Assessment
20(6) 764–775

DOI: 10.1177/1073191112436670
asm.sagepub.com
SAGE

Psychometric Evaluation of the Thought–Action Fusion Scale in a Large Clinical Sample

Joseph F. Meyer[1] and Timothy A. Brown[2]

Abstract

This study examined the psychometric properties of the 19-item Thought–Action Fusion (TAF) Scale, a measure of maladaptive cognitive intrusions, in a large clinical sample (N = 700). An exploratory factor analysis (n = 300) yielded two interpretable factors: TAF Moral (TAF-M) and TAF Likelihood (TAF-L). A confirmatory bifactor analysis was conducted on the second portion of the sample (n = 400) to account for possible sources of item covariance using a general TAF factor (subsuming TAF-M) alongside the TAF-L domain-specific factor. The bifactor model provided an acceptable fit to the sample data. Results indicated that global TAF was more strongly associated with a measure of obsessive-compulsiveness than measures of general worry and depression, and the TAF-L dimension was more strongly related to obsessive-compulsiveness than depression. Overall, results support the bifactor structure of the TAF in a clinical sample and its close relationship to its neighboring obsessive-compulsiveness construct.

Keywords

thought–action fusion, obsessive-compulsive disorder, obsessions, psychometrics, bifactor analysis

A central tenet of current cognitive-behavioral theories is that pathological anxiety arises from and is maintained by disorder-specific misinterpretations of physical sensations, situations, and thoughts (Clark, 1999). Although misinterpretations of physiological reactions are the core feature of panic disorder and distorted situational appraisals are a central difficulty in social phobia, ego-dystonic cognitive intrusions are thought to play a key role in the maintenance of obsessive-compulsive disorder (OCD; American Psychiatric Association, 2000). The quality of intrusive, negative thoughts in OCD (e.g., preoccupation with distressing cognitions and subsequent neutralization behaviors) implicates the role of thought–action fusion (TAF) as a relevant cognitive construct (Berle & Starcevic, 2005; Obsessive Compulsive Cognitions Working Group, 1997). The term *TAF* was coined by Rachman (1993) and was first discussed in the clinical literature as an observation that patients with OCD tend to view their thoughts as actions taking place in the external world, often assigning personal responsibility or special meaning to intrusive thoughts and engaging in self-blame for perceived harm inflicted on themselves or others (Salkovskis, 1985).

Contemporary studies linking TAF and OCD typically rely on a two-part assumption (Abramowitz, Whiteside, Lynam, & Kalsy, 2003): (a) subjective distress is generated when an individual conflates thinking about a harmful or immoral action with actually committing the act, and (b) compulsive behaviors emerge from an individual's attempts to neutralize perceived negative outcomes, which are viewed as more likely by simply having thought of them. Some contemporary researchers stress that the central pernicious element in TAF is the person's *belief* about both the meaning and the consequences of negative thoughts (Abramowitz et al., 2003). A potential precursor to this belief formation pinpointed by Salkovskis, Richards, and Forrester (1995) is a sense of personal responsibility or agency that emerges in tandem with intrusive thoughts. This sense of personal agency confers an inflated sense of authorship over event outcomes, which has been viewed as a reversal of the omission bias (e.g., doing nothing yields a stronger as opposed to a weaker sense of personal responsibility for an outcome). Heightened responsibility is then lessened via neutralization and other compulsive acts or rituals (Salkovskis et al., 1995).

[1]University of Rhode Island, Kingston, RI, USA
[2]Boston University, Boston, MA, USA

Corresponding Author:
Timothy A. Brown, Center for Anxiety and Related Disorders, Department of Psychology, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215, USA.
Email: tabrown@bu.edu

Thought–Action Fusion Scale

The first attempt at operationalization and psychometric exploration of TAF occurred with the development of the Responsibility Appraisal Questionnaire (Rachman, Thordarson, Shafran, & Woody, 1995). Using a sample of Canadian psychology undergraduates ($N = 291$), the investigators sought to develop a reliable measure of perceived responsibility devoid of items of overt OCD symptoms. One of the five content areas contained four items measuring TAF (e.g., "For me, having a mean thought is as bad as doing something mean"), and the principal components analysis (PCA)–derived TAF component (Cronbach's $\alpha = .51$) was moderately associated with obsessive symptoms as measured by the Inventory of Beliefs Related to Obsessionality ($r = .50$), the Maudsley Obsessional Compulsive Inventory (MOCI; $r = .45$), as well as state and trait guilt (rs = .34 and .37, respectively). The investigators thus implicated the role of TAF in OCD, observing that distressing cognitive intrusions can exacerbate guilt if there is a belief that thinking something negative is tantamount to acting on it (Rachman et al., 1995).

The advent of the Thought–Action Fusion Scale (TAFS; Shafran, Thordarson, & Rachman, 1996) followed closely after the Responsibility Appraisal Questionnaire study (Rachman et al., 1995). Elaboration of the TAF construct was based on clinical observations that intrusive thoughts may elicit a heightened sense of personal responsibility surrounding moral and probabilistic themes, and the TAFS was explicitly designed with these dimensions in mind (Shafran et al., 1996). Specifically, the moral thematic content, or TAF-M, refers to the confusion between thinking about and actually committing morally questionable acts, such as the belief that having thoughts of violence toward others is equally as reprehensible as physically committing them. The probabilistic content, or TAF-L, refers to the conflation of thought and likelihood of event occurrence, such as the belief that thinking about harm befalling a friend increases the probability that harm will occur in reality. This example refers to a specific type of TAF-L designated Likelihood-Other (TAF-LO) because the thought is about someone else. Probabilistic confusion about harm befalling oneself, for example, would be subsumed under Likelihood-Self (TAF-LS; Shafran et al., 1996).

Using an undergraduate student sample ($n = 190$), Shafran et al. (1996) conducted a PCA of the 19-item TAFS, which yielded three components (TAF-M, TAF-LS, and TAF-LO). However, this study also showed that a two-component solution (TAF-M and TAF-L) best characterized the data from a sample of community participants ($n = 147$) who endorsed OCD symptoms as defined by exceeding a cutoff score of 11 on the MOCI (see Hodgson & Rachman, 1977). The three-component structure was replicated in another student sample ($n = 272$) and an adult community sample ($n = 122$), whereas the two-component solution was replicated in another community sample endorsing OCD symptoms ($n = 118$; Shafran et al., 1996). Across two studies, participants with OCD symptoms (MOCI score ≥11) endorsed stronger TAF-L than participants without substantial OCD symptoms (MOCI score ≤11), with TAF-L features correlating positively with obsessive-compulsive symptoms (Shafran et al., 1996).

In addition to the original psychometric investigation (Shafran et al., 1996), other studies have provided evidence in support of the reliability and validity of the TAFS to varying degrees. Specifically, (a) all three TAFS domains have demonstrated acceptable reliability as evidenced by Cronbach's αs ranging from .75 to .96 (e.g., Marino, Lunt, & Negy, 2008; Rassin, Merckelbach, Muris, & Schmidt, 2001; Rassin, Muris, Schmidt, & Merckelbach, 2000; Shafran et al., 1996; Yorulmaz, Karanci, Bastug, Kisa, & Goka, 2008), (b) total TAF scores have evidenced moderate associations with reported obsessions as measured by a revised Padua Inventory (PI; Sanavio, 1988) and the MOCI (Rassin, Merckelbach, et al., 2001), and (c) patients with anxiety disorders endorse significantly stronger TAF features than healthy controls (Rassin, Merckelbach, et al., 2001). Correlations between the TAFS total and subscale scores of the MOCI range from weak to modest (rs = .20 to .38) across mostly undergraduate samples (Berle & Starcevic, 2005), suggesting questionable convergent validity. Although TAF-L scores were found to significantly (albeit weakly) correlate with MOCI scores ($r = .23$) in one study (Rassin, Merckelbach, et al., 2001), the TAFS total scores were not significantly higher for patients with OCD compared with patients with mixed anxiety diagnoses (e.g., panic disorder, social phobia, and posttraumatic stress disorder), indicating that TAF may not be specific to OCD.

TAF-L and TAF-M Relations to OCD

Across several studies, the TAF-L subscale appears to be more strongly related to OCD features relative to TAF-M. For example, in one study, TAF-L was more strongly associated with obsessions (as measured by MOCI and PI scores) compared with the TAF-M (Rassin, Merckelbach, et al., 2001). In another study, nonclinical participants endorsing OCD symptoms ($N = 424$ undergraduates) also demonstrated higher TAF-L compared with controls and an exaggerated belief that thoughts of undoing harm would reduce harm likelihood in reality (Amir, Freshman, Ramsey, Neary, & Brigidi, 2001).

With reference to more specific OCD symptoms, TAF-LO was found to significantly correlate ($r = .30$) with compulsive checking behaviors while controlling for depressive symptoms across both student and community samples endorsing OCD symptoms (Shafran et al., 1996). In another study (Cougle, Lee, Horowitz, Wolitzky-Taylor,

& Telch, 2008), the Washing factor of the Mental Pollution Questionnaire, which gauges adherence to hand washing in response to obscene thoughts, guilt, and committing perceived immoral actions, correlated significantly with TAF-LS (r = .20) and TAF-LO (r = .33) but not TAF-M (r = .11), even after controlling for depression and trait anxiety.

Despite some consistency across these findings (e.g., the stronger relation of TAF-L to OCD symptoms relative to TAF-M), few studies to date have drawn from adult clinical samples, which limits TAF-related inferences in the context of the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (*DSM-IV*; American Psychiatric Association, 1994) anxiety disorders. As evident from a comprehensive review of TAF research by Shafran and Rachman (2004), formally diagnosed patients were represented in only 5 of 20 peer-reviewed adult studies, none of which consisted of formal psychometric investigations of the TAFS. In contrast, 14 of the 20 adult studies consisted exclusively of undergraduate students. Since this 2004 review, the majority of studies continue to be conducted on undergraduate samples, and no further psychometric research has attempted to replicate the original Shafran et al. (1996) principal component structure in an exploratory factor analysis (EFA)/confirmatory factor analysis (CFA) framework in a clinical sample. Of note, the largest clinical samples used to date consisted of 95 patients (e.g., Abramowitz et al., 2003; Yorulmaz et al., 2008).

Present Study

A number of limitations are evident in the TAF literature, including (a) the utilization of PCA instead of common factor analysis, (b) the lack of latent structure replications in larger clinical samples, and (c) the use of the TAF total score in the absence of adequate psychometric support. With regard to Point (a), the Shafran et al. (1996) study and others (e.g., Pourfaraj, Mohammadi, & Taghavi, 2008; Yorulmaz, Yimaz, & Gençöz, 2004) used PCA instead of relying on factor analysis in their initial efforts to evaluate the TAFS. PCA is best characterized as a data reduction technique that does not distinguish common from unique variance, which means that components retain random error that would otherwise be removed in common FA. This has the effect of attenuating component intercorrelations, thus resulting in misleading statistical inferences (e.g., false positive conclusions about orthogonality of dimensions). PCA can yield such misleading results when the aim is to reproduce indicator intercorrelations with a smaller range of factors and when hypothesized factors subsume small numbers of indicators (Brown, 2006, p. 22). Because both these conditions apply to the TAFS, the use of common factor analysis was deemed more appropriate.

Concerning Point (b) above, previous studies have relied most heavily on undergraduate participants, who may or may not have met formal *DSM-IV* criteria for anxiety and/or mood disorders. In addition, although the original Shafran et al. (1996) study included community-dwelling adults, these participants were not formally diagnosed using validated semistructured diagnostic interviews; rather, their inclusion in an "obsessional" sample was solely determined on the basis of MOCI cutoff scores. Finally, with reference to Point (c), although multifactorial structures were reported, the TAF total score was still used in some studies (e.g., see Berle & Starcevic, 2005; Rassin, Merckelbach, et al., 2001) in the absence of an evidence-based rationale.

The purpose of the current study was to redress these shortcomings by examining the psychometric properties of the 19-item TAFS in a large clinical sample (N = 700) using an exploratory and confirmatory factor analytic framework. Given that no compelling reasons were encountered in the literature for expecting TAFS measurement properties to differ as a function of anxiety disorder diagnosis (e.g., size of item–factor relationships varying as a function of disorder type; cf. Brown, White, Forsyth, & Barlow, 2004), a clinically heterogeneous sample of patients was used in accord with the philosophy of the continuous nature of most psychopathological features (cf. Brown & Barlow, 2002, 2005, 2009). Moreover, in light of the broad bandwidth of symptom expression inevitably accompanying a large, heterogeneous clinical sample, we reasoned that a commensurate heterogeneity of TAF expression would arise in the form of a tripartite factor structure (i.e., TAF-M, TAF-LS, and TAF-LO in lieu of a domain-general TAF-L factor; see Abramowitz et al., 2003). Furthermore, if a tripartite structure were to hold, it would permit more fine-grained assessment of possible differential relationships among the domain-specific TAF-L factors and OCD, worry, and depressive symptoms. Therefore, it was hypothesized that (a) a three-factor solution would provide the best fit to the data and (b) the TAF dimensions would be more strongly associated with theoretically relevant constructs (i.e., obsessive-compulsive symptoms) compared with less theoretically relevant constructs (i.e., general depression and worry symptoms).

Method

Participants

Participants were 700 treatment-seeking adult outpatients who were assessed for anxiety and mood disorders at the Center for Anxiety and Related Disorders (CARD) at Boston University. Females (n = 426) constituted the majority of the sample (61%), and the average age was 32.68 (SD = 12.16). Most (90%) of the sample identified as Caucasian (n = 632), 5.1% identified as Asian (n = 36), 3.7% identified as African American (n = 26), and 0.8% identified as Other (n = 6). Per the selection criteria employed at CARD, patients were excluded from participation if they evidenced current suicidal or homicidal intent and/or plan, psychotic symptoms, or marked neurocognitive impairment (i.e., mental retardation, traumatic brain injury, or dementia).

Clinical diagnoses at intake were established using the Anxiety Disorders Interview Schedule for *DSM-IV*–Lifetime Version (ADIS-IV-L; Di Nardo, Brown, & Barlow, 1994), which was administered by trained PhD-level psychologists and advanced clinical doctoral students. The ADIS-IV-L is a semistructured diagnostic interview used to assess the presence and severity of *DSM-IV* anxiety, mood, substance use, and somatoform disorders and to screen for the presence of other major psychopathology (e.g., psychotic disorders). The ADIS-IV-L allows for dimensional assessment (i.e., using 0-8 ratings) of primary and associated disorder features across all diagnostic sections, and in most sections, symptoms are dimensionally rated regardless of whether a formal diagnosis is under consideration. Immediately after the interview, interviewers also dimensional rate (0-8) each *DSM-IV* diagnostic criterion for each major anxiety and mood disorder covered in the assessment. The ADIS-IV-L evidences good-to-excellent interrater reliability across current anxiety disorders (range of κs =.67-.86) and associated dimensional ratings (Brown, Di Nardo, Lehman, & Campbell, 2001). Each diagnosis is assigned a clinical severity rating (CSR), which indicates the level of impairment and/or distress associated with the particular disorder (0 = *none* to 8 = *very severely disturbing/disabling*). When patients are assigned two or more current diagnoses, the one with the highest CSR is referred to as the *principal* diagnosis, and nonprincipal diagnoses are labeled *additional* diagnoses. CSRs of 4 (*definitely disturbing/disabling*) or higher are assigned to disorders that meet or surpass the formal *DSM-IV* diagnostic threshold. In this sample, the *DSM-IV* principal diagnoses assigned most frequently to patients were social phobia (n = 152), generalized anxiety disorder (n = 111), panic disorder with agoraphobia (n = 98), OCD (n = 54), major depressive disorder (n = 52), and specific phobia (n = 44). Collapsing across principal and additional diagnoses, the rates of these *DSM-IV* disorders in the sample were as follows: social phobia (n = 327), generalized anxiety disorder (n = 220), panic disorder with or without agoraphobia (n = 173), OCD (n = 110), major depressive disorder (n = 195), and specific phobia (n = 102).

Measures

Thought–Action Fusion Scale. The TAFS (Shafran et al., 1996) consists of 19 items rated on a 5-point scale (0 = *strongly disagree* to 4 = *strongly agree*) from which a total score is derived (range = 0-76). These items are subdivided into TAF-M (12 items) and TAF-L (7 items). TAF-M refers to the belief that merely thinking about a moral transgression is tantamount to actually acting on it (e.g., "If I wish harm on someone, it is almost as bad as doing harm"), whereas TAF-L describes the belief that merely thinking about a particular event increases the probability of its occurrence. TAF-L is further subdivided into TAF-LS (3 items) and TAF-LO (4 items), which refer to thought–probability conflation regarding events occurring to individuals (e.g., "If I think of myself being injured in a fall, this increases the risk that I will have a fall and be injured") and significant others (e.g., "If I think of a relative/friend being in a car accident, this increases the risk that he/she will have a car accident"), respectively. Higher TAF scores reflect a stronger tendency toward TAF-like cognitions, and mean TAF total scores were stable across the original samples (Shafran et al., 1996).

Obsessive-Compulsive Inventory–Revised (OCI-R). The OCI-R (Foa et al., 2002) is an 18-item scale that has been used to effectively distinguish people with and without pathological levels of OCD. Each item is rated using a 4-point Likert-type scale where 0 = *not at all*, 2 = *moderately*, and 4 = *extremely*. Measured symptoms fall into six subscales: (a) obsessing (e.g., "I feel I have to repeat certain numbers"), (b) checking (e.g., "I check things more often than necessary"), (c) neutralizing (e.g., "I frequently get nasty thoughts and have difficulty getting rid of them"), (d) hoarding (e.g., "I collect things I don't need"), (e) ordering (e.g., "I get upset if objects are not arranged properly"), and (f) washing (e.g., "I wash my hands more often and longer than necessary"). Total OCI-R scores were moderately to highly correlated with subscale scores (rs = .63-.80) and evidenced good internal consistency (Cronbach's α = .90; Foa et al., 2002).

Beck Depression Inventory–II (BDI-II). Severity of depression symptoms was measured using the BDI-II (Beck, Steer, & Brown, 1996), a multiple-choice, 21-item self-report instrument that quantifies the severity of depressed mood symptoms and anhedonia. Response options range from 0 (*absence of symptoms*) to 3 (*intense symptoms*). BDI-II total score ranges are characterized as follows: (a) scores in the 1 to 10 range indicate the absence or negligible levels of depressed mood; (b) scores in the 11 to 16 range indicate mild mood disturbance; (c) scores in the 17 to 20 range reveal mild-to-moderate depression; (d) scores in the 21 to 30 range indicate moderate depression; (e) scores in the 31 to 40 range indicate severe depression; and (f) scores exceeding 40 indicate extreme depression. The BDI-II has garnered evidence for validity and reliability (e.g., α = .92 in an outpatient sample; Beck et al., 1996; Beck & Steer, 1984).

Penn State Worry Questionnaire (PSWQ). The PSWQ (Meyer, Miller, Metzger, & Borkovec, 1990; Molina & Borkovec, 1994) is a 16-item questionnaire using a 5-point scale (i.e., 1 = *not at all typical of me* and 5 = *very typical of me*) to measure degree of worry intensity and uncontrollability. Five items are reverse-scored to reduce response acquiescence, and total summed scores (range = 16-80) reflect overall degree of worry (e.g., higher scores indicate strong levels of worry). The PSWQ evidences high internal consistency (Cronbach's α = .93) for outpatients with anxiety disorders (Brown, Antony, & Barlow, 1992).

Procedure

Participants completed the TAFS as part of a battery of questionnaires (including the OCI-R, BDI-II, and PSWQ) administered prior to the ADIS-IV-L interview at CARD. The sample (N = 700) was randomly divided into two subsamples (Sample 1: N = 300, 123 males, 177 females; Sample 2: N = 400, 151 males, 249 females) for the purpose of examining and cross-validating the TAFS factor structure. The first sample was used to conduct an EFA on the TAFS given the absence of prior common factor analyses of this measure in clinical samples. A CFA was conducted in Sample 2 to corroborate the TAFS latent structure obtained in Sample 1 and to examine the concurrent validity of the TAFS dimensions (i.e., convergent and discriminant validity).

Data Analysis

A latent variable software program using maximum likelihood fitting functions (Mplus 5.0; Muthén & Muthén, 1998-2009) was used to analyze the raw data. Goodness of fit for the CFA models was evaluated using the following indices: (a) the root mean square error of approximation (RMSEA; Steiger, 1990) with accompanying 90% confidence interval (90% CI; MacCallum, Browne, & Sugawara, 1996), (b) the comparative fit index (CFI; Bentler, 1990), (c) the Tucker–Lewis index (TLI; Tucker & Lewis, 1973), and (d) the standardized root mean square residual (SRMR) statistic (Jöreskog & Sörbom, 1986). Acceptable model fit was defined as follows: RMSEA (close to or <.08, upper 90% CI close to or <.08), CFI (close to or <.95), TLI (close to or <.95), and SRMR (close to or <.08; Hu & Bentler, 1999). Multiple goodness-of-fit indices were used because they provide different information for the evaluation of model fit (e.g., absolute fit, fit adjusting for model parsimony, fit relative to a null model); when considered together, these indices provide a conservative and reliable means of model fit evaluation (cf. Brown, 2006).

To guide factor selection, a scree test (Cattell, 1966) and a parallel analysis (Horn, 1965) were conducted using an EFA program (Factor 8.02; Lorenzo-Seva & Ferrando, 2011). Although the scree test requires the researcher to determine the optimal number of factors based on the elbow of a plot of eigenvalues, parallel analysis is a procedure that compares eigenvalues derived from sample data with eigenvalues generated from random data to ascertain whether factors from the sample data account for more variance than expected by chance. The scree test is generally accurate, but its performance is variable (Zwick & Velicer, 1986) and may result in under- or overfactoring (Fava & Velicer, 1992). Parallel analysis is a highly recommended method of factor extraction (Worthington & Whittaker, 2006) and has generated accurate results across numerous studies, especially those using large samples (see Zwick & Velicer, 1986).

Results

An EFA was conducted on Sample 1 using maximum likelihood estimation and oblique rotation (geomin). Factor selection and acceptability were guided by the scree test, parallel analysis, goodness of model fit, solution interpretability, and strength of parameter estimates (i.e., primary factor loadings >.30). Scree test results suggested either a two- or three-factor solution. However, a parallel analysis using 500 random correlation matrices from a normal distribution indicated a two-factor solution. Eigenvalues for the unreduced correlation matrix were 9.2, 2.9, 1.1, 0.84, and 0.81 (variance explained = 49%, 15%, 6%, 4%, and 4%, respectively). In view of these results, both two-and three-factor models were pursued, a decision that was further based on the fact that two- and three-component solutions had been reported in previous psychometric studies of the TAFS.

The two- and three-factor solutions yielded factors comprising (a) TAF-M and TAF-L and (b) TAF-M, TAF-LO, and TAF-LS, respectively. Primary factor loadings for all 19 items were well more than .30 for both the two-factor (range = .54-.97) and three-factor (range = .59-.96) EFA solutions. Factor correlations ranged from weak to moderate across the two-factor (r = .51) and three-factor (rs = .28-.49) solutions. Although the three-factor model provided a better fit to the data than the two-factor model (e.g., TLIs = .92 and .83, respectively), all three TAF-LS items (i.e., 12, 14, and 16) in the three-factor model evidenced salient (i.e., >.30) cross-loadings on LO (range = .39-.51). However, the two-factor model did not contain any salient cross-loadings (see Table 1 for details). Based on these results, the two-factor solution was judged the most acceptable model based on best simple structure and parsimony.

A CFA was conducted on Sample 2 to further evaluate the two-factor model. Measurement error was specified as random. The confirmatory model did not fit the data well, $\chi^2(151) = 947.6$; $p < .001$; RMSEA = .12; 90% CI = .11, .12; SRMR = .06; CFI = .88; TLI = .86; Bayesian information criterion (BIC) = 17842.8. Fit diagnostics indicated that localized strains in the solution corresponded to the error covariances of Items 12 and 14, Items 12 and 16, and Items 14 and 16 (range of modification indices = 104.75-177.02).

Table 1. Factor Structure of the Thought–Action Fusion Scale: Exploratory Factor Analysis ($N = 300$).

TAFS Item	Factor	
	TAF-M	TAF-L
1. Thinking of making an extremely critical remark to a friend is almost as unacceptable to me as actually saying it.	**.579**	.003
3. Having a blasphemous thought is almost as sinful to me as a blasphemous action.	**.614**	.164
4. Thinking about swearing at someone else is almost as unacceptable to me as actually swearing.	**.733**	.011
6. When I have a nasty thought about someone else, it is almost as bad as carrying out a nasty action.	**.725**	.156
8. Having violent thoughts is almost as unacceptable to me as violent acts.	**.781**	−.044
10. When I think about making an obscene remark or gesture in church, it is almost as sinful as actually doing it.	**.751**	.035
11. If I wish harm on someone, it is almost as bad as doing harm.	**.766**	−.002
13. When I think unkindly about a friend, it is almost as disloyal as doing an unkind act.	**.775**	−.055
15. If I think about making an obscene gesture to someone else, it is almost as bad as doing it.	**.855**	−.025
17. If I have a jealous thought, it is almost the same as making a jealous remark.	**.709**	.042
18. Thinking of cheating in a personal relationship is almost as immoral to me as actually cheating.	**.647**	−.056
19. Having obscene thoughts in a church is unacceptable to me.	**.580**	.072
2. If I think of a relative/friend losing their job, this increases the risk that they will lose their job.	−.018	**.853**
5. If I think of a relative/friend being in a car accident, this increases the risk that he/she will have a car accident.	−.001	**.958**
7. If I think of a relative/friend being injured in a fall, this increases the risk that he/she will have a fall and be injured.	.017	**.970**
9. If I think of a relative/friend falling ill, this increases the risk that he/she will fall ill.	−.041	**.974**
12. If I think of myself being injured in a fall, this increases the risk that I will have a fall and be injured.	.102	**.653**
14. If I think of myself being in a car accident, this increases the risk that I will have a car accident.	.117	**.644**
16. If I think of myself falling ill, this increases the risk that I will fall ill.	.126	**.539**

	Factor Correlation	
	TAF-M	TAF-L
TAF-M	—	
TAF-L	.507	—

Note. TAFS = Thought–Action Fusion Scale; TAF-M = TAF-Moral; TAF-L = TAF-Likelihood. The exploratory factor analysis was conducted using maximum likelihood estimation and oblique (geomin) rotation. Factor loadings ≥.30 are in boldface.

Based on these results, the model could be revised in two ways: (a) the retention of a two-factor model with the specification of error covariances for Items 12, 14, and 16 or (b) the specification of a three-factor model splitting TAF-L into TAF-LS and TAF-LO.

The two-factor model with error covariances was pursued given (a) its parsimony with regard to number of factors, (b) the fact that a separate TAF-LS factor would be defined by only three items (and the TAF-LO factor would consist of four items), and (c) the lack of compelling reasons to distinguish TAF-LS and TAF-L as separate, substantively meaningful dimensions. With regard to (c), it was deemed likely that the three TAF-LS items measured the general TAF-L dimension but shared additional covariance because of a method effect (e.g., because of differential item content). Thus, the CFA solution was respecified with correlated measurement errors for Items 12, 14, and 16.

Although the revised CFA model provided a better fit to the data than the initial two-factor model: that is, $\chi^2(148) = 565.0$, $p < .001$, BIC = 17478.2, modification indices revealed two salient areas of ill fit that corresponded to the error covariances of Items 8 and 11 and Items 10 and 19 (modification indices = 52.44 and 57.35, respectively). Because this unexplained indicator covariance was likely because of the highly similar wording and content of these items (e.g., Items 8 and 11 pertain to harming others, and Items 10 and 19 assess obscene thoughts, remarks, and/or gestures in church), the CFA solution was respecified with correlated measurement errors for these two pairs of items.

The second revised CFA model with additional correlated residuals provided an acceptable fit to the data, $\chi^2(146) = 454.8$; $p < .001$; RMSEA = .07; 90% CI = .07, .08; SRMR = .05; CFI = .95; TLI = .94; BIC = 17380.0, and fit diagnostics did not detect other salient areas of strain. As hypothesized on the basis of overlapping item content, significant error covariances were observed among Items 12, 14, and 16 (r = .54-.67, ps < .001), between Items 8 and 11 ($r = .37$, $p < .001$), and between Items 10 and 19 ($r = .39$, $p < .001$). All factor loadings were exceeded .30 (range = .59-.95), and the factor correlation was moderate in strength ($r = .55$; see Table 2 for details).

Table 2. Factor Structure of the Thought–Action Fusion Scale: Confirmatory Factor Analysis With Correlated Residuals (N = 400).

TAFS Item	Factor	
	TAF-M	TAF-L
1. Thinking of making an extremely critical remark to a friend is almost as unacceptable to me as actually saying it.	.593	
3. Having a blasphemous thought is almost as sinful to me as a blasphemous action.	.767	
4. Thinking about swearing at someone else is almost as unacceptable to me as actually swearing.	.738	
6. When I have a nasty thought about someone else, it is almost as bad as carrying out a nasty action.	.808	
8. Having violent thoughts is almost as unacceptable to me as violent acts.	.751	
10. When I think about making an obscene remark or gesture in church, it is almost as sinful as actually doing it.	.801	
11. If I wish harm on someone, it is almost as bad as doing harm.	.743	
13. When I think unkindly about a friend, it is almost as disloyal as doing an unkind act.	.757	
15. If I think about making an obscene gesture to someone else, it is almost as bad as doing it.	.858	
17. If I have a jealous thought, it is almost the same as making a jealous remark.	.735	
18. Thinking of cheating in a personal relationship is almost as immoral to me as actually cheating.	.715	
19. Having obscene thoughts in a church is unacceptable to me.	.669	
2. If I think of a relative/friend losing their job, this increases the risk that they will lose their job.		.761
5. If I think of a relative/friend being in a car accident, this increases the risk that he/she will have a car accident.		.939
7. If I think of a relative/friend being injured in a fall, this increases the risk that he/she will have a fall and be injured.		.954
9. If I think of a relative/friend falling ill, this increases the risk that he/she will fall ill.		.952
12. If I think of myself being injured in a fall, this increases the risk that I will have a fall and be injured.		.722
14. If I think of myself being in a car accident, this increases the risk that I will have a car accident.		.734
16. If I think of myself falling ill, this increases the risk that I will fall ill.		.636

	Factor Correlation	
	TAF-M	TAF-L
TAF-M	—	
TAF-L	.551	—

Note. TAFS = Thought–Action Fusion Scale; TAF-M = TAF-Moral; TAF-L = TAF-Likelihood. The confirmatory factor analysis was conducted using maximum likelihood estimation. For all values, $p < .001$.

Based on these CFA results, which corroborated that the majority of items (i.e., 12 of 19) loaded onto the TAF-M factor, the final CFA solution was respecified as a bifactor model (cf. Chen, West, & Sousa, 2006; Holzinger & Swineford, 1937). One key advantage of the bifactor model is the ability to evaluate the importance of domain-specific factors. In this case, the TAF-M domain-specific factor may not be relevant to the prediction of the TAFS items when a general TAF factor is included in the model. This situation seemed possible because most items loaded onto a broad TAF-M factor (with some additional covariation explained by a smaller TAF-L dimension). Although second-order factor models allow for the inclusion of superordinate factors, these analyses do not permit indicators to load directly on the higher order factors nor do they directly permit the validation of subdomain factors. Bifactor models, on the other hand, allow a general factor to directly account for item covariation, along with possible orthogonal subdomain factors that account for additional covariance over and beyond the general factor. Stated another way, the bifactor technique is appropriate given its ability to provide a clear picture of both (a) the homogeneity or meaningful overlap of scale items as captured by the specification of a general factor and (b) the heterogeneity or diversity of items loading onto domain-specific factors. The unique question answered by this technique, which is not addressed by traditional factor analysis or hierarchical factor analysis methods, is the following: Does a global factor account for covariance among an entire set of scale indicators, with the possible addition of a subdomain factor (or factors) that explains additional indicator covariance not accounted for by the general factor? A bifactor confirmatory model thus appropriately tests the hypothesis that a TAF general factor will account for common item variance among the 19 TAFS items (and subsume the TAF-M factor), and the TAF-L domain-specific factor will account for unique variance not explained by the TAF general factor.

Thus, a confirmatory bifactor analysis was conducted on Sample 2 using a general TAF factor as well as the TAF-L domain-specific factor. The interfactor covariance was fixed

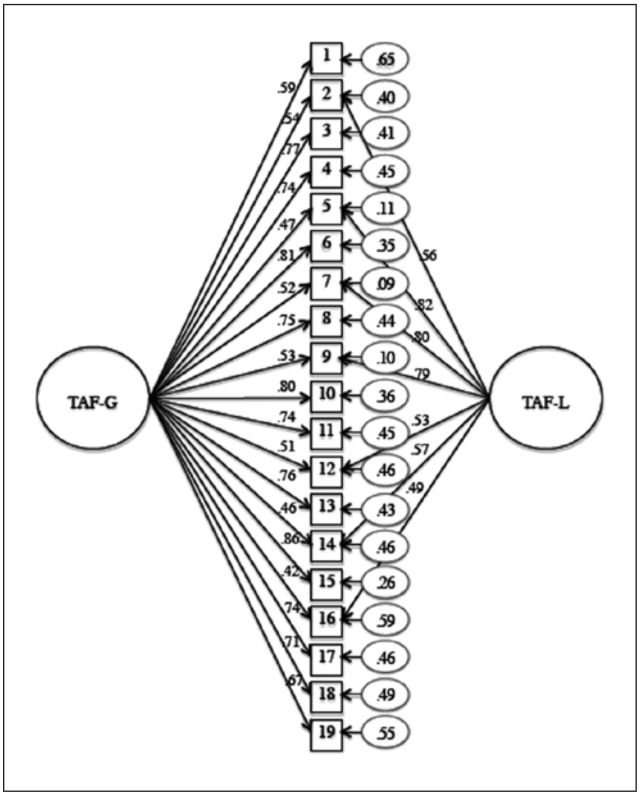

Figure 1. Bifactor measurement model of the TAFS.
*Note***.** TAFS = Thought–Action Fusion Scale; TAF-G = thought–action fusion general factor; TAF-L = thought–action fusion likelihood domain-specific factor. Completely standardized maximum likelihood parameter estimates. Squares denote TAFS Items 1 to 19. Circles denote TAFS factors. For purposes of clarity, the error covariances among five items have been omitted from the figure but are provided in the Results section. For all values, $p < .001$.

to zero, and the five previously noted item error covariances were retained. The bifactor model provided an acceptable fit to the data, $\chi^2(140) = 414.0$; $p < .001$; RMSEA = .07; 90% CI = .06, .08; SRMR = .04; CFI = .96; TLI = .95; BIC = 17375.2. No salient points of ill model fit were detected. The completely standardized parameter estimates from the bifactor solution are presented in Figure 1. The factor loadings for both the general TAF factor and the TAF-L domain specific factor exceeded .30 (range = .42-.86). Correlated residuals among Items 12, 14, and 16 were again significant and moderate in strength (r = .53-.67, ps < .001) as were correlated residuals between Items 8 and 11 (r = .38, $p < .001$) and between Items 10 and 19 (r = .39, $p < .001$).[1]

Scale Reliability

Given the limitations of Cronbach's alpha when certain measurement model conditions are violated (e.g., tau equivalence and absence of correlated residuals; cf. Raykov, 2001a, 2001b), scale reliabilities (ρ) of the two factors within the bifactor CFA model (n = 400) were computed using Raykov's (2004) CFA-based estimation method. This method revealed high scale reliability for general TAF (ρ = .97) and the TAF-L subdomain (ρ = .95).

Table 3. Convergent and Discriminant Validity of TAF Factors With Measures of OCD, Generalized Worry, and Depression.

	OCI-R	PSWQ	BDI-II
TAF	.384[a]	.153[b]	.187[b]
TAF-L	.240[a]	.204[a]	.085[b]

Note. TAF = Thought–Action Fusion general factor; OCD = obsessive-compulsive disorder; TAF-L = TAF-Likelihood; OCI-R = Obsessive-Compulsive Inventory–Revised; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory–II. Correlations across rows with different superscripts are significantly different in magnitude ($p < .001$) as indicated by Steiger's z tests.

Concurrent Validity of TAF Factors

To evaluate the convergent and discriminant validity of the TAF factors, correlations with anxiety and depression measures were estimated by including the single indicators of OCI-R, PSWQ, and BDI-II as covariates in the bifactor CFA model (n = 400). Although evidence regarding concurrent validity of TAF factors has been mixed (Berle & Starcevic, 2005; Coles, Mennin, & Heimberg, 2001; Rassin, Merckelbach, et al., 2001), no study has examined the validity of a general TAF factor. On the basis of Shafran et al.'s (1996) original theoretical rationale (viz., proposed link between OCD and intrusive thoughts with thought–action conflation), it was predicted that the general TAF factor and TAF-L domain would evidence stronger correlations with OCD features (OCI-R) than with the general worry (PSWQ) and depression (BDI-II) measures.

Results partially supported these predictions, and correlations are provided in Table 3. Convergent rs were moderate (rs = .24 and .38), and discriminant rs were weak in magnitude (r = .09-.20). Global TAF was significantly more strongly correlated with OCI-R scores than BDI-II and PSWQ scores as evidenced by Steiger's z tests of differential magnitude, $z = 3.04$, $p < .01$; and $z = 3.53$, $p < .001$, respectively. However, although TAF-L was significantly more strongly correlated with OCI-R scores than BDI-II scores, $z = 2.25$, $p < .05$, TAF-L was not significantly more strongly correlated with OCI-R scores than PSWQ scores.

Finally, correlations among all six OCI-R subscales and TAF factors (i.e., general TAF and TAF-L) were estimated. Single indicators of the OCI-R subdomains were included as covariates in the bifactor CFA model, and the results are provided in Table 4. Although a priori hypotheses about the nature of these differential relationships could not be formulated at present given certain limitations in the literature

Table 4. Correlations Among TAF Factors and OCI-R Subscales.

	Checking	Hoarding	Neutralizing	Obsessions	Ordering	Washing
TAF	.279***	.239***	.241***	.272***	.289***	.220***
TAF-L	.180**	.064	.209***	.205***	.172**	.106*

Note. TAF = Thought–Action Fusion general factor; TAF-L = TAF-Likelihood; OCI-R = Obsessive-Compulsive Inventory–Revised.
*p < .05. **p < .01, ***p < .001.

(e.g., the inconsistency of OCI-R subdomain relations to TAF factors and lack of clinical samples), these exploratory analyses were conducted to aid future research attempting predictions at the level of the OCI-R subscales.

Discussion

These findings represent an extension of the psychometric basis of the TAFS through a combinatory EFA and bifactor CFA framework applied to a large clinical outpatient sample. Contrary to our three-factor hypothesis (i.e., TAF-M, TAF-LO, and TAF-LS), a two-factor structure (i.e., TAF-M and TAF-L) was consistent between the current clinical samples, which aligns with results from clinical samples (e.g., Shafran et al., 1996). CFA results also supported Shafran et al.'s (1996) original two-factor model, yet further provided an empirical rationale for the hypothesis that all 19 items tap a single, broader TAF construct. This subsequently led to a bifactor CFA model to capture simultaneously the homogeneity (i.e., overlap) and heterogeneity (i.e., diversity) of the TAF-L subdomain. Bifactor CFA results indicated that a general TAF factor accounted for covariation among all indicators, whereas the TAF-L domain-specific factor (orthogonal to the general factor) explained additional item covariance *not* explained by the general TAF factor. Moreover, both the general TAF factor and TAF-L subdomain evidenced strong reliability (ρs = .97 and .95, respectively). These findings call into question the necessity of separately specifying the TAF-M factor in future studies using the TAFS in heterogeneous clinical samples and suggest that a considerable amount of TAFS item covariance can be more parsimoniously accounted for by a global TAF dimension.

Key reasons for pursuing the two-factor bifactor solution in the current study included (a) better parsimony, (b) the small numbers of items defining the TAF-LS and TAF-LO subdomains (3 and 4 items, respectively), and (c) lack of a compelling conceptual basis for distinguishing TAF-LS and TAF-LO (i.e., sparse evidence and paucity of consistent theoretical justification for the differential importance of separating these two subdomains from TAF-L). Moreover, given that the majority of TAFS items (i.e., 12 of 19) loaded onto the TAF-M factor (whereas the domain-general TAF-L factor accounted for most of the remaining indicator variance), the relevance of separating TAF-M from the global TAF factor (i.e., for prediction purposes) should be questioned in future research. Informative results may be obtained in clinical settings by relying on a total TAFS score, which accounts for variance in all TAFS items, although additional unique variance *not* explained by the global factor can be captured through the separate assessment of TAF-L as evidenced by our bifactor solution (as well as a handful of error covariances). However, these inferences should be heeded cautiously since further psychometric inquiry is needed using large clinical samples.

Also worthy of discussion is whether it is tenable to split the TAF-L factor into smaller, more specific factors (TAF-LO, TAF-LS). Although the three-factor EFA model (i.e., TAF-M, TAF-LO, and TAF-LS) fit the Sample 1 data better than the two-factor EFA model (i.e., TAF-M and TAF-L), all three TAF-LS items cross-loaded onto TAF-LO in the three-factor model. This finding calls into question the substantive distinctiveness of the TAF-LS and TAF-LO dimensions. That is, are TAF-LS and TAF-LO distinct and substantively important constructs, or is the differential covariance among TAF-LS and TAF-LO items better construed as method variance (i.e., artifactual covariance introduced by highly similar item content or direction of item wording; cf. Brown, 2003, Marsh, 1996)? For the reasons noted earlier, our subsequent TAFS measurement model specifications were pursued under the assumption that the additional covariance among TAF-LS items was because of a method effect. However, future research should address this issue further. For instance, the separation of TAF-L into smaller and specific dimensions (TAF-LS, TAF-LO) should be bolstered by sound conceptual reasoning, as well as evidence attesting to the discriminant validity of these subfactors and their distinct contribution to the prediction of clinical outcomes.

Regarding concurrent validity, the current results revealed that general TAF was more strongly related to obsessive-compulsive features than depression symptoms or general worry. General TAF had a significantly stronger correlation with the OCI-R compared with (a) its weaker correlation with BDI-II scores and (b) its weaker correlation with PSWQ scores. Relative to previous studies detecting more modest correlations between TAF and OCD (e.g., Rassin, Merckelbach, et al., 2001; Sanavio, 1988), the stronger TAF-OCD relationship found in this study furnishes more compelling evidence for convergent validity. Lower correlations in previous studies may arise from the absence of a global TAF factor. Also, past studies have relied on different OCD measures (e.g., the PI and MOCI) and have primarily investigated nonclinical samples.

Although the TAF-L subdomain evidenced high reliability, evidence for its convergent validity was mixed. Specifically, TAF-L was significantly more strongly related to OCI-R scores than BDI-II scores as expected, but the differential magnitude between TAF-L and OCI-R and between TAF-L and PSWQ was nonsignificant. The latter finding suggests TAF-L concerns are similarly related to obsessive-compulsiveness and general worry, in accord with the notion that anxiety associated with a perceived increased probability of harm inflicted on the self or close others may occur in the context of uncontrollable worry (i.e., in GAD). This type of worry is conceptually consistent with TAF-L content whereas depressive symptoms, such as self-blame and low self-worth, are more consonant with non–TAF-L content relevant to self-deprecation.

Taken together, these results provide encouraging evidence for a more parsimonious latent structure of the TAFS. Given that general TAF (a) evidenced high reliability and (b) was more strongly related to OCD features than symptoms of worry and depression endorsed by clinical participants, the use of a total TAF score may prove more valuable than domain-specific scores in future studies. These observations attest to the notable overlap among the different qualitative features of TAF, be they related to self or to others, and imply that the subdomains synergistically contribute to observed heterogeneity in TAF expression. To date, this study marks the first application of a bifactor CFA to the TAFS and the first empirical support for the TAF total score. Future replications using clinical and nonclinical samples are needed to cross-validate this more parsimonious factor structure.

With regard to clinical implications, evidence for the presence and superiority of a TAF total score (general factor) promotes a more parsimonious use of this measure in future treatment studies (e.g., using TAFS as a process or outcome measure). Optimal measurement of TAF is paramount, especially since this construct continues to be (a) investigated in applied clinical research settings and (b) used by practicing clinicians (and not strictly psychometrically oriented psychologists) in the assessment of maladaptive cognitive correlates of OCD that may require special attention and intervention strategies (e.g., cognitive challenging and restructuring) during the course of treatment. If future research is able to replicate the present findings, which represent an application of Occam's razor to potentially unnecessary construct specification complexity, cognitive-behavioral clinical researchers can look forward to conceptualizing the latent structure of TAF in a more economically compressed manner. This, in turn, will ultimately benefit practicing clinicians through the provision of cleaner, more efficient measurement of the scope and severity of these distressing and potentially impairing cognitive distortions.

To the extent that treatment-seeking patients endorse distressing and interfering TAF-like symptoms, cognitive-behavioral interventions may be used effectively. For example, a study of undergraduate students endorsing elevated TAF scores revealed a significant reduction in reported anxiety and TAF severity following an educational intervention relative to controls (Zucker, Craske, Barrios, & Holguin, 2002). General "anti-TAF" strategies, such as the one used in the Zucker et al. (2002) study, may suffice in place of more time-consuming strategies addressing domain-specific intrusions (e.g., as recommended by Freeston, Rhéaume, & Ladouceur, 1996). Furthermore, when accompanying a diagnosis of OCD, TAF symptoms may only require direct attention in treatment-resistant cases (Shafran & Rachman, 2004). It was observed, for example, that TAF symptoms significantly improved following successful treatment of OCD, suggesting that mainstream CBT may be sufficient depending on diagnostic profile (Rassin, Diepstraten, Merckelbach, & Muris, 2001). Nevertheless, additional research is needed to assess the efficacy of specific cognitive-behavioral interventions (e.g., psychoeducation and general behavioral exposures) for alleviating anxiety associated with TAF.

Despite the strengths of the current study, some limitations warrant attention. First, with regard to demographic characteristics, the majority (90%) of the study sample comprised Caucasian outpatients, which limits the generalizability of the results across different ethnic and racial groups. Although the three-factor solution has been obtained in Turkish samples (Yorulmaz et al., 2004, 2008), additional investigations of the TAFS should test this factor structure (as well as a bifactor conceptualization) across more diverse samples. Second, although 110 patients in our sample were diagnosed with OCD above the *DSM-IV* threshold, the nature of TAF expression in larger, more focused OCD clinical samples (e.g., from specialized OCD clinics and research centers) deserves further clarification in future psychometric studies given the consistent OCD–TAF relationship (Berle & Starcevic, 2005). However, researchers should bear in mind that although TAF shares a robust, modest-to-moderate relationship with OCD symptoms (Rassin, Merckelbach, et al., 2001), TAF is not *exclusively* associated with OCD symptoms meeting *DSM-IV* diagnostic criteria (Rassin, Diepstraten, et al., 2001). Rather, TAF-like cognitive intrusions have been detected in clinical depression as well as a broad range of anxiety conditions including pathological worry, social anxiety, and panic (Berle & Starcevic, 2005; Hazlett-Stevens, Zucker, & Craske, 2002).

In closing, further research is needed to delineate the interrelationships among general TAF, TAF subdomains, general worry, depression, and intrusive thoughts, which may share varying degrees of overlap across OCD and GAD (e.g., Lee, Cougle, & Telch, 2005). Differential expression of TAF features across different anxiety disorders (i.e., degrees of specificity and commonality of TAF) also requires more in-depth consideration (Hazlett-Stevens et al., 2002). Hopefully, research into distinct cognitive processes underlying disorder constructs will aid in enriching treatment strategies targeting maladaptive thought content and subsequent behavioral repercussions.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by Grant MH039096 from the National Institute of Mental Health.

Note

1. A bifactor CFA including general TAF, TAF-M, and TAF-L was also conducted to determine whether meaningful variance was attributable to TAF-M after accounting for the general factor. Although this solution converged, results indicated (a) an out of range estimate for the residual variance of Item 3, which resulted in a nonpositive definite theta matrix and (b) statistical nonsignificance for all TAF-M factor loadings, with all but one loading <.30. Thus, the domain-specific TAF-M factor was no longer relevant to the prediction of the TAFS items when the general factor was included in the model (as evidenced by the preponderance of weak and statistically nonsignificant factor loadings).

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*, 1069-1079.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (text revision). Washington, DC: Author.

Amir, N., Freshman, M., Ramsey, B., Neary, E., & Brigidi, B. (2001). Thought-action fusion in individuals with OCD symptoms. *Behaviour Research and Therapy, 39*, 765-776.

Beck, A. T., & Steer, R. A. (1984). Internal consistencies of the original and revised Beck Depression Inventory. *Journal of Clinical Psychology, 40*, 1365-1367.

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

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107*, 238-246.

Berle, D., & Starcevic, V. (2005). Thought-action fusion: Review of the literature and future directions. *Clinical Psychology Review*, *25*, 263-284.

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy, 41*, 1411-1426.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1992). Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. *Behaviour Research and Therapy, 30*, 33-37.

Brown, T. A., & Barlow, D. H. (2002). Classification of anxiety and mood disorders. In D. H. Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed., pp. 292-327). New York, NY: Guilford Press.

Brown, T. A., & Barlow, D. H. (2005). Categorical vs. dimensional classification of mental disorders in *DSM-IV* and beyond. *Journal of Abnormal Psychology, 114*, 551-556.

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the *DSM-IV* anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment, 21*, 256-271.

Brown, T. A., Di Nardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of *DSM-IV* anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 110*, 49-58.

Brown, T. A., White, K. S., Forsyth, J. P., & Barlow, D. H. (2004). The structure of perceived emotional control: Psychometric properties of a revised Anxiety Control Questionnaire. *Behavior Therapy, 35*, 75-99.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41*, 189-225.

Clark, D. M. (1999). Anxiety disorders: Why they persist and how to treat them. *Behaviour Research and Therapy, 37*, 5-27.

Coles, M. E., Mennin, D. S., & Heimberg, R. G. (2001). Distinguishing obsessive features and worries: The role of thought-action fusion. *Behaviour Research and Therapy, 39*, 947-959.

Cougle, J. R., Lee, H. J., Horowitz, J. D., Wolitzky-Taylor, K. B., & Telch, M. J. (2008). An exploration of the relationship between mental pollution and OCD symptoms. *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 340-353.

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV: Lifetime version (ADIS-IV-L)*. New York, NY: Oxford University Press.

Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research, 27*, 387-415.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive–Complusive Inventory: Development and validation of a short version. *Psychological Assessment, 14*, 485-495.

Freeston, M. H., Rhéaume, J., & Ladouceur, R. (1996). Correcting faulty appraisals of obsessional thoughts. *Behaviour Research and Therapy, 34*, 433-446.

Hazlett-Stevens, H., Zucker, B. G., & Craske, M. G. (2002). The relationship of thought-action fusion to pathological worry and generalized anxiety disorder. *Behaviour Research and Therapy, 40*, 1199-1204.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy, 15*, 389-395.

Holzinger, K. J., & Swineford, F. (1937). The bifactor method. *Psychometrika, 2*, 41-54.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Jöreskog, K., & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood and least square methods*. Mooresville, IN: Scientific Software.

Lee, H. J., Cougle, J. R., & Telch, M. J. (2005). Thought-action fusion and its relationship to schizotypy and OCD symptoms. *Behaviour Research and Therapy, 43*, 29-41.

Lorenzo-Seva, U., & Ferrando, P. J. (2011). Factor 8.02 [Computer software]. Tarragona, Spain: Universitat Rovira i Virgili.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

Marino, T. L., Lunt, R. A., & Negy, C. (2008). Thought-action fusion: A comprehensive analysis using structural equation modeling. *Behaviour Research and Therapy, 46*, 845-853.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*, 810-819.

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*, 487-495.

Molina, S., & Borkovec, T. D. (1994). The Penn State Worry Questionnaire: Psychometric properties and associated characteristics. In G. C. L. Davey & F. Tallis (Eds.), *Worrying: Perspective on theory assessment and treatment* (pp. 265-283). New York, NY: Wiley.

Muthén, L. K., & Muthén, B. O. (1998-2009). Mplus 5.0 [Computer software]. Los Angeles, CA: Author.

Obsessive Compulsive Cognitions Working Group. (1997). Cognitive assessment of obsessive-compulsive disorder. *Behaviour Research and Therapy, 35*, 667-681.

Pourfaraj, M., Mohammadi, N., & Taghavi, M. (2008). Psychometric properties of Revised Thought-Action Fusion Questionnaire (TAF-R) in an Iranian population. *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 600-609.

Rachman, S. (1993). Obsessions, responsibility and guilt. *Behaviour Research and Therapy, 31*, 149-154.

Rachman, S., Thordarson, D. S., Shafran, R., & Woody, S. R. (1995). Perceived responsibility: Structure and significance. *Behaviour Research and Therapy, 33*, 779-784.

Rassin, E., Diepstraten, P., Merckelbach, H., & Muris, P. (2001). Thought-action fusion and thought suppression in obsessive-compulsive disorder. *Behaviour Research and Therapy, 39*, 757-764.

Rassin, E., Merckelbach, H., Muris, P., & Schmidt, H. (2001). The Thought-Action Fusion Scale: Further evidence for its reliability and validity. *Behaviour Research and Therapy, 39*, 537-544.

Rassin, E., Muris, P., Schmidt, H., & Merckelbach, H. (2000). Relationships between thought-action fusion, thought suppression and obsessive-compulsive symptoms: A structural equation modeling approach. *Behaviour Research and Therapy, 38*, 889-897.

Raykov, T. (2001a). Bias of Cronbach's alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement, 25*, 69-76.

Raykov, T. (2001b). Estimation of congeneric scale reliability via covariance structure models with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology, 54*, 315-323.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35*, 299-331.

Salkovskis, P. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy, 23*, 571-583.

Salkovskis, P. M., Richards, H. C., & Forrester, E. (1995). The relationship between obsessional problems and intrusive thoughts. *Behavioural and Cognitive Psychotherapy, 23*, 281-299.

Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy, 26*, 169-177.

Shafran, R., & Rachman, S. (2004). Thought-action fusion: A review. *Journal of Behavior Therapy and Experimental Psychiatry, 35*, 87-107.

Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 10*, 379-391.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173-180.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806-838.

Yorulmaz, O., Karanci, A. N., Bastug, B., Kisa, C., & Goka, E. (2008). Responsibility, thought-action fusion, and thought suppression in Turkish patients with obsessive-compulsive disorder. *Journal of Clinical Psychology, 64*, 308-317.

Yorulmaz, O., Yimaz, A. E., & Gençöz, T. (2004). Psychometric properties of the Thought-Action Fusion Scale in a Turkish sample. *Behaviour Research and Therapy, 42*, 1203-1214.

Zucker, B. G., Craske, M. G., Barrios, V., & Holguin, M. (2002). Thought-action fusion: Can it be corrected? *Behaviour Research and Therapy, 40*, 652-664.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.