Original Article

The Assessment of Thought Fusion Beliefs and Beliefs About Rituals

Psychometric Properties of the Thought Fusion Instrument and Beliefs About Rituals Inventory

Kim Melchior[1,2], Ingmar H. A. Franken[2], Richard Vuijk[3], Vivian Peerbooms[1,2], and Colin van der Heiden[1,2]

[1]Outpatient Treatment Centre PsyQ, Rotterdam, The Netherlands
[2]Department of Psychology, University Rotterdam, The Netherlands
[3]Sarr Autism Rotterdam, The Netherlands

Abstract. *Background:* According to the metacognitive model, two domains of metacognitive beliefs play a role in the development and maintenance of obsessive-compulsive disorder (OCD). The Thought Fusion Instrument (TFI) has been developed to measure metacognitive beliefs about the significance of intrusive thoughts. The Beliefs About Rituals Inventory (BARI) assesses metacognitive beliefs about the necessity of performing ritual behaviors. Studies assessing the psychometric properties of the TFI and BARI are scarce. There are no studies assessing the factor structure of the TFI and the BARI. *Methods:* In this study, we assessed the psychometric properties of the TFI and the BARI in nonclinical ($n = 141$) and clinical populations (OCD [$n = 60$], anxiety disorder [$n = 30$], and autism spectrum disorder [$n = 50$]). In the nonclinical population, the factor structure is also explored. *Results:* For both the TFI and the BARI, an explorative factor analysis revealed a one-factor solution, which now needs further exploration using confirmative approaches. The internal consistency appeared good, and they had a moderate test–retest reliability. Convergent and divergent validity of the instruments appeared sufficient, but more research is required to draw firm conclusions. The criterion validity turned out to be moderate for the BARI but low for the TFI in measuring OCD-specific metacognitions. *Conclusions:* Based on the explorative factor analysis, we hypothesize the TFI and the BARI to measure a single-factor construct. The current study shows that the TFI and the BARI are potentially suitable questionnaires to assess metacognitions in clinical and nonclinical populations. More research is required before clear recommendations can be made for the utility and use in clinical practice.

Keywords: obsessive-compulsive disorder, fusion beliefs, beliefs about rituals, validation, factor structure, TFI, BARI

Obsessive-compulsive disorder (OCD) is characterized by distressing and intrusive recurring thoughts or urges, the obsessions, that individuals try to neutralize by engaging in repetitive actions, called compulsions (American Psychiatric Association, 2013). Unwanted intrusions are considered almost universal and occur in healthy individuals just as they do in OCD patients (Muris et al., 1997). OCD, however, occurs in only 1.6% of the general population (Kessler et al., 2005). The metacognitive model (Wells, 1997) suggests that metacognitive beliefs contribute to the development of normal occurring intrusions into OCD.

In the metacognitive model of OCD, two types of metacognitions are distinguished (Wells, 2009). The first type concerns beliefs about the meaning and power of intrusive thoughts, the fusion beliefs. This concept contains three different subdomains, Thought Action Fusion

(TAF), Thought Event Fusion (TEF), and Thought Object Fusion (TOF). TAF refers to the belief that certain thoughts will lead irrevocably to carrying out an act (e.g., "If I think about harming myself, I will do so"). TEF, the second subdomain, implies that solely thinking about an event is responsible for causing it in the future or is a sign that the event actually did happen or is happening (e.g., "If I think of an unpleasant event, then it will happen"). Finally, TOF refers to the belief that it is possible to transfer thoughts and feelings to objects (e.g., "My memories can be passed into objects"). According to the metacognitive model, thought fusion beliefs are activated by a trigger (a normally occurring intrusion, urge, or doubt) and consequently cause OCD patients to negatively appraise these triggers as overly important or even potentially dangerous, leading to phenomena like worrying and thought monitoring. Consequently, OCD patients

experience distress. In turn, beliefs about the necessity of performing rituals are activated (e.g., "I have to perform my rituals, otherwise I will never have peace of mind again"), the second class of metacognitive beliefs in OCD, leading to the performance of ritual behaviors to get rid of the intrusive thought and experienced distress. As a result, because of the feeling of safety and rest after performing ritual behaviors, the positive metacognitive beliefs about the usefulness and necessity of the neutralizing rituals are exacerbated (Fisher & Wells, 2005). The metacognitive model is depicted in Figure 1.

It is assumed that all people experience some degree of fusion beliefs and beliefs about rituals, but there is empirical evidence that a high degree of these beliefs contribute to the development and maintenance of OCD (for an overview, see Fisher, 2009). Cross-sectional studies have shown that both fusion beliefs and beliefs about the necessity of performing rituals are significantly associated with symptoms of OCD (Gwilliam et al., 2004; McNicol & Wells, 2012; Myers & Wells, 2005). Moreover, support for a causal relationship between fusion beliefs and OCD symptomatology can be found in experimental research. Fisher and Wells (2005) manipulated thought-fusion in OCD patients, which led to changes in their obsessive-compulsive symptoms. Also, Myers and Wells (2013) found that experimental manipulation of fusion beliefs leads to an accretion of intrusive experiences. Myers et al. (2009) conducted a study to directly investigate the metacognitive model in a student sample and found that when they controlled for cognitive beliefs, worry, and threat, fusion beliefs and beliefs about rituals remained as independent predictors of obsessive-compulsive symptoms. In contrast, cognitive factors such as responsibility and perfectionism that have been frequently associated

with OCD (Frost & Steketee, 1997) did not turn out to be independent predictors of the symptomatology of OCD. This finding was replicated by Hansmeier et al. (2016) in a large sample of patients with OCD and also in a clinical sample by Myers et al. (2017). Finally, support for the importance of metacognitive beliefs in OCD comes from therapy effect studies. Metacognitive therapy for OCD tailored to both fusion beliefs and beliefs about the necessity of performing rituals leads to significant reductions in OCD symptoms (Fisher & Wells, 2008; Rees & van Koesveld, 2008; Van der Heiden et al., 2016). To summarize, there is empirical evidence that a high degree of fusion beliefs and beliefs about rituals contribute to the development and maintenance of OCD.

Since there is empirical evidence for the importance of metacognitive beliefs in OCD, there is a demand for well-validated instruments that measure these constructs in both normal and clinical populations, for research purposes and for individual assessment. For the assessment of fusion beliefs, multiple scales have been developed containing only items about TAF, for example, the TAF scale (Shafran et al., 1996), and as such, do not cover the full range of beliefs about thoughts across the three fusion domains that are considered relevant for OCD according to the metacognitive model. Also, in comparative studies, no differences between OCD and other clinical groups (e.g., anxiety disorders and depression) were found regarding fusion beliefs, as measured with the TAF scale (Abramowitz et al., 2003; Hansmeier et al., 2016; Rassin, Diepstraten et al., 2001; Solem et al., 2010). This raises questions about the disorder-specificity of TAF as a single construct and underlines the need for instruments that measure the whole range of OCD-specific metacognitive beliefs as disorder-specific parameter.

The Thought Fusion Instrument (TFI; Wells et al., 2001) was designed to measure fusion beliefs across fusion domains that are considered relevant in the metacognitive formulation and treatment of OCD. The questionnaire consists of 14 items to assess beliefs about thoughts on a single scale, and no separate subscales are formulated. Although it can be hypothesized that the three classes of fusion beliefs as described above, TEF, TAF, and TOF, are represented in the TFI, this has not been investigated yet. There are some preliminary investigations addressing the psychometric properties of the TFI. Gwilliam et al. (2004) found a Cronbach's α for the overall scale of .89, suggesting good internal consistency. Positive correlations between the TFI and theoretically related measurements, for example, the Meta-Cognition Questionnaire (MCQ; Cartwright-Hatton & Wells, 1997), a well-validated self-report scale for measuring metacognitive parameters in psychopathology in general, and positive correlations with obsessive-compulsive symptoms suggest good convergent
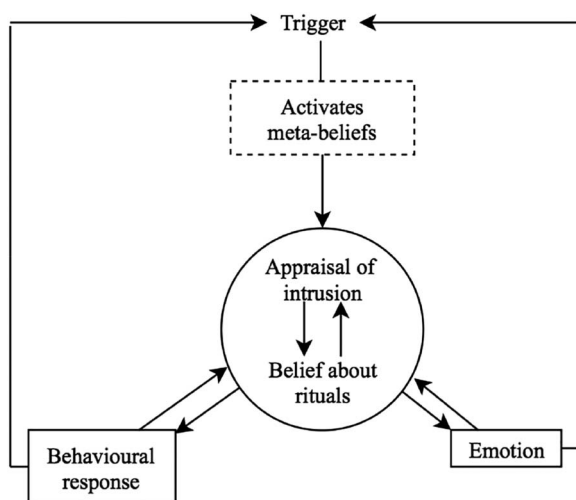

Figure 1. Metacognitive model for OCD (Wells, 1997).

validity. However, these investigations were only carried out in a general population. Also, there are no studies available addressing the factor structure and the test–retest reliability of the TFI.

The Beliefs About Rituals Inventory (BARI; McNicol & Wells, 2012) is a 12-item unidimensional measurement to assess beliefs about the necessity of performing rituals. McNicol and Wells describe a Cronbach's α of .86 in their paper in which they introduce the BARI. They also reported preliminary evidence for the convergent and divergent validity since they found a higher correlation between the BARI score and the theoretically related Obsessive-Compulsive Inventory (Foa et al., 1998) than with a measure of worry (Penn State Worry Questionnaire; Meyer et al., 1990). No further studies have addressed the psychometric properties of the BARI so far, for example, in clinical groups. There are also no studies available addressing the factor structure or the test–retest reliability of the BARI.

In sum, although there are some preliminary investigations into the psychometric properties of the TFI and the BARI, previous studies leave room for further research. Since the TFI and the BARI are of value for research purposes, for example, for considering them as mechanisms of change in treatment studies, more research into the psychometric properties in both normal and clinical populations is necessary, especially into the factor structure and test–retest reliability. Second, since the questionnaires are also of value for clinicians, for example, for individual assessment in psychological treatment for patients with OCD, investigation of the criterion validity of the measurements is also necessary.

The goal of the present study was to evaluate the psychometric properties of the Dutch versions of both the TFI and the BARI (identical to the English versions). Three substudies were set up to accomplish this goal. In the first study, we used a nonclinical sample. First, we employed an explorative factor analysis on the data. We chose for an explorative factor analysis, rather than confirmative factor analysis, since the factor structures of both the TFI and the BARI have not yet been investigated. Second, the reliability coefficients of the scales were determined using both Cronbach's α and McDonald's ω coefficients. Finally, the convergent validity and divergent validity were assessed by studying the correlation coefficients between the scales and other constructs. To investigate the convergent validity, we chose theoretically related constructs, to wit, obsessive compulsive symptoms (for which we use a measurement that assesses both obsessions and repetitive behaviors), a general measurement for metacognitive beliefs in psychopathology, and intolerance of uncertainty, which is also proposed as a disorder-specific cognitive domain involved in the development and maintenance of

OCD (Freeston et al., 1994). The correlation between the TFI and the BARI is also an indicator of the convergent validity evidence. Next, the correlation coefficient with a divergent construct (depressive symptoms) is calculated. Although there is a high comorbidity between OCD and depression, there is no theoretical reason why OCD-specific metacognitive beliefs are directly associated with depressive symptoms. In a second study, a subpopulation of study one was asked to complete the questionnaires a second time after 4–6 weeks to determine the test–retest reliability coefficient of the scales. In the third study, we explore the validity and reliability coefficients of the TFI and BARI in an OCD sample. We also used two other clinical samples and a nonclinical control group to explore the criterion validity of the measurements by making a comparison between OCD and other anxiety disorders, and with patients with autism spectrum disorder (ASD). Since both OCD and ASD are characterized with repetitive thoughts and the urge to perform repetitive actions (McDougle et al., 1995), this offers a stringent test of the disorder-specific contribution of the factors.

Study 1

Reliability, Validity, and Factor Structure in a Normal Population

Participants and Procedure

A group of 141 healthy subjects volunteered in this first study (28.8% male, 71.2% female). We used Shoukri et al. (2004; Table 3) to determine the appropriate sample size needed in our studies, considering the number of questionnaires and measurement points in our study. Participants were recruited using snowball sampling (authors asked three relatives to complete the questionnaires and also asked these three participants to ask three relatives in their own lives to participate, and so on). The mean age of the participants was 46.7 (SD = 14.42, range 19–90 years). Eighty-nine percent of the sample had completed high school or a higher education level and 75.4% had a full-time job. In the sample, 43.9% was married and another 25.2% lived together. A subsample of these participants (n = 55) were asked to complete the questionnaires twice, a second time after 4–6 weeks to assess the test–retest reliability (see Study 2). There was no reward for participation.

Instruments

Thought Fusion Instrument
The TFI was used for measuring metacognitive beliefs about the influence and meaning of thoughts (Wells et al.,

2001). The TFI consists of 14 items (e.g., "If I think about an unpleasant event, it will make it more likely to happen"), which have to be rated on an 11-point scale from 0 (= *I do not believe this at all*) to 100 (= *I am completely convinced this is true*). A total score is computed by summing up the scores on all 14 items. It is hypothesized that the three classes of fusion beliefs as described earlier, TEF, TAF, and TOF, are represented in the TFI, but this has not been investigated yet. Also, there are no studies available addressing the factor structure of the TFI yet, so in our studies only the total score of the TFI is used.

The original English version of the TFI was translated into Dutch by one of the authors. This Dutch translation was also translated back into the original English language by a native speaker, which showed that the meaning of the items stayed intact.

Beliefs About Rituals Inventory

The BARI was used as a self-report measure for assessing beliefs about the necessity to perform rituals (McNicol & Wells, 2012). It consists of 12 items (e.g., "I must perform my rituals, otherwise I will never find peace again") that are rated on a 4-point scale from 1 (*disagree*) to 4 (*agree*). A total score is computed by summing up the scores on all 12 items. Again, no information about the factor structure is available yet. The original English version of the BARI was translated into Dutch by one of the authors. Again, the Dutch translation was translated back into the original English language by a native speaker, which showed that the meaning of the items stayed intact.

Padua Inventory-Revised

The Padua Inventory-Revised (PI-R) was included as a self-report measure to assess obsessive and compulsive symptoms, and the degree to which these symptoms interfere with daily functioning (Dutch version: Van Oppen et al., 1995; Burns et al., 1996). The PI-R incorporates 41 items, which have to be rated on a 5-point scale from 0 (= *not at all*) to 4 (= *very much*). Both the original PI-R and the Dutch translation were found to have satisfactory reliability, as well as good convergent and divergent validity (Sternberger & Burns, 1990; Van Oppen et al., 1995). The Cronbach's α in the present study was .93.

The Beck Depression Inventory, 2nd Version

The Beck Depression Inventory, 2nd Version (BDI-II), is a widely used self-report instrument for measuring depressive symptoms (Beck et al., 1996). The BDI-II consists of 21 items that have to be rated on a scale from 1 to 4, describing the severity of depressive feelings in the past week. The BDI-II has adequate psychometric properties (Beck et al., 1996), which is also true for the Dutch version

of the BDI-II (Van der Does, 2002). The Cronbach's α in the present study was .91.

Intolerance of Uncertainty Scale

The Intolerance of Uncertainty Scale (IUS) is a 27-item self-report measure to assess emotional, cognitive, and behavioral reactions to uncertain or ambiguous situations (Freeston et al., 1994). Each item is scored on a scale from 1 to 5. A higher score indicates a higher amount of intolerance toward ambiguous situations. The Dutch version of this questionnaire demonstrated good psychometric properties (De Bruin et al., 2006). The Cronbach's α in the present study was .96.

Meta-Cognition Questionnaire

The MCQ-30 measures five domains of metacognition which are central in metacognitive models of psychological disorders (Wells & Cartwright-Hatton, 2004). The MCQ-30 is the shortened and refined version of the original MCQ-65. The five subscales, which have been identified with explorative factor analysis and which have been verified by confirmatory factor analysis in multiple studies, are positive beliefs about worry, negative beliefs about worry, low cognitive confidence, need to control thoughts, and cognitive self-consciousness (Cartwright-Hatton & Wells, 1997). Cronbach's α's for the subscales range from .72 to .93. For our study, only the total score was used. This total score showed adequate psychometric properties in terms of test–retest correlation (.75), and a positive correlation was found between the MCQ-30 and other theoretically appropriate measures (Wells & Cartwright-Hatton, 2004). Although the psychometric properties of the Dutch version of the MCQ-30 are unknown, the Cronbach's α in the present study was good (.92).

Statistical Analysis

Data were obtained using Qualtrics, an online self-report program for surveys. After reading the participant information and providing informed consent, the online survey was e-mailed to the participants. All statistical analyses were performed using SPSS, version 25.

As a first investigation of the factor structure of the questionnaires, we perform an explorative factor analysis using the principal axis factoring (PAF) method and using an oblique rotation to allow for correlations among potential factors. The item-total score correlations for the TFI ranged from .49 to .91, and all appeared significant. For the BARI, the item-total score correlations ranged from .47 to .85, and also all significant. Also, prior to conducting the factor analyses, the distribution of all items was inspected. None of the items were excessively skewed or kurtotic, so no items were dropped from the further analysis. All items were submitted to two separate PAF analyses. Since

multiple researchers have shown that the Kaiser criterion (Kaiser, 1960) can misidentify the number of factors (Gaskin & Happell, 2014) and the scree plot can also be subjective in case of a more gradual slope, we used parallel analysis (PA; Horn, 1965) based on Monte Carlo simulations on randomly generated data. We used 5,000 Monte Carlo simulations to generate a 95th percentile cutoff line that was displayed in a scree plot together with the eigenvalues from our data. Factors above this cutoff line were considered as meaningful and were used in a subsequent explorative factor analysis. The PA for our current study was run in SPSS by using the SPSS code rawpar.spss developed by O'Connor (2000).

To evaluate the reliability of the scales, we used both Cronbach's α and McDonald's ω coefficients. We chose to use both of these since the Cronbach's α is the most widely used method for the evaluation of reliability, which makes comparison with earlier studies possible; however, Cronbach's α has been criticized recently since it is susceptible to problems with inflation and attenuation. Dunn et al. (2014) propose McDonald's ω as a more accurate estimation of reliability. For the calculation of McDonald's omega, the macro for SPSS by Prof. Andrew F. Hayes was used.

To explore the construct validity of the TFI and the BARI, Pearson's correlation coefficients were calculated with both convergent constructs (obsessive-compulsive symptoms [PI-R], a general measurement for metacognitive beliefs [MCQ30], and IUS) and a divergent construct (depressive symptoms [BDI-II]). We hypothesize that the correlation coefficients between the TFI, the BARI, and the related convergent constructs will be substantially larger than the correlation with depressive symptoms (divergent construct).

Results

Factor Structure
Thought Fusion Instrument
First, a PA was performed to identify the number of factors to extract for our PAF. In this procedure, eigenvalues from the raw data were produced, next to eigenvalues of a 95th percentile based on Monte Carlo simulations. A scree plot of this 95th percentile cutoff line overlaid onto the eigenvalues of the raw data is presented in Figure 2. Our PA indicates three factors from the raw data which are above the 95th percentile cutoff line. However, the slope of the scree plot for this three-factor solution dramatically shifts after the first factor. Based on suggestions by Cattell (1966), who states that a dramatic change in the raw data eigenvalues indicates the cutoff for the amount of factors to extract, a one-factor solution can also be justified. This is also in line with the assumption that PA tends
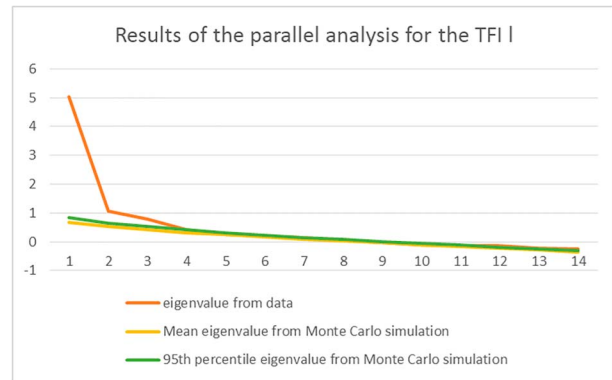

Figure 2. Scree plot of the results of the parallel analysis for the TFI. TFI = Thought Fusion Instrument.

to indicate more factors than warranted and that additional procedures should be used to trim trivial factors (Buja & Eyuboglu, 1992). Nevertheless, based on the PA combined with the assumption that there is a theoretical reason for a three-factor structure, a first inspection of this three-factor solution needs further investigation.

We conducted a PAF with an oblique rotation on the 14 items of the TFI specifying a three-factor solution. The Kaiser–Meyer–Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = .80. Bartlett's test of sphericity also indicated that the correlations between items were sufficiently large for PAF (χ^2 = 894.238, df = 91, p < .001). In our first analysis based on the PA results, a three-factor solution accounts for 58.9% of the variance. Table 1 shows the factor loadings of the 14 items into this three-factor solution. When this three-factor solution was

Table 1. Pattern matrix of the exploratory factor analysis of the TFI for a three-factor solution

Item	Factor 1	Factor 2	Factor 3
1	.037	.907	.002
2	.225	.261	.384
3	−.131	.316	.301
4	−.046	.368	.566
5	−.019	−.045	.472
6	−.014	−.072	.813
7	.122	.113	.473
8	.169	.230	.295
9	.165	.222	.005
10	.020	.037	.556
11	.812	.256	−.097
12	.408	−.087	.475
13	.339	.331	.177
14	.987	−.149	.135

Note. TFI = Thought Fusion Instrument.

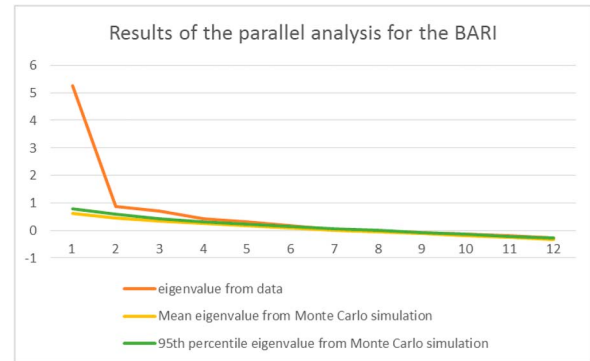**Table 2.** Summary of the exploratory factor analysis of the TFI for a one-factor solution

Item	Factor loadings
1	.645
2	.666
3	.442
4	.688
5	.393
6	.601
7	.605
8	.595
9	.339
10	.543
11	.644
12	.620
13	.685
14	.656

Note. TFI = Thought Fusion Instrument.

analyzed, multiple observations made the solution difficult to interpret. First of all, only two items loaded on the first factor and another two items loaded on the second factor. This can hardly be seen as sufficient to measure a meaningful construct. Also, multiple items loaded on more than one factor. Finally, inspection of the three factors gives no theoretically reason for meaningful factors since the factors did not, for example, correspond with the three classes of fusion beliefs. Based on the scree plot inspection that shows only one inflection, we also performed a second PAF for a one-solution structure. This one-solution structure appeared to explain 39.23% of the variance. Table 2 shows the factor loadings for each item for this one-factor solution. Applying Stevens' (2002) guidelines for substantive importance of factor loadings, a loading of .4 can be considered as significant. Based on this criterion, two items can be considered as less meaningful for the total score.

Beliefs About Rituals Inventory

For the BARI, a PA was also performed first to explore the number of factors to extract for our PAF. A scree plot of the results from this PA can be found in Figure 3. Our PA indicates five factors from the raw data which are above the 95th percentile cutoff line. However, just as with the TFI, the slope of the scree plot for this five-factor solution dramatically shifts after the first factor. Based on suggestions by Cattell (1966) and the assumption that PA tends to indicate more factors than warranted, a five-factor solution can hardly be meaningful in a 12-item questionnaire and additional procedures should be used to trim trivial factors (Buja & Eyuboglu, 1992); in the subsequent analysis, we explore a one-factor structure. We conducted

Figure 3. Scree plot of the results of the parallel analysis for the BARI. BARI = Beliefs About Rituals Inventory.

a PAF on the 12 items of the BARI. The Kaiser–Meyer–Olkin measure verified the sampling adequacy for the analysis, KMO = .81. Also, Bartlett's test of sphericity indicated that the correlation between items was sufficiently large for PAF (χ^2 = 962.35, *df* = 66, *p* < .001). The one-solution structure appeared to explain 42.64% of the variance. Table 3 shows the factor loadings for each item for this one-factor solution. Applying Stevens' (2002) guidelines for substantive importance of factor loadings, all items can be considered as meaningful.

Reliability

The Cronbach's α for the TFI was .86, indicating good internal consistency. The item-total score correlations ranged from .49 to .91. The mean item-total score correlation turned out to be .80. The Cronbach's α for the BARI turned out to be .89. The mean item-total score correlation was .67 with a range from .47 to .85. This result suggests good internal consistency. In addition,

Table 3. Summary of the exploratory factor analysis of the BARI for a one-factor solution

Item	Factor loadings
1	.792
2	.828
3	.694
4	.737
5	.449
6	.754
7	.833
8	.521
9	.487
10	.532
11	.570
12	.444

Note. BARI = Beliefs About Rituals Inventory.

McDonald's omega showed similar results for both scales. McDonald's ω for the TFI turned out to be .86, and for the BARI .90.

Construct Validity

Table 4 shows the correlations between the TFI, the BARI, and various other related scales to determine the construct validity. Pearson's correlations for the TFI ranged from .39 with the BDI-II to .53 with the MCQ-30. For the BARI, Pearson's correlation ranged from .28 with the BDI-II to .62 with the PI-R. The correlations with the convergent constructs turned out to be moderate to strong, and all appeared to be significant. The highest correlations were found between the TFI and the MCQ-30 and between the BARI and the PI-R, both convergent constructs. This suggests that the TFI and the BARI draw on strongly related constructs to the measurements chosen as convergent constructs. Correlations with the BDI-II turned out to be the lowest. For the BARI, the correlation with the BDI-II turned out to be .28. The correlation between the TFI and the BDI-II also turned out to be the lowest ($r = .39$).

Study 2

Test–Retest Reliability in a Normal Population

Participants and Procedure

A subsample consisting of the first 55 participants who completed the questionnaires for Study 1 was asked to complete the TFI and the BARI for a second time with an interval of 4–6 weeks to determine the test–retest reliability coefficients. Fifty-one participants completed this study (attrition rate of 7%). The mean age of these participants was 40.6 ($SD = 16.7$), 31% of the sample were male, 84% of the sample had a minimum education level of high school and were fully employed at the time of administration.

Table 4. Pearson's correlation coefficients between the TFI, BARI, and various other scales in a normal population ($n = 141$)

	TFI	BARI	PI-R	BDI-II	MCQ-30	IUS
TFI		.44**	.44**	.39**	.53**	.50**
BARI	.		.62**	.28**	.48**	.42**
PI-R				.58**	.75**	.70**
BDI-II					.68**	.67**
MCQ-30						.77**

Note. BARI = Beliefs About Rituals Inventory, BDI-II = Beck Depression Inventory, 2nd Version, IUS = Intolerance of Uncertainty Scale, MCQ-30 = Metacognition Questionnaire, PI-R = Padua Inventory, TFI = Thought Fusion Instrument. **$p < .001$.

Statistical Analysis

To investigate the test–retest reliability, mean scores at the two measurement points were calculated for our questionnaires. A t test was used to determine a possible significant difference between mean scores of the two measurement points. The intraclass correlation coefficient (ICC) was calculated to determine the test–retest reliability coefficients of the instruments. Based on the selection flowchart suggested by Koo and Li (2016), we chose an ICC estimate based on a single rating, absolute agreement, 2-way mixed-effects model. ICCs will be interpreted as follows: Values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively (Koo & Li, 2016).

Results

The mean scores at the first measurement point were 214.6 ($SD = 147.2$) for the TFI and 15.1 for the BARI ($SD = 4.4$). On the second occasion, this was 141.7 for the TFI ($SD = 145.9$) and 14.6 for the BARI ($SD = 4.02$). The mean scores at the two measurement points were significantly different for the TFI ($t[50] = 5.53$, $p < .001$). The mean scores at the two measurement points for the BARI were not different ($t[50] = 0.99$, $p = .324$). The ICC estimate based on a single rating, absolute agreement, 2-way mixed-effects model for the TFI was .84 with a 95% CI [.47–.93] ($p < .001$), suggesting good test–retest reliability. For the BARI, this was .63 with a 95% CI [.43–.77] ($p < .001$), suggesting moderately adequate test–retest reliability.

Study 3

Reliability and Validity in Clinical Samples

Participants and Procedure

To assess the reliability and validity of the TFI and the BARI in an OCD sample, data were collected in four groups, three clinical groups, and a nonclinical control group. First, measurements were administered to 61 patients with a primary diagnosis of OCD. They were diagnosed with OCD based on the criteria of the fourth version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) by means of the structured clinical interview for DSM-IV axis-I disorders (SCID-I; First et al., 2001), which was administered by a trained psychologist. At the time of testing, they were on the waiting list for psychological treatment at the therapeutic center PsyQ in Rotterdam. Also, they did not receive any other form of psychological treatment for at least 3 months, and there

was no change in medication dose or type in the 6 weeks before testing.

To determine the criterion-related validity of the TFI and the BARI, data were also collected in two other patient groups. Thirty patients with an anxiety disorder, also on the waiting list at the therapeutic center PsyQ and not receiving any form of treatment at the time of testing, volunteered for this purpose. These group consisted of patients who had a primary diagnosis of panic disorder ($n = 7$), social phobia ($n = 13$), a generalized anxiety disorder ($n = 5$), or specific phobia ($n = 5$). Diagnoses were established using the SCID-I interview. Patients with a comorbid axis-I disorder next to the abovementioned primary diagnoses were excluded from participation in this study. A third patient group consisted of 50 patients with an ASD, formerly classified as DSM-IV-TR autistic disorder ($n = 24$), Asperger syndrome ($n = 15$), and pervasive developmental disorder – not otherwise specified ($n = 7$). The diagnoses in this sample were based on the Dutch Interview for Diagnosing Autism Spectrum Disorder (Vuijk, 2016), a structured clinical interview including an evaluation of autism-specific behaviors by direct observation and a developmental history report provided by a parent or other caregiver. All diagnoses were verified by a registered psychologist or psychiatrist. Patients in the ASD group were selected based on the criteria of not having a comorbid diagnosis on axis-I, but this was not verified with a structured clinical interview before testing. Finally, a control group ($n = 63$) was collected in the general population, again using snowball sampling among relatives of the authors. In all groups, the TFI and the BARI and also the PI-R and the BDI-II were administered. Participant characteristics and outcome variables of the four groups are displayed in Table 5.

Statistical Analysis

To evaluate the reliability of the scales in clinical samples, we again used both Cronbach's α and McDonald's ω

coefficients. To explore the construct validity of the TFI and the BARI, Pearson's correlation coefficients were calculated between the TFI, the BARI, and a theoretically related construct (obsessive-compulsive symptoms [PI-R]) for the convergent validity and also with a divergent construct (depressive symptoms [BDI-II]). We hypothesize that the correlation coefficient between the TFI, the BARI, and the convergent construct will be substantially larger than with the discriminant construct (depressive symptoms).

To explore the criterion validity of the tests, we compared the TFI and BARI scores among participants with OCD, anxiety disorders, ASD, and a normal control condition. Differences in test scores among the groups were first examined by calculating the ANOVA with the TFI and the BARI as the dependent variables. Also, Bonferroni corrected post hoc tests were performed to make separate comparisons between each of the groups. Significantly higher ratings on the TFI and the BARI were expected in the OCD group in comparison with the other groups. Second, the accuracy of the instruments will be explored with area under the curve (AUC) analysis of receiver operating characteristics (ROC) curves. In a first analysis, data of all four groups will be used. We assessed whether the instruments are capable of distinguishing those with OCD from those without OCD based on the total score on the TFI and the BARI. Based on Swets (1986), ROC analysis results were interpreted as follows: AUC <0.70, low accuracy; AUC in the range of 0.70–0.90, medium accuracy; and AUC ≥0.90, high accuracy.

Results

Reliability

For the total group of patients ($n = 141$), Cronbach's α of the TFI turned out to be .89. The mean item-total score correlation was .76 (ranging from .59 to .78). More interestingly, separate analyses were carried out for the three patient groups. Cronbach's α for the anxiety disorder

Table 5. Descriptives and differences in metacognitive beliefs among the samples

	OCD	ASD	Anxiety disorder	Controls
	$n = 61$	$n = 50$	$n = 30$	$n = 63$
Variable	M (SD)	M (SD)	M (SD)	M (SD)
Age	30.87 (10.7)	41.73 (12.7)	27.63 (7.3)	45.52 (10.8)
Gender (% female)	64	34	80	69
TFI	368.56 (259.0)	386.40 (267.4)	386.33 (280.1)	272.54 (201.9)
BARI	26.25 (7.9)	23.18 (11.0)	18.97 (8.4)	40.61 (27.3)
PI-R	60.69 (24.9)	52.55 (32.9)	14.56 (3.86)	25.78 (17.6)
BDI-II	21.21 (11.0)	23.86 (14.7)	22.37 (13.6)	10.78 (9.3)

Note. BARI = Beliefs About Rituals Inventory, BDI-II = Beck Depression Inventory, 2nd Version, OCD = obsessive-compulsive disorder, PI-R = Padua Inventory-Revised, TFI = Thought Fusion Instrument.

group was .94 for the TFI, .90 for the group of patients with ASD, and .88 for the group of OCD patients. Assessing reliability with McDonald's omega, similar coefficients were found. For the total group of patients, McDonald's ω turned out to be .89. For the separate patient groups, McDonald's ω showed coefficients ranging from .86 to .91. In sum, the TFI seems to be a reliable measurement in both a general psychiatric sample and an OCD sample.

The same appears true for the BARI. For the total group of patients, the Cronbach's α was .96. The mean item-total score correlation was .87 with a range from .55 to .89. The Cronbach's α's for the separate patient groups were .95 for the anxiety disorder group, .96 for the group of patients with ASD and, .89 for the OCD group. Additionally, McDonald's ω for the total group turned out to be .96, and for the separate patient groups, coefficients ranged from .90 to .96.

Construct Validity

Table 6 shows the correlations between the TFI, the BARI, and two other scales in our OCD sample (PI-R as a convergent construct and BDI-II as a divergent construct). We found a moderate and significant correlation between the score on the TFI and the PI-R ($r = .37$, $p = .001$). On the other hand, as hypothesized, the correlation between the TFI and the BDI-II appeared not to be significant and was smaller ($r = .25$, $p = .052$). The same pattern is found for the BARI. The Pearson's correlation between the BARI and the PI-R appeared substantially larger ($r = .44$, $p < .001$) than the correlation coefficient between the BARI and the BDI-II ($r = .19$, $p = .149$).

Criterion Validity

A one-way ANOVA showed a significant between-groups effect, $F(3, 200) = 2.71$, $p = .046$ for the TFI. However, Bonferroni corrected post hoc tests revealed no significant differences between the four groups. A Bonferroni corrected post hoc test between the healthy controls and the total group of patients appeared to be significant ($p < .041$).

The criterion validity of the BARI appeared moderate. Here, a one-way ANOVA also showed a significant between-groups effect, $F(3, 200) = 24.57$, $p < .001$, and post hoc tests with the Bonferroni correction revealed

significant differences between all three patient groups and the control group ($p < .001$). Also, the OCD group had higher BARI scores in comparison with the group with anxiety disorders ($p < .001$). The difference between the OCD group and the ASD group was not significant ($p = .266$).

In line with the ANOVA results, the AUC analysis of the ROC curves for the BARI appeared to be .78, showing medium accuracy of the BARI in identifying those with OCD from those without OCD in our sample. The AUC of the TFI turned out to be .54, which can be interpreted as low accuracy of the TFI in distinguishing those with OCD from those without OCD in our sample. In a second AUC analysis using the combined clinical group, the AUC turned out to be .71, suggesting a medium accuracy for the TFI in distinguishing those with psychopathology from those without.

Discussion

In this study, the psychometric properties of the TFI and the BARI, two short questionnaires for measuring metacognitive beliefs in both the general population and clinical populations, were assessed.

First of all, by using a general population, for both the TFI and the BARI, an explorative factor analysis using PA and subsequent PAF was carried out. Although the PAs at first revealed a three-factor solution for the TFI and a five-factor solution for the BARI, after inspection of the scree plots and pattern matrices, a one-factor solution seems to connect closer to the data for both questionnaires. Since explorative factor analysis only allows constructing a hypothesis regarding the factor structure, future research must strengthen this conclusion by performing confirmative approaches.

Second, we found excellent reliability in both clinical and nonclinical populations for both scales. Cronbach's α's and McDonald's ω correlations were high for both questionnaires in a sample of healthy controls and in an OCD sample. Next, we found meaningful correlations between the TFI and the BARI and other theoretically related constructs. There were statistically significant correlations between the two scales and questionnaires for theoretically related concepts of obsessive-compulsive symptoms, intolerance of uncertainty, and metacognitive parameters in psychopathology in general. Correlations were substantially large, suggesting that the scales measure related constructs, pointing toward convergent validity. The lowest correlations were found between the two scales and the measurement for depressive symptoms, which was chosen as a divergent construct. The correlation between the BARI and the BDI-II

Table 6. Pearson's correlation coefficients between the TFI, BARI, and various other scales in an OCD sample ($n = 61$)

	TFI	BARI	PI-R	BDI-II
TFI		.28*	.37**	.25
BARI		.	.44**	.19
PI-R				.69**

Note. BARI = Beliefs About Rituals Inventory, BDI-II = Beck Depression Inventory, 2nd Version, OCD = obsessive-compulsive disorder, PI-R = Padua Inventory, TFI = Thought Fusion Instrument. *$p < .05$, **$p < .001$.

appeared the lowest in both the healthy control group and the OCD group, which suggests that the scale measures a distinct construct from depression. For the TFI, the correlations with the BDI-II also appeared the lowest of all correlations in both the healthy control condition and the OCD group, pointing to divergent validity. Although the results regarding the construct validity are in line with our hypothesis (larger correlation coefficients with convergent constructs than with divergent ones), it must be mentioned that caution is warranted. First of all, although the convergent correlations are indeed larger than the divergent one in Study 1, the correlations with depressive symptoms are also moderate and significant. The same appears to be true in the OCD sample of Study 3. Here too, some of the convergent correlations are only slightly larger than the divergent one. Second, when analyzing the entire pattern of correlations, it must be acknowledged that some of the correlations between convergent and divergent constructs are also high, for example, between the MCQ and the BDI. Although we state that there is no theoretical reason why OCD-specific metacognitive beliefs are directly associated with depressive symptoms, and therefore the BDI-II is suitable as a divergent construct, the MCQ is developed as a measurement for assessing metacognition over a broad range of psychological disorders (e.g., depressive disorder). This might explain the fact that the entire pattern of correlations is relatively high and of questionable utility to draw firm conclusions about the convergent and divergent validity of the measurements. In sum, although the results are in line with our hypothesis, more research is needed to draw more definitive conclusions, for example, by adding more measurements that are meant to capture the same and similar constructs as the TFI and the BARI for the examination of the convergent validity (e.g., the TAF scale) and by using questionnaires distinguishing metacognitive beliefs from theoretically unrelated constructs for the examination of the divergent validity (e.g., measurements for distinct cognitive processes).

Third, we assessed the test–retest reliability over a period of 4–6 weeks. For the TFI, the test–retest reliability as determined with the ICC turned out to be good (ICC = .84), for the BARI, it was less stable over time (ICC = .63), but still satisfactory (Koo & Li, 2016). For the TFI, there was however a significant difference between the mean score on the first and second occasion. Apparently, although the intraclass correlation was high, the stability of the item scores over time was low in our study, raising questions about whether the item scores on the TFI are affected by other variables (e.g., mood). Additional research is needed to test this hypothesis, and also warranted before clear recommendations for the use in clinical practice and research can be made.

Finally, the criterion validity of the TFI and the BARI was assessed by making a comparison of total scores on the TFI and the BARI between both a sample of healthy controls and various clinical groups. The BARI was able to discriminate between the healthy controls and the group of OCD patients, ASD patients, and patients with an anxiety disorder. Only the difference between the OCD and ASD groups was not significant. A possible explanation can be found in the fact that, although there is a clear difference between OCD and ASD, both are characterized by repetitive thoughts and the urge to perform repetitive actions (McDougle et al., 1995; Zandt et al., 2007). In summary, the criterion-related validity of the BARI seems to be moderate. Differences in total scores on the TFI appeared not to be significant between patient groups. Only the difference between the healthy control group and the total group of patients appeared significant. Also, the AUC of the TFI turned out to be .54, which can be interpreted as low accuracy of the TFI in distinguishing those with OCD from those without OCD in our sample, suggesting a low criterion validity in this study. A second AUC analysis using the combined clinical group turned out to be .71, suggesting a medium accuracy for the TFI in distinguishing those with psychopathology from those without.

Concerning the fusion beliefs of TAF, this is in line with earlier research which finds no differences between individuals with OCD versus other clinical samples (Abramowitz et al., 2003; Hansmeier et al., 2016; Rassin, Diepstraten et al., 2001; Solem et al., 2010). However, as we assumed that the TFI assesses fusion beliefs across the three fusion domains that are considered relevant and specific for OCD, this finding is contradictory to our hypothesis. Although more research is needed at this point, our study indicates that the TFI did not differentiate individuals with OCD from individuals with other disorders, which implies restrictions for the use in clinical practice. The TFI might be a useful instrument for setting up an individual case conceptualization or monitoring the treatment process, but not as a diagnostic instrument. Also, more research is needed before clear recommendations can be made about the use of both instruments as primary outcome measurement in treatment effect studies.

Our study is not without limitations. We chose an explorative factor analysis, rather than confirmative factor analysis, since the factor structure of both the TFI and the BARI had not been investigated yet. Although explorative factor analysis is the most suitable method for a first investigation of the factor structure of the measurements at this point and corresponds to the purpose of the current study, explorative approaches are not without limitations. The analyses in our study are only a first investigation into the factor structure of the measurements and do not provide a test of the unidimensionality of the scales. Since an explorative approach only allows generating hypothesis

about the underlying structure, confirmative approaches are now needed to draw more definitive conclusions about the factorial validity of the TFI and the BARI. Also, for the TFI, it can be hypothesized that there is a reason for a multidimensional scale since the TFI is considered to assess fusion beliefs across three theoretically meaningful classes of fusion beliefs. Further research using confirmatory approaches should test whether a three-dimensional model would fit the data better. Also, since the PA also revealed solutions with multiple factors for the BARI, confirmative approaches should take a closer look into this. Finally, future studies in larger populations could test whether the scales behave similarly in clinical and nonclinical samples.

Related to the limitations of our factor analysis is the fact that we have used the total scores of the measurements in the subsequent correlational analysis on convergent and divergent validity. Since our explorative factor analysis only allows making hypothesis about the underlying structure, this is a limitation of our subsequent analysis. Additional limitations are the lack of information about comorbidity in some of the clinical samples, while high rates of comorbidity might affect the outcome of our analysis, especially in the third study. Another limitation is that there were no data on mental health status of participants in the control conditions. Also, a limitation with the chosen sampling procedures is that they can generate a homogenous sample of individuals with specific demographic characteristics. For example, 89 participants of Study 1 had completed high school or a higher education level. Such specific demographic characteristics cause limitations to the generalization of the results. Other limitations are the use of only self-report measures of metacognitive beliefs and the higher number of females in the samples.

To summarize, both short questionnaires seem to have promising potential for research purposes and for use in clinical settings. More research into the psychometric properties is however necessary. Since some critical observations are made about the criterion validity and test–retest reliability of the TFI, conscientious use in clinical practice is required. Also, confirmative evidence regarding the factor structure of the measurements and additional research with regard to the concept of convergent and divergent validity is needed.

References

Abramowitz, J. S., Whiteside, S., Lynam, D., & Kalsy, S. (2003). Is thought-action fusion specific to obsessive-compulsive disorder? A mediating role of negative affect. *Behaviour Research and Therapy*, 41(9), 1069–1079. https://doi.org/10.1016/s0005-7967(02)00243-7

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509–540. https://doi.org/10.1207/s15327906mbr2704_2

Burns, G. L., Keortge, S. G., Formea, G. M., & Sternberger, L. G. (1996). Revision of the Padua Inventory of obsessive compulsive disorder symptoms: Distinctions between worry, obsessions, and compulsions. *Behaviour Research and Therapy*, 34(2), 163–173. https://doi.org/10.1016/0005-7967(95)00035-6

Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: The meta-cognitions questionnaire and its correlates. *Journal of Anxiety Disorders*, 11(3), 279–296. https://doi.org/10.1016/s0887-6185(97)00011-x

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

De Bruin, G. O., Rassin, E., van der Heiden, C., & Muris, P. (2006). Psychometric properties of a Dutch version of the intolerance of uncertainty scale. *Netherlands Journal of Psychology*, 62(2), 91–97. https://doi.org/10.1007/bf03061055

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412. https://doi.org/10.1111/bjop.12046

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2001). Structured clinical interview for DSM-IV-TR axis-I disorders (Research Version, Patient Edition; SCID-I/P). Biometrics Research Department, NY State Psychiatric Institute.

Fisher, P. L. (2009). Obsessive compulsive disorder: A comparison of CBT and the metacognitive approach. *International Journal of Cognitive Therapy*, 2(2), 107–122. https://doi.org/10.1521/ijct.2009.2.2.107

Fisher, P. L., & Wells, A. (2005). Experimental modification of beliefs in obsessive–compulsive disorder: A test of the metacognitive model. *Behaviour Research and Therapy*, 43(6), 821–829. https://doi.org/10.1016/j.brat.2004.09.002

Fisher, P. L., & Wells, A. (2008). Metacognitive therapy for obsessive–compulsive disorder: A case series. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(2), 117–132. https://doi.org/10.1016/j.jbtep.2006.12.001

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. https://doi.org/10.1037/1040-3590.10.3.206

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences*, 17(6), 791–802. https://doi.org/10.1016/0191-8869(94)90048-5

Frost, R. O., & Steketee, G. (1997). Perfectionism in obsessive-compulsive disorder patients. *Behaviour Research and Therapy*, 35(4), 291–296. https://doi.org/10.1016/s0005-7967(96)00108-8

Gaskin, C. J., & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use. *International Journal of Nursing Studies*, 51(3), 511–521. https://doi.org/10.1016/j.ijnurstu.2013.10.005

Gwilliam, P., Wells, A., & Cartwright-Hatton, S. (2004). Dose meta-cognition or responsibility predict obsessive-compulsive symptoms: A test of the metacognitive model. *Clinical Psychology Psychotherapy*, 11(2), 137–144. https://doi.org/10.1002/cpp.402

Hansmeier, J., Exner, C., Rief, W., & Glombiewski, J. A. (2016). A test of the metacognitive model of obsessive-compulsive disorder.

Journal of Obsessive-Compulsive and Related Disorders, *10*(3), 42–48. https://doi.org/10.1016/j.jocrd.2016.05.002

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/bf02289447

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Koo, T. K., & Li, M. Y. (2016). A Guideline of selecting and reporting Intraclass Correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

McDougle, C. J., Kresch, L. E., Goodman, W. K., Naylor, S. T., Volkmar, F. R., Cohen, D. J., & Price, L. H. (1995). A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. *American Journal of Psychiatry*, *152*(5), 772–777. https://doi.org/10.1176/ajp.152.5.772

McNicol, K., & Wells, A. (2012). Metacognition and obsessive-compulsive symptoms: The contribution of thought-fusion beliefs and beliefs about rituals. *International Journal of Cognitive Therapy*, *5*(3), 330–340. https://doi.org/10.1521/ijct.2012.5.3.330

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy*, *28*(6), 487–495. https://doi.org/10.1016/0005-7967(90)90135-6

Muris, P., Merckelbach, H., & Clavan, M. (1997). Abnormal and normal compulsions. *Behaviour Research and Therapy*, *35*(3), 249–252. https://doi.org/10.1016/s0005-7967(96)00114-3

Myers, S. G., Fisher, P. L., & Wells, A. (2009). An empirical test of the metacognitive model of obsessive-compulsive symptoms: Fusion beliefs, beliefs about rituals, and stop signals. *Journal of Anxiety Disorders*, *23*(4), 436–442. https://doi.org/10.1016/j.janxdis.2008.08.007

Myers, S. G., Grøtte, T., Haseth, S., Guzey, I. C., Hansen, B., Vogel, P. A., & Solem, S. (2017). The role of metacognitive beliefs about thoughts and rituals: A test of the metacognitive model of obsessive-compulsive disorder in a clinical sample. *Journal of Obsessive-Compulsive and Related Disorders*, *13*(2), 1–6. https://doi.org/10.1016/j.jocrd.2017.01.007

Myers, S. G., & Wells, A. (2005). Obsessive-compulsive symptoms: The contribution of metacognitions and responsibility. *Journal of Anxiety Disorders*, *19*(7), 806–817. https://doi.org/10.1016/j.janxdis.2004.09.004

Myers, S. G., & Wells, A. (2013). An experimental manipulation of metacognition: A test of the metacognitive model of obsessive-compulsive symptoms. *Behaviour Research and Therapy*, *51*(4–5), 177–184. https://doi.org/10.1016/j.brat.2013.01.007

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396–402. https://doi.org/10.3758/bf03200807

Rassin, E., Diepstraten, P., Merckelbach, H., & Muris, P. (2001). Thought-action fusion and Thought suppression in obsessive-compulsive disorder. *Behaviour Research and Therapy*, *39*(7), 757–764. https://doi.org/10.1016/s0005-7967(00)00051-6

Rees, C. S., & van Koesveld, K. E. (2008). An open trial of group metacognitive therapy for obsessive-compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(4), 451–458. https://doi.org/10.1016/j.jbtep.2007.11.004

Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders*, *10*(4), 379–391. https://doi.org/10.1016/0887-6185(96)00018-7

Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of reliability study: Review and new results. *Statistical Methods in Medical Research*, *13*(4), 251–271. https://doi.org/10.1191/0962280204sm365ra

Solem, S., Myers, S. G., Fisher, P. L., Vogel, P. A., & Wells, A. (2010). An empirical test of the metacognitive model of obsessive-compulsive symptoms: Replication and extension. *Journal of Anxiety Disorders*, *24*(1), 79–86. https://doi.org/10.1016/j.janxdis.2009.08.009

Sternberger, L. G., & Burns, G. L. (1990). Obsessions and compulsions: Psychometric properties of the Padua Inventory with an American college population. *Behaviour Research and Therapy*, *28*(4), 341–345. https://doi.org/10.1016/0005-7967(90)90087-y

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Erlbaum.

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin*, *99*(1), 100–117. https://doi.org/10.1037/0033-2909.99.1.100

Van der Does, A. J. W. (2002). *Handleiding bij de Nederlandse versie van Beck Depression Inventory – second edition (BDI–II–NL)*. Swets Test Publishers.

Van der Heiden, C., van Rossen, K., Dekker, A., Damstra, M., & Deen, M. (2016). Metacognitive therapy for obsessive-compulsive disorder: A pilot study. *Journal of Obsessive-Compulsive and Related Disorders*, *9*(2), 24–29. https://doi.org/10.1016/j.jocrd.2016.02.002

Van Oppen, P., Hoekstra, R. J., & Emmelkamp, P. M. G. (1995). The structure of obsessive-compulsive symptoms. *Behaviour Research and Therapy*, *33*(1), 15–23. https://doi.org/10.1016/0005-7967(94)e0010-g

Vuijk, R. (2016). *Nederlands Interview ten behoeven van Diagnostiek Autismespectrumstoornis bij volwassenen (NIDA)* [Dutch interview for diagnosing autism spectrum disorder]. Sarr Expertisecentrum Autisme/Dare to Design.

Wells, A. (1997). *Cognitive therapy of anxiety disorders: A practice manual and conceptual guide*. Wiley.

Wells, A. (2009). *Metacognitive therapy for anxiety and depression*. Guilford Press.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: Properties of the MCQ-30. *Behavior Research and Therapy*, *42*(4), 385–396. https://doi.org/10.1016/s0005-7967(03)00147-5

Wells, A., Gwilliam, P., & Cartwright-Hatton, S. (2001). *The Thought Fusion Instrument* [Unpublished self-report scale]. University of Manchester, UK.

Zandt, F., Prior, M., & Kyrios, M. (2007). Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder. *Journal of Autism and Developmental Disorders*, *37*(2), 251–259. https://doi.org/10.1007/s10803-006-0158-2

History

Kim Melchior
Outpatient Treatment Centre PsyQ
Max Euwelaan 70
3062 MA Rotterdam
The Netherlands
k.melchior@psyq.nl