed specifically to evaluate the elevated requirement for evidence in a more structured way using ratings tailored to these concepts. The hypotheses therefore addressed in this study are: (i) for obsessional washers, internally referenced criteria will be rated as more important than for non-washing obsessionals and healthy controls. These group differences should be more pronounced in the most important wash relative to the least important wash. Objective criteria should be at least as important for washers as for controls. (ii) Obsessional washers will have *more* criteria with higher ratings than non-washing obsessionals and non-clinical controls, the difference being greater for the most important wash. In addition, it is hypothesized that (iii) the decision making process will be rated as more effortful and conscious for the obsessional washers than the two control groups, the difference being less pronounced for the least relevant wash. Controlled and automatic processing in this study are defined as two ends on a continuum with controlled processing being characterized as effortful, conscious and subject to voluntary control (McNally, 1995). We included this construct as a subjective measure in order to get a first indication of the perception of automaticity, knowing that more objective measures of controlled processing are desirable in future studies. Finally, we predicted that the importance of the situation interacts with the use of elevated evidence requirements.

3.1. Methods

3.1.1. Overview

Participants were asked to complete two self-report questionnaires about the importance of a series of

factors that could have influenced their decision to stop washing at that time. The inventories were administered only *after* the interview so that the interview would assess the participants' views about the termination of a wash in a relatively unbiased way. However, it was also intended that the interview would *prime* the participants' memory for specific reasons used to stop a wash, increasing the ecological validity of the ratings.

For "Participants", and "Material and laboratory equipment" please refer to study Part 1, Sections 2.1.2 and 2.1.3.

3.1.2. Instruments

3.1.2.1. The washing inventory. The washing inventory followed each part of the interview (compulsive and non-compulsive wash) to take advantage of the increased salience of the concepts discussed in the interview. The first part of the washing inventory consisted of several reasons which people might give for deciding to stop washing. Participants had to indicate on a VAS from 0 to 100 how important each individual reason was for them to stop washing during the just discussed wash. 0 was labeled 'not at all important' and 100 was labeled 'extremely important'.

The five reasons or criteria which were important for our study included a question on 'feeling right', two further more internally referenced criteria ('to have tried hard enough', 'to feel physically or mentally exhausted') and two more objective criteria ('hands looked completely clean'; 'hands didn't feel sticky/dirty any more'). The final part of the inventory consisted of two questions referring to the nature of the decision. Participants had to rate on a 0–100 VAS (a) to what extent they were consciously deciding to stop washing (0 = completely automatic, required no attention; 100 = required all my attention) and (b) to what extent the decision to stop washing was made with deliberate effort (0 = no effort required; 100 = extremely effortful).

The following composite variables were based on the washing inventory:

- *Subjective or internally referenced criteria:* Means of the following variables: 'feel right', 'tried hard enough' and 'physically or mentally exhausted'?
- *Objective or externally referenced criteria:* Means of the two items 'hands looked completely clean' and 'hands did not feel sticky or dirty any more'.
- *Overall number of criteria:* Sum of those criteria on the washing inventory that had been given a rating of greater than zero, the new variable represents the total number of dimensions that were considered.

- *Controlled processing*: Means of the two items ‘consciously deciding’ and ‘with deliberate effort’.

The 24-h test–retest reliability of the washing inventories was established with $N = 29$ participants (10 washer, 17 non-washing OCD and 2 non-clinical controls), using Pearson Correlations as the distributions met criteria for bivariate normal distributions. Correlation coefficients were: most needed wash: objective criteria 0.749; subjective criteria 0.760; number of criteria 0.787; automaticity (controlled processing) 0.762; least needed wash: objective criteria 0.630; subjective criteria 0.632; number of criteria 0.888; automaticity (controlled processing) 0.882. Additionally, all means remained stable over time. Thus, most test–retest reliability coefficients are moderate.

3.1.3. Procedure

After completion of the first part of the interview about the compulsive or ‘most needed’ wash, participants were asked to complete a washing inventory referring to the compulsive or most needed wash just discussed. The washing inventory referring to the non-compulsive or ‘least needed’ wash was completed once the interview about the second part of the interview was finished. Before participants were dismissed, 29 participants were asked to fill in the washing inventories again the next day and to return it by post to establish the re-test reliability.

3.1.4. Treatment of data

The overall type one error probability was set at $\alpha = 0.05$. Mixed-model ANOVA was the main procedure for the parametric data. The Tukey Honestly Significant Difference Test (Tukey HSD) was employed for post hoc analysis when parametric differences between all three groups were of interest. Helmert contrasts were used

when the difference between the washers and the mean of both control groups was investigated. They test simultaneously whether the difference between washers and the two control groups is significant, and whether the two control groups are significantly different from each other. Kruskal–Wallis and Mann–Whitney tests for follow-up analysis were used for non-parametric data. With the exception of Tukey HSD tests, follow-up group tests were Bonferroni corrected.

3.2. Results

3.2.1. Subjective and objective criteria

Subjective and objective criteria measures were both subjected to a mixed model ANOVA with group (W versus NW versus HC) as the between subjects factor and type of criteria (subjective versus objective) and importance of the situation (most needed wash versus least needed wash) as within subject factors. Means and standard deviations of objective and subjective criteria are shown in Table 4. There was a significant main effect of group, $F(2,119) = 8.675$, $p < 0.001$, main effect of type of criteria, $F(1,119) = 46.827$, $p < 0.001$, and main effect of importance of the situation, $F(2,119) = 54.822$, $p < 0.001$; with $M = 53.22$ and 38.93 for most and least needed wash. There was also a significant two way interaction between type and group, $F(2,119) = 3.806$, $p < 0.05$. The three-way interaction between group, type and importance was not significant $F(2,119) = 1.975$, $p = 0.146$.

The two-way interaction between group and type of criteria was investigated further by analyzing group differences separately for subjective and objective criteria for the troublesome wash only using Helmert contrasts. For subjective criteria, the overall comparison between groups was significant, $F(2,119) = 20.587$,

Table 4

Questionnaire data: importance of objective and subjective criteria on VAS (0 = not at all important; 100 = extremely important), and total number of criteria, means and standard deviations

	<i>M</i> (S.D.)		
	W	NW	HC
Most needed wash			
Subjective criteria	60.48 (20.19)	37.52 (21.36)	34.30 (17.65)
Objective criteria	70.39 (34.06)	52.32 (33.73)	64.30 (30.68)
Least needed wash			
Subjective criteria	42.50 (25.72)	29.27 (19.29)	23.33 (17.68)
Objective criteria	55.39 (36.14)	38.66 (30.35)	44.42 (34.18)
Total number			
Most needed wash	7.58 (1.43)	6.78 (1.35)	6.05 (1.84)
Least needed wash	6.71 (2.08)	6.41 (1.63)	5.33 (2.50)

Note: Total number of criteria is the sum of all criteria rated as >0 , maximum number was nine.

$p < 0.001$. The difference between washers and the two control groups was significant (mean difference of 24.57, $p < 0.001$; 97.5% confidence interval: 15.81–33.34), the difference between the two control groups was not significant (mean difference of 3.22, $p = 0.457$). For objective data, the overall comparison between groups did not reach significance when α was adjusted, $F(2,119) = 3.137$, $p = 0.047$). In Helmert contrasts, the difference between washers and the control groups was a non-significant trend (mean difference of 12.085, $p = 0.062$), as was the difference between the two control groups (mean difference of -11.985 , $p = 0.097$).

These results indicate that, consistent with hypothesis, for the washers, the subjective reasons for stopping were rated as significantly higher than for both control groups. The personal significance of the situation influenced all ratings: during the most important wash, all groups rated all criteria as more important.

3.2.2. Number of criteria being considered

The variable ‘overall number of criteria’ (defined as the sum of all criteria being rated greater than 0) was subjected to a mixed model ANOVA with importance of the situation (most needed versus last needed wash) as within subject factor and group (W versus NW versus HC) as between subjects factor. The maximum number of criteria that could be considered was nine. Table 4 contains means and standard deviations per group and type of wash. There was a significant main effect of group, $F(2,119) = 7.907$, $p < 0.001$ (with $M = 7.2$, 6.6, and $M = 5.7$ for W, NW and HC, respectively) and main effect of importance of the situation, $F(1,119) = 20.199$, $p < 0.001$, indicating that the number of criteria was higher for the most needed wash than for the least needed wash ($M = 6.80$ and 6.15). The interaction was not significant, $F(2,119) = 1.043$, $p = 0.356$. Post hoc multiple comparison tests demonstrated that washers used more criteria than healthy controls, Tukey HSD, $p < 0.001$, but not more than other obsessionals, Tukey HSD; $p = 0.318$. Other obsessional also used more criteria than healthy controls, Tukey HSD; $p < 0.05$.

This means that washers reported using more criteria than non-psychiatric controls, but washers and non-washing obsessionals did not differ on the number of criteria they used. Again, results were influenced by the importance of the situation: all participants used more criteria for the most needed wash than for the least needed wash.

3.2.3. Controlled processing

The variable ‘Controlled processing’ was subjected to a mixed model ANOVA with importance of the

situation (most needed versus last needed wash) as within subject factor and group (W versus NW versus HC) as between subjects factor. There was a significant main effect of group, $F(2,117) = 33.587$, $p < 0.001$, and main effect of importance of the situation, $F(1,117) = 19.254$, $p < 0.001$. Both effects were modified by a significant interaction between importance and group, $F(2,117) = 5.294$, $p < 0.01$. In order to investigate the nature of the two-way interaction, Helmert contrasts were calculated separately for each wash. For the compulsive wash, there was a significant overall difference between groups, $F(2,119) = 38.054$, $p < 0.001$, with $M = 60.61$ (S.D. = 34.91), $M = 20.18$ (S.D. = 24.43), and $M = 14.40$ (S.D. = 17.7) for W, NW and HC, respectively. The mean difference of 44.46 (97.5% confidence interval: 32.78–56.213) between washers and the two controls groups was significant, $p < 0.001$. The mean difference of 5.99 between the control groups was not significant, $p = 0.298$. For the non-compulsive wash, the overall differences between groups were also significant, $F(2,119) = 17.393$, $p < 0.001$, with $M = 40.27$ (S.D. = 34.56), $M = 16.83$ (S.D. = 22.09) and $M = 8.45$ (S.D. = 14.25), for W, NW and HC, respectively, indicating that the difference between washers and the two control groups was significant, $p < 0.001$, albeit smaller than for the compulsive wash (mean difference of 27.63; 97.5% confidence interval: 16.58–38.68). Non-washing obsessionals and normal controls were not significantly different from each other (mean difference of 8.39; $p = 0.124$).

3.3. Discussion of the washing inventory

The main analysis supports the predictions concerning the use of subjective and objective criteria by obsessionals. It indicates that obsessional washers regard subjective but not objective criteria to be more important for the decision when to stop washing than other obsessionals or healthy controls. These results are consistent with the results of the semi-structured interview in terms of the importance of subjective criteria for the termination of a *compulsive* wash.

We also expected a specific effect for quantity of criteria, i.e. that washers use more criteria than controls. However, we found that *both* obsessional groups used more criteria than healthy controls in *both* types of wash. This suggests that the use of multiple criteria may be a general strategy in obsessionals.

Although the study did not include an objective measure of controlled processing, the subjective view of

this subject is that – consistent with predictions – the decision when to stop a compulsive wash is much more deliberate and involves more conscious effort for washers than for the other participants.

How does the importance of the situation influence the decision making process? As predicted, the nature of the situation influenced the use of elevated evidence requirements. All type of criteria were regarded as more important for the compulsive/most needed wash than for the non-compulsive/least needed wash to the same extent in all groups; all participants used more criteria for the most needed wash than for the last needed wash. Finally, the importance of the situation interacted with group differences between washers and both controls on the perceived automaticity of the decision. In the least needed situation, group differences were smaller than for the most needed wash. These findings point towards a ‘normal’ (i.e. common in most people) employment of elevated evidence requirements, which are used by everyone given the right circumstances. The use of elevated evidence requirements is flexible and adaptable to the situation, yet operated on a higher level for obsessionals than for healthy individuals. There is some indication that perceived automaticity of the decision varies more with the importance of the situation than type and number of criteria, since importance of the situation interacted with group for the variable “automaticity”. This could point towards the importance of a deliberate decision making process at the end of a compulsive action which needs to be addressed in future studies.

The inventory study has some limitations. Order and carry over effects could have compromised ecological validity. Participants were always asked about the compulsive or most needed wash first thereby prompting responses that were again easily accessible for the non-compulsive or least needed wash. We have no psychometric data for the validity of the items we used. However, this was a first attempt to assess subjective and objective reasons and the answers from the semi-structured interview confirm our more structured assessment of subjective and objective data in the washing inventories. Finally, memory biases could have influenced results since participants were asked about past washes. Recollection was enhanced by asking participants to form a clear image of the wash before being interviewed about it, yet it remains possible that recollection was biased by present ideas about how the compulsion was terminated. We therefore conducted a behavioral test in which we asked participants “on line” about their reasons for stopping a wash.

4. Behavioral test

Washing was elicited in the laboratory in order to evaluate aspects of elevated evidence requirements “on line”, and to seek to relate this strategy to observational measures of the wash itself. All participants washed at baseline and after rubbing wax into their hands as part of a standardized procedure. Self-report measures (the washing inventories) of the importance of various stopping criteria were taken. In order to test whether counterproductive stopping criteria are related to the length of the wash, the wash was videotaped and timed. Using behavioral measures and examining a standard wash supplements and validates the interview data gathered from the washing interview and inventory study. An experimental manipulation to increase the salience of the wash was also introduced. However, the experimental manipulation was not successful in making participants more anxious and will be described only briefly. The hypotheses were as follows: (i) for obsessional washers, internally referenced criteria will be rated as more important than for non-washing obsessionals and healthy controls. Objective criteria should be at least as important for washers as for controls. (ii) Obsessional washers will have *more* criteria with higher ratings than non-washing obsessionals and non-clinical controls. Finally, it is hypothesized that (iii) the decision making process will be rated as more effortful and conscious for the obsessional washers than the two control groups.

4.1. Methods

For “Participants” and “Instruments” please refer to Part 1, Section 2.1.2 and Part 2, Section 3.1.2.1 “The washing inventory”.

4.1.1. Dependent variables: behavioral

During the experiment participants were asked to wash their hands twice. The first being a baseline wash, the second the experimental wash. Both occasions were videotaped (vision and sound). The time taken to complete the wash was defined as the time from the first act of the wash (usually turning the tap on) to putting the towel down (coding rules of the observable units and definition and computation of length are available from authors on request). The rater was blind to the group and condition the participants belonged to. The reliability of the coding system was checked by having a second blind rater evaluate 21 participants (seven washers, seven non-washing OCDs and seven healthy controls), with intra-class correlation coefficients used to calculate reliability. For length of the wash, the *agreement* intra-class

correlation coefficient was used, which is close to 1 if each object is rated very similarly by both raters. It was 0.9870 for the present study, which indicates an extremely high agreement.

4.1.2. Material and laboratory equipment

The experiment was conducted in a laboratory containing two chairs, a table, a sink and a video-camera. The experimental stimulus was a clean white cotton cloth covered in light yellow floor wax. It was presented in a plastic bag lying next to another new cloth on a small tray. The other cloth was not covered in wax and was used by the experimenter for demonstration purposes. The sink had a mixer tap, a soap dispenser and a holder for paper towels and cotton towels. Participants were informed that a video-camera would be used to record the hand wash, but that it would not be switched on until they were told so. For the experimental manipulation, one of two tins was placed visibly at the rear end of the table. In the contamination condition, the tin was a 50-year-old rusty tin with dents and stains. In the control condition, the tin was an obviously new floor wax tin.

4.1.3. Procedure

The washing experiment was carried out after completing the washing interview. Before the wash started, participants were asked to rate (a) how uncomfortable they felt, (b) how clean they felt, (c) how much they felt the urge to wash, (d) how worried they were about the state of their hands, and (e) how anxious they were *right now*, on a VAS (*pre-baseline* wash ratings). They were then asked to wash their hands. This baseline wash was videotaped. Immediately after the baseline wash, participants were asked to rate the five dimensions described above again (*post-baseline* wash ratings). Afterwards the cloth with the floor polish was presented. Half of the participants received the cloth with the contamination instruction, the other half with the control instruction. In the contamination condition, an old tin was put on the table

and participants were told that the wax came from a local hospital where it had been used around the wards. In the control condition, a new tin was used and participants were told that the wax came from a brand new tin.

Participants were asked to rub the cloth with the floor wax over the hands by copying the experimenter. A delay period of one minute followed. Subjects were asked just to sit there for a minute and then the experiment would go on. After one minute the participants' ratings on the five scales were taken again (*pre-experimental* wash ratings) and they were then asked to wash their hands. The experimenter switched on the camera and left the room. Immediately after the wash participants ratings on the VAS were taken once more (*post experimental* wash ratings) and they were asked to fill in the washing inventory. The five ratings were taken again after a five minutes interval and after a ten minutes interval. After the last ratings, participants were thanked for their participation and debriefed.

4.2. Results

The overall type one error probability was set at $\alpha = 0.05$. The Tukey Honestly Significant Difference Test (Tukey HSD) was employed in multiple comparisons of between subject factors in parametric procedures. Post hoc univariate ANOVAS and post hoc tests for non-parametric procedures were Bonferroni corrected.

4.2.1. Subjective versus objective criteria

The variables 'subjective criteria' and 'objective criteria' were subjected to a mixed model ANOVA with type of criterion (subjective versus objective) as within subject factor and group (W versus NW versus HC) and experimental condition (neutral versus contamination) as between subject factors. Table 5 shows means and standard deviations divided by group and condition. There was a significant main effect of group, $F(2,115) = 10.156$, $p < 0.0001$, and a significant interaction between group and type of criterion,

Table 5

Standardized wash: importance of subjective and objective criteria on VAS (0 = not at all important; 100 = extremely important) and total number of criteria, means and standard deviations

	W		NW		HC	
	Neutral	Contamination	Neutral	Contamination	Neutral	Contamination
Subjective	78.95 (20.59)	80.42 (21.77)	49.50 (24.92)	51.31 (29.51)	50.34 (24.81)	49.76 (26.48)
Objective	66.97 (27.62)	61.11 (28.40)	53.25 (31.38)	53.33 (32.95)	51.59 (26.34)	58.33 (29.17)
Number>0	7.52 (2.06)	7.83 (1.42)	7.50 (1.40)	7.05 (2.08)	6.50 (2.02)	6.19 (2.48)

Note: Means are shown with standard deviations in parentheses. Maximum number of criteria was eight.

$F(2,115) = 4.891, p < 0.01$. The three way interaction between group, type and experimental condition was not significant, $F(2,115) = 0.544, p = 0.582$, nor was any other effect.

The two-way interaction was investigated further by analyzing objective and subjective criteria separately, with the data pooled over experimental conditions. The ANOVAs with the between subjects factor ‘group’ (W versus NW versus HC) resulted in a significant group effect for the subjective criteria, $F(2,119) = 17.421, p < 0.0001$, but not for the objective criteria, $F(2,118) = 1.547, p = 0.217$. Tukey HSD tests showed that the washers rated subjective criteria as significantly more important than non-washing obsessionals and healthy controls, $p < 0.0001$, the two control groups were not significantly different from each other, $p = 0.997$; with $M = 79.69, 50.41$ and $M = 50.05$ for washers, other OCD and healthy controls, respectively. This means that obsessional washers rated the use of subjective criteria for the experimental wash as more important than the two control groups, whereas there was no difference for the objective criteria.

4.2.2. Number of dimensions considered

The number of criteria being considered were analyzed using a factorial ANOVA with group (W versus NW versus HC) and experimental conditions (neutral versus contamination) as between subjects factors. The maximum number of criteria that could be considered was eight. Means and standard deviations are displayed in Table 5. A significant effect of group, $F(2,121) = 4.925, p < 0.01$, was accounted for by the difference between the washers and non-clinical controls (Tukey $p < 0.01$). Washers and non-washing obsessionals were not significantly different from each other ($p = 0.632$), with the two control groups showing a trend only ($p = 0.086$), with $M = 7.7$ and 7.3 , and $M = 6.3$ for W, NW, and HC, respectively. The experimental condition was not significant, $F(1,121) = 0.179, p = 0.673$, nor was the group by condition interaction, $F(2,120) = 0.405, p = 0.668$. In sum, both obsessional groups used more criteria than the healthy control group.

4.2.3. Controlled processing

‘Controlled processing’ was subjected to a factorial ANOVA with group (W versus NW versus HC) and experimental condition (neutral versus contamination) as between subjects factors. Automaticity was significantly different between groups, $F(2,117) = 12.812, p < 0.001$. Multiple comparison Tukey HSD tests demonstrated that washers judged the decision making

process as more strategic than non-washing obsessionals, $p < 0.001$, and healthy controls, $p < 0.001$. The two control groups were not significantly different from each other, $p = 0.693$; with $M = 56.39$ (S.D. = 35.79), $M = 29.38$ (S.D. = 26.6) and $M = 23.86$ (S.D. = 26.6), for W, NW, and HC, respectively). The experimental condition did not have a significant effect, $F(1,117) = 0.52, p = 0.821$, and nor did the group by condition interaction, $F(2,117) = 1.033, p = 0.359$.

4.2.4. Length of the wash

Length of the washes carried out during the experimental session was analyzed using a mixed model ANOVA with phase of the wash (baseline versus experimental) as within subject factor and two between subjects factors: group (W versus NW versus HC) and experimental condition (neutral versus contamination). Due to technical problems, data were missing for five washers and six other obsessionals for length of the wash. Table 6 presents means and standard deviations of length of the washes for each cell. For length of the wash, there was a significant effect of phase, $F(1,103) = 71.167, p < 0.001$; with $M = 49.23$ s (S.D. = 31.05) and $M = 75.71$ s (S.D. = 42.50) for baseline and experimental wash, respectively, and group, $F(2,103) = 7.071, p < 0.01$; with $M = 81.22$ s (S.D. = 49.89), $M = 54.62$ s (S.D. = 27.71) and $M = 51.56$ s (S.D. = 20.78) for W, NW and HC, respectively. Multiple comparison Tukey HSD tests demonstrated that washers washed significantly longer than other obsessionals and healthy controls, $p < 0.05$. Control groups were not significantly different from each other, $p = 0.922$. Experimental condition was not significant, $F(1,103) = 0.195, p = 0.660$, nor were any of the interactions, viz, phase by group, $F(2,103) = 2.536, p = 0.084$; phase by experimental condition, $F(1,103) = 0.860, p = 0.356$; group by experimental condition, $F(2,103) = 0.536, p = 0.587$; phase by group by experimental condition, $F(2,103) = 1.277, p = 0.283$.

Table 6
Length of baseline and experimental wash, means and standard deviations

Length in second	M (S.D.)		
	W (N = 33)	NW (N = 35)	HC (N = 43)
Baseline			
Neutral	58.19 (51.69)	46.00 (22.50)	43.32 (21.51)
Contamination	67.93 (56.95)	39.58 (17.62)	40.38 (16.03)
Experimental wash			
Neutral	92.44 (50.52)	60.44 (34.41)	65.32 (29.03)
Contamination	106.33 (74.89)	72.47 (47.67)	57.24 (20.30)

Table 7
Correlations between length of experimental wash and main dependent variables

	Subjective criteria	Objective criteria	Number >0	Automaticity
Length				
All (N = 111)	0.363 ^a	0.425 ^a	0.307 ^a	0.477 ^a
W (N = 34)	0.455 ^a	0.504 ^a	0.351	0.582 ^a
NW (N = 36)	0.274	0.374 ^a	0.268	0.386 ^a
HC (N = 43)	0.145	0.367 ^a	0.118	0.318 ^a

Note: Spearman rank correlation coefficients are shown.

^a Correlation is significant at the 0.01 level.

4.2.5. Correlational analysis

Spearman rank correlations were used. α was set at 0.01. Table 7 presents the Spearman rank correlations. Over all participants there is a positive relationship between the behavioral measures of time and elevated evidence requirements. Length of the wash was correlated with the use of subjective criteria (Spearman rank correlation = 0.363), the number of criteria being considered (Spearman rank correlation = 0.307), and controlled processing mode (Spearman rank correlation = 0.477). This means that the more criteria are being considered, the more important subjective criteria are, and the more controlled the decision making process is, the longer the wash is.

5. Overall discussion

This study is the first investigating the use of elevated evidence requirements in obsessional washers, patients with other obsessional problems and healthy controls. The investigation involved three connected studies: a standardized interview with open ended questions, a questionnaire study which assessed the importance of subjective and objective criteria quantitatively and finally, a behavioral test which assessed elevated evidence requirements immediately after a standardized laboratory wash.

The prediction that subjective criteria with an internal reference point such as a “feeling of rightness” are important for the termination of a compulsive wash was largely supported by the data. Consistently across all studies we found that subjective criteria were rated as having an important role in the termination of a compulsive wash. In the open-ended interview obsessional washers reported an internal feeling of rightness more often than non-washing obsessionals or healthy controls. In the questionnaire study, subjective criteria were rated as more important by the washers than by the two control groups. Laboratory data from an actual wash confirmed these findings. This is also consistent with other work that has investigated the use of internal

feelings (Richards, 1995) or “Not just right experiences” (Coles et al., 2003) in obsessional patients or healthy controls.

We predicted that obsessional washers would use more criteria than controls before reaching a decision about the termination of a wash. However, we found that both obsessional groups (washer and other obsessional patients) used more criteria than healthy controls. This points towards a *general obsessional strategy* even in situations which are not characterized by strong obsessional fears. Results are consistent with early studies which found that obsessional patients required more trials than psychiatric controls before reaching a decision in a visual discrimination task (Walker, 1967) or a tone discrimination task (Milner, Beech, & Walker, 1971).

Our third hypothesis was that obsessional washers would perceive the decision making process as more deliberate and as requiring more mental effort than controls. Although the study did not include an objective measure of controlled processing or automaticity of the decision, we found that – consistent with predictions – the decision when to stop a compulsive wash was more deliberate and involves more conscious effort for washers than for other participants in the questionnaire data as well as in the laboratory wash. Future studies might include more refined measures of automaticity in order to see whether results hold up for more objective measures.

Behavioral measures of the length of the wash supplement the subjective ratings. All aspects of elevated evidence requirements were positively correlated with the length of the laboratory wash which is consistent with our idea that the use of elevated evidence requirements extends an obsessional wash.

Finally, we predicted that the decision making process and the use of elevated evidence requirements depend on the perceived personal significance of the situation. Consistent with this hypothesis we found that all type of criteria were regarded as more important

for the compulsive/most needed wash than for the non-compulsive/least needed wash, and that all participants used more criteria for the most needed wash than for the last needed wash; additionally, perceived automaticity varied with the importance of the situation. Whereas the questionnaire data suggests a clear interpretation of the importance of the situation and supports a “normalizing” approach of elevated evidence requirements – i.e. elevated evidence requirements are used by everyone given the right circumstances – one finding does not fit entirely into the picture. Interview data suggests that the use of subjective criteria is also important for the termination of a non-compulsive wash in washers.

Our data is partly consistent with the security motivation model by [Szechtman and Woody \(2004\)](#). The study confirms the role of an internal state (such as a “feeling of knowing”) as being crucial for the termination of a compulsive washing action which is in line with the above model. However, we also found evidence for the importance of external cues that are taken into consideration before stopping a wash. The decision making process was also described as being a conscious process made with deliberate effort. These two findings are not consistent with the model by [Szechtman and Woody](#) which assumes that the “feeling of knowing” is based on internal cues and operates automatically. It remains to be seen how the use of elevated evidence requirements and other strategies such as memory based strategies or internal rules (e.g. a certain amount of time or number of times) interact and eventually lead to the termination of a wash. The interview data suggests that predetermined rules and memory-based cues may be crucial to the termination of a compulsive wash although they were mentioned less often than subjective criteria. In our view this finding is consistent with the notion of elevated evidence requirements since we assume that a number of criteria – internal, external, memory based, etc. – are considered during the wash but that in the end it is a (combined?) feeling of “rightness” or “completeness” which results in the end of the wash as the “top level”

The study has several limitations. A problem of the current design is that participants were not directly asked how many criteria they had to consider before they stopped washing. Instead, numbers were inferred from the importance ratings based on an arbitrary cut off point. Since obsessional problems have been associated with perfectionist tendencies (e.g. [Rasmussen & Eisen, 1989](#)), it may be that obsessional patients were over inclusive. If that were the case, the present

ratings resulted in an overinclusion of criteria for all obsessional patients. When asked about it directly, they might have given fewer reasons. Future studies would have to assess the number of criteria directly and clarify whether it is influenced by the method of assessment.

Correlational data do not establish causal connections between elevated evidence requirements and the length of the compulsion. Although the positive correlations between components of elevated evidence requirements and length of the wash are consistent with the proposed causal link in the theory of elevated evidence requirements, future studies would have to systematically vary the use of elevated evidence requirements and observe the effects on the length of the wash to clarify this issue. The interview and inventory studies required a recollection of a past wash, so the results are vulnerable to memory effects. Recollection was enhanced by asking participants to form a clear image of the wash or check before being interviewed about it, yet it remains possible that recollection was biased by present ideas about how the compulsion was terminated. The subsequent behavioral test addressed this issue by asking participants to rate criteria immediately after a standardized wash in the laboratory. This raises the possibility of order effects, as the laboratory wash was preceded by the interview about the past wash. Participants’ ratings of this current wash may have been influenced by their recall of the “historical” wash.

Aspects of the validity of the study may have been affected by the way in which the concepts subjective and objective criteria were operationalized. New strategies had to be devised from scratch as there was no previous work on this issue. This issue is presently under investigation by our group. It can be argued that differences between washers and control obsessionals emerged because washers were more severely affected (they had higher OCI total score than obsessionals). This is a possible interpretation of the results, but in the authors view insufficient to explain the consistent findings between different parts of the study.

Finally, one could argue that positive feelings may have been associated with anxiety reduction at the end of the wash as the main motivating factor. This question was explicitly addressed in a follow up study investigating stopping criteria in obsessional checkers ([Wahl & Salkovskis, submitted for publication](#)). There was a tendency for checkers to experience positive mood changes at the end of the check more frequently than controls; however, these results were not significant.

There was, however, a low but significant correlation between mood changes and “it felt right” (corr. $r = 0.249$). Correlation between the variables “it felt right” and “reduction in anxiety”, on the other hand, were not significant ($r = 0.147$).

Taken together, the present findings are largely consistent with the predictions of the elevated evidence requirements modification to cognitive theories of OCD. Clinical implications include a tailored treatment for patients who rely mainly on some sort of internally referenced criteria. Dysfunctional beliefs about the appropriateness of those criteria can be challenged in behavioral experiments by purposefully alternating strategies to stop washing and observing the short and long term effects.

Appendix A. Categories used for group comparison statistics in study 1

Main categories

1. *External, perception*: Comprising statements that were based on a sensory perception
 - tactile information;
 - visual information;
 - olfactory information;
 - decision was based entirely on soap being used or being rinsed off (it was not clear which of the senses was employed to make the decision);
 - based on sensi-motor information as a consequence of an action, such as ‘rinsed hands’.
2. *Internal, feeling/mood*: Indication that a *general* feeling of rightness/completeness, satisfaction, cleanliness, mood or an epistemological sense of ‘just knowing’ was considered for the decision to stop
 - feels right;
 - feels clean;
 - feels satisfied;
 - knowing;
 - mood.
3. *Internal rules, memory*:
 - Having a clear picture or memory of the wash;
 - having absolutely no doubts that hands had been washed;
 - time or number of times.
4. *Internal, effort*: Indication that a certain amount of effort had to be put into the wash
 - tried hard enough;
 - physically/mentally exhausted by the wash.

Appendix B. The questions used in the washing inventory

Now answer all the questions below thinking of that particular wash.

On this questionnaire is a list of reasons that people might give for deciding to stop to washing. Please read each one carefully. Then rate on a scale from 0 to 100, using the rating scales below each statement, how important you think each one was to *you* in deciding to stop wash *on that particular occasion*. Some may be more relevant to you than others but please rate each statement:

- (1) How important was it that you knew from the way that you felt about things that they were right or complete?
 - (2) How important was it that you had tried hard enough?
 - (3) How important was it that you were physically or mentally exhausted?
 - (4) How important was it that your hands looked completely clean?
 - (5) How important was it that your hands did not feel sticky or waxy?
- Additional items not reported in this study:
- (6) How important was it that you had no doubts that your hands were clean?
 - (7) How important was it that you had checked your hands thoroughly?
 - (8) How important was it that you had done things for as long or for as many times as you felt that you needed to?
 - (9) How important was it that someone else told you that it was OK to stop at that point?

Now think back at the moment when you were about to complete the washing. . .

To what extent was the decision to stop washing something you were consciously deciding?

To what extent was the decision that you made to stop washing made with deliberate effort?

References

- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100.
- Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 18, 561–571.
- Bobes, J., Gonzalez, M. P., Bascaran, M. T., Arango, C., Saiz, P. A., & Bousono, M. (2001). Quality of life and disability in patients with

- obsessive–compulsive disorder. *European Psychiatry*, 16(4), 239–245.
- Coles, M., Frost, R. O., Heimberg, R. G., & Rheume, J. (2003). Not just right experiences”: Perfectionism, obsessive–compulsive features and general psychopathology. *Behaviour Research and Therapy*, 41, 681–700.
- Davey, G. C. L., Startup, H. M., Zara, A., MacDonald, C. B., & Field, A. P. (2003). The perseveration of checking thoughts and mood-as-input hypothesis. *Journal of Behavior Therapy and Experimental Psychiatry*, 34(2), 141–160.
- Ecker, W., & Engelkamp, J. (1995). Memory for actions in obsessive–compulsive disorder. *Behavioural and Cognitive Psychotherapy*, 23: 349–371.
- First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., et al. (1995). The structured clinical interview for DSM-III-R personality disorders (SCID-II): II. Multi-site test-retest reliability study. *Journal of Personality Disorders*, 9(2), 92–104.
- Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a new obsessive–compulsive disorder scale: The Obsessive–compulsive Inventory. *Psychological Assessment*, 10(3), 206–214.
- Freeston, M. H., Rheume, J., & Ladouceur, R. (1996). Correcting faulty appraisals of obsessional thoughts. *Behaviour Research and Therapy*, 34(5–6), 433–446.
- Koran, L. M. (2000). Quality of life in obsessive–compulsive disorder. *Psychiatric Clinics of North America*, 23(3), 509–517.
- Koran, L. M., Thienemann, M. L., & Davenport, R. (1996). Quality of life for patients with obsessive–compulsive disorder. *American Journal of Psychiatry*, 153(6), 783–788.
- MacDonald, C. B., & Davey, G. C. L. (2005). A mood-as-input account of perseverative checking: The relationship between stop rules, mood and confidence in having checked successfully. *Behaviour Research and Therapy*, 43, 69–91.
- McNally, R. J. (1995). Automaticity and the anxiety disorders. *Behaviour Research and Therapy*, 33, 747–754.
- Milner, A., Beech, H. R., & Walker, V. J. (1971). Decision processes an obsessional behaviour. *British Journal of Social and Clinical Psychology*, 10, 88–89.
- Parkin, R. (1997). Obsessive–compulsive disorder in adults. *International Review of Psychiatry*, 9(1), 73–82.
- Rachman, S. (1998). A cognitive theory of obsessions: Elaborations. *Behaviour Research and Therapy*, 36(4), 385–401.
- Rachman, S. J., & Hodgson, R. J. (1980). *Obsessions and compulsions*. New Jersey: Prentice Hall.
- Rachman, S., & Shafran, R. (1998). Cognitive and behavioral features of obsessive–compulsive disorder. In R.P. Swinson, M.M. Antony, S. Rachman, M.A. Richter (Eds.), *Obsessive–compulsive disorder: Theory, research, and treatment*.
- Radomsky, A. S., Rachman, S., & Hammond, D. (2001). Memory bias, confidence and responsibility in compulsive checking. *Behaviour Research and Therapy*, 39(7), 813–822.
- Rasmussen, S. A., & Eisen, J. L. (1989). Clinical features and phenomenology of obsessive compulsive disorder. *Psychiatric Annals*, 19(2), 67–73.
- Richards, H. C. (1995). The cognitive phenomenology of OCD repeated rituals. *Paper presented at the world congress of behavioural and cognitive therapies*.
- Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, 23, 571–583.
- Salkovskis, P. M. (1999). Understanding and treating obsessive–compulsive disorder. *Behaviour Research and Therapy*, 37, s29–s52.
- Salkovskis, P., Thorpe, S., Wahl, K., Wroe, A., & Forrester, E. (2003). Neutralizing increases discomfort associated with obsessional thoughts: An experimental study with obsessional patients. *Journal of Abnormal Psychology*, 112(4), 709–715.
- Spielberger, C. D. (1983). *Manual for the state-trait inventory STAI (form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Sookman, D., & Pinard, G. (2002). Overestimation of threat and intolerance of uncertainty in obsessive compulsive disorder. In: R. O. Frost & G. Steketee (Eds.), *Cognitive approaches to obsessions and compulsions: Theory, assessment and treatment*. Amsterdam: Pergamon.
- Szechtman, H., & Woody, E. (2004). Obsessive–compulsive disorder as a disturbance of security motivation. *Psychological Review*, 111, 111–127.
- Tallis, F., Eysenck, M., & Mathews, A. (1991). Elevated evidence requirements and worry. *Personality and Individual Differences*, 12(1), 21–27.
- Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., Amir, N., Street, G. P., & Foa, E. B. (2001). Memory and memory confidence in obsessive–compulsive disorder. *Behaviour Research and Therapy*, 39, 913–927.
- Wahl, K. U., & Salkovskis, P. M. (submitted for publication). ‘It feels right’ and other cognitive factors in the decision to stop obsessional behaviour.
- Walker, V. J. (1967). *An investigation of ritualistic behaviour in obsessional patients*. Unpublished Ph.D. Thesis. Institute of Psychiatry, University of London.
- Wirtz, M., & Caspar, F. (1994). *Ratings: Bestimmung und Verbesserung der Übereinstimmung und der Reliabilität*. Forschungsberichte aus dem psychologischen Institut der Universität Bern.
- Zaudig, M. (2006). Behandlung von Zwangsstörungen. In: H.-J. Möller (Ed.), *Therapie psychischer Erkrankungen*. Stuttgart: Thieme Verlag.
- Zitterl, W., Urban, C., Linzmayer, L., Aigner, M., Demal, U., Semler, B., et al. (2001). Memory deficits in patients with DSM-IV obsessive–compulsive disorder. *Psychopathology*, 34(3), 113–117.