The Metacognitions Questionnaire–30: An Examination of a Bifactor Model and Measurement Invariance Among Men and Women in a Community Sample

Thomas A. Fergus¹ and Joseph R. Bardeen²

Abstract
The Metacognitions Questionnaire–30 (MCQ-30) is a self-report measure that assesses metacognitive beliefs (i.e., beliefs about thinking). Prior research has supported a correlated five-factor model, but no known published study has examined the tenability of second-order or bifactor models of the MCQ-30. Results supported a bifactor model of the MCQ-30 in a sample of community adults from the United States (N = 785), as well as separately among men (n = 372) and women (n = 413). Multiple-groups confirmatory factor analysis supported the configural and metric/scalar invariance of the bifactor model among men and women. Results further supported the incremental validity of one of the MCQ-30 domain-specific factors in accounting for unique variance in an index of health anxiety beyond the general metacognition factor. Results provide support for a bifactor conceptualization of the MCQ-30 and the invariance of that model across men and women.

Keywords
bifactor, gender, measurement invariance, metacognition, Metacognitions Questionnaire–30, MCQ-30

The self-regulatory executive function (S-REF) model posits that metacognition (i.e., beliefs about thinking) is central to emotional disorders, as metacognitive beliefs contribute to inflexible engagement in self-regulation strategies (e.g., rumination, worry) that exacerbate and maintain emotional distress (Wells, 2009; Wells & Matthews, 1996). For example, individuals may hold beliefs about the benefits of worry (e.g., “If I worry, I will be more prepared”) that increase the use of worry to remedy inconsistencies between a current situation and a desired outcome. When worry is frequently used to self-regulate, the likelihood that one will perceive relatively benign situations as threatening increases, which in turn, results in greater emotional distress and further use of worry in an attempt to alleviate the distress. With time, individuals may develop beliefs about the uncontrollability of their thoughts (e.g., “If I cannot control my thoughts it means that I am going crazy”) that lead to negative interpretations of thinking and greater emotional distress. Persistent emotional distress leads to continued efforts to self-regulate that can further strengthen beliefs about the uncontrollability of thinking. Altering metacognitive beliefs is central to terminating this maladaptive self-regulatory process (Wells, 2009).

Extant data suggest that metacognitive beliefs are important across a wide variety of symptom presentations, including depression (Papageorgiou & Wells, 2003), generalized anxiety (Wells & Carter, 1999), health anxiety (Bailey & Wells, 2013), obsessive-compulsive (Gwilliam, Wells, & Cartwright-Hatton, 2004), panic (Wells & Carter, 2001), posttraumatic stress (Roussis & Wells, 2006), and social anxiety (Wells & Carter, 2001) symptoms. Metacognitive beliefs are relevant to other forms of psychopathology as well, including addictive behaviors (Spada, Caselli, Nikčević, & Wells, 2015), eating disorders (Cooper, Grocutt, Deepak, & Bailey, 2007), and psychosis (Morrison, French, & Wells, 2007). Wells (2009) developed metacognitive therapy (MCT) to alter metacognitive beliefs and extant data support MCT as an efficacious treatment, particularly for anxiety and depression (Normann, van Emmerik, & Morina, 2014).

To evaluate MCT and its underlying theory, it is necessary to have valid measures of metacognitive beliefs. Cartwright-Hatton and Wells (1997) created the 65-item Metacognitions Questionnaire (MCQ) to assess metacognitive beliefs in adults. Because its length made the routine use of the MCQ prohibitive, Wells and Cartwright-Hatton...
Assessment

(2004) developed a 30-item short form of the MCQ (i.e., the MCQ-30). Like the MCQ items, the MCQ-30 items are rated using a 4-point ordered-category scale ranging from 1 (do not agree) to 4 (agree very much). Wells and Cartwright-Hatton (2004) found that the MCQ-30 assesses the same five metacognitive beliefs as the MCQ: (a) positive beliefs about worry (e.g., “worrying helps me to avoid problems in the future”); (b) negative beliefs about uncontrollability and danger of worry (e.g., “my worrying is dangerous for me”); (c) cognitive confidence (e.g., “I have little confidence in my memory for words and names”); (d) need for control (e.g., “If I did not control a worrying thought, and then it happened, it would be my fault”); and (e) cognitive self-consciousness (e.g., “I think a lot about my thoughts”). The MCQ-30 scales showed adequate internal consistency (Cronbach’s $\alpha$s ranging from .72 to .93), satisfactory test-retest reliability ($r$s ranging from .59 to .87 after 30 days), and medium-to-large correlations ($r$s ranging from .25 to .73) with indices of anxiety in Wells and Cartwright-Hatton’s (2004) study. The adequacy of the five-factor model of the MCQ-30 has since been replicated in independent samples (Grøtte et al., 2016; Ramos-Cejudo, Salguero, & Cano-Vindel, 2013; Spada, Mohiyeddini, & Wells, 2008).

Studies using the MCQ-30 have tended to either examine the MCQ-30 subscales separately (e.g., Bailey & Wells, 2013; Irak & Tosun, 2008; McEvoy, Mahoney, & Moulds, 2010; Roussis & Wells, 2006) or develop a metacognition latent construct using a subset of MCQ-30 subscales for use in structural equation modeling (e.g., Spada, Nikcicvic, Moneta, & Wells, 2007; Spada, Nikcicvic, Moneta, & Wells, 2008). These approaches make at least two assumptions about the MCQ-30. One assumption is that each MCQ-30 subscale is representative of the same general construct, as suggested by the analytic strategy of using a subset of MCQ-30 subscales as indicators of a metacognition construct. The second assumption is that the MCQ-30 subscales capture unique information beyond a total scale, as suggested by the data analytic strategy of using separate MCQ-30 subscales rather than a total scale.

However, extant studies examining the structure of the adult version of the MCQ-30 have yet to conduct analyses that adequately test these two assumptions. To adequately test these assumptions, second-order or bifactor models warrant examination (Chen, West, & Sousa, 2006). A second-order model can assess the degree to which the covariation among factors is accounted for by a higher-order factor (Brown, 2015). A second-order model has been supported for a child version of the MCQ-30 (Esbjørn et al., 2013), suggesting each MCQ-30 factor is representative of the same general construct (i.e., metacognition). The tenability of such a model, though, has not yet been examined using the adult version. A bifactor model similarly suggests the presence of a general factor. A fundamental difference between second-order and bifactor models is that the bifactor model specifies direct effects of the general factor on the indicators of domain-specific factors, whereas there are no direct effects of the general factor on the indicators in a second-order model (Reise, 2012). A graphical depiction of a bifactor model of the MCQ-30 is presented in Figure 1. Bifactor models can allow for an investigation of the presence of a general factor and the degree to which domain-specific factors are meaningfully distinct from the general factor. The degree to which the domain-specific factors appear redundant with the general factor helps determine if a total scale should be used instead of subscales (Reise, 2012).

Support for a bifactor model of the MCQ-30 has important implications for researchers and clinicians. As reviewed, extant data support the adequacy of a correlated five-factor model of the MCQ-30 (Grøtte et al., 2016; Ramos-Cejudo

Figure 1. Graphical depiction of bifactor model of the MCQ-30.
Note. POS = positive beliefs about worry; NEG = negative beliefs about uncontrollability and danger of worry; CC = cognitive confidence; NC = need for control; and CSC = cognitive self-consciousness. “….” signifies intermediary indicators not depicted. Indicator error variances ($\epsilon$) not depicted.
et al., 2013; Spada, Mohiyeddini, et al., 2008; Wells & Cartwright-Hatton, 2004). As noted by Reise, Bonfiay, and Haviland (2013), support for a correlated trait model is often interpreted to indicate that subscales should be used and may call into question the use of a total scale alone. Importantly, and as further noted by Reise et al. (2013), such correlated trait models “. . . seldom inform researchers on the degree to which multidimensionality is severe enough such that the total score is uninterpretable as an indicator of a single construct or whether subscales are psychologically justified” (p. 13). Bifactor modeling more fully addresses those important questions about multidimensional item response data than do correlated trait models (Reise et al., 2013).

A bifactor model of the MCQ-30 has not yet been examined; however, an accumulating body of research has found that bifactor models tend to provide a good fit to the data when examining anxiety-related measures (e.g., Ebesutani, McLeish, Luberto, Young, & Maack, 2014; Fergus, Kelley, & Griggs, 2016; Olatunji, Ebesutani, & Abramowitz, 2016). A second-order model is nested within the bifactor, and thus, can be compared using $\chi^2$ difference evaluation to quantify the degree of improvement in model fit (Chen et al., 2006). In addition to a $\chi^2$ difference evaluation, it is important to examine statistical indices from bifactor models that help quantify the degree to which information is gained by modeling and scoring total versus subscale scores (Rodriguez, Reise, & Haviland, 2016). Through the examination of a bifactor model of the MCQ-30, the present study can help explicate the degree to which scoring and reporting MCQ-30 subscales is justified and how future structural models examining metacognitive beliefs consist of five related, but distinct, subfactors (Grøtte et al., 2016; Ramos-Cejudo et al., 2013; Spada, Mohiyeddini, et al., 2008; Wells & Cartwright-Hatton, 2004), and (b) data showing good fit of bifactor models of anxiety-related measures (e.g., Ebesutani et al., 2014; Fergus et al., 2016; Olatunji et al., 2016). Measurement invariance analyses of a bifactor model were considered exploratory, given that the present study provided the first known examination of a bifactor model of the MCQ-30.

In the present study, we examined the structure and measurement invariance of the adult version of the MCQ-30 in a community sample of adult respondents from the United States. We predicted that a bifactor model of the MCQ-30 would provide a better fit to the data than competing models. This prediction was based on (a) data suggesting that metacognitive beliefs consist of five related, but distinct, subfactors (Grotter et al., 2016; Ramos-Cejudo et al., 2013; Spada, Mohiyeddini, et al., 2008; Wells & Cartwright-Hatton, 2004), and (b) data showing good fit of bifactor models of anxiety-related measures (e.g., Ebesutani et al., 2014; Fergus et al., 2016; Olatunji et al., 2016). Measurement invariance analyses of a bifactor model were considered exploratory, given that the present study provided the first known examination of a bifactor model of the MCQ-30.

Finally, through the use of bifactor modeling, researchers can determine if a domain-specific factor relates to a criterion variable while holding the general factor constant (Brown, 2015). As such, as a preliminary investigation, we used bifactor modeling to examine if the domain-specific factors of the MCQ-30 accounted for unique variance in scores on a criterion measure (i.e., health anxiety) after accounting for the general factor (i.e., metacognition). Health anxiety—defined by cognitive-behavioral researchers as the wide range of
worry individuals can have about their health (Asmundson & Taylor, 2005)—was chosen as the criterion variable in the present study because preliminary research suggests that the MCQ-30 total and subscale scores share medium-to-large associations (rs ranging from .47 to .72) with health anxiety (Bailey & Wells, 2013). Based on those findings, using health anxiety as the criterion variable would allow for an adequate test of the incremental utility of the MCQ-30 subscales beyond the total score and provide a further examination of Bailey and Wells’s (2013) preliminary findings.

Method

Participants

The total sample consisted of 785 adults recruited from across the United States through an online crowdsourcing website. The average age was 35.3 years (SD = 11.1, range 18-69). The sample was primarily White (80.8%) and not Hispanic or Latino (91.0%). Approximately 7.8% of the sample self-identified as African American, 7.4% as Asian, and 4.0% as “other.” The sample consisted of 413 women (52.6%) and 372 men (47.4%). Women (M = 36.7, SD = 11.8) were significantly older than men (M = 33.7, SD = 10.1) and the difference was small in size, t(783) = 3.93, p < .001, Cohen’s d = 0.28. There were no significant racial, χ2(5) = 9.74, p = .083, or ethnic, χ2(1) = 1.78, p = .182, differences among men and women. In addition to the MCQ-30 (Wells & Cartwright-Hatton, 2004), participants completed the following measure.

Measure

Whiteley Index-6 (WI-6; Asmundson, Carleton, Bovell, & Taylor, 2008). The WI-6 is a six-item revised version of the original 14-item measure of the WI (Pilowsky, 1967) that assesses health anxiety. Asmundson et al.’s (2008) revised version addresses factorial instability of the original version of the WI (Welch, Carleton, & Asmundson, 2009). Following the recommendations of Welch et al. (2009), items were rated using a 5-point scale (1 = not at all to 5 = a great deal) rather than using the traditional true/false rating system (Pilowsky, 1967). The WI-6 evidences large correlations (rs of .63 to .80) with other measures of health anxiety (Fergus, 2013) and Ferguson et al. (2016) found that the WI-6 is best operationalized using a bifactor model, consisting of a general health anxiety factor and two domain-specific factors (health worry, somatic/bodily preoccupation). The WI-6 scales evidenced good internal consistency in the present study (Cronbach’s α ranging from .86 to .92).

Procedure

Participants were recruited using Amazon’s Mechanical Turk (MTurk), an Internet-based platform that allows individuals to request the completion of jobs (e.g., survey completion) for monetary compensation. Respondents completing surveys through MTurk have been found to produce high quality data and tend to be more demographically diverse than other internet samples or undergraduate samples (see Chandler & Shapiro, 2016, for a review). The present research was approved by the local institutional review board. Recruitment was limited to MTurk workers over 18 years of age and located in the United States. We followed Paolacci and Chandler’s (2014) recommendation and sought to improve data quality by restricting MTurk worker approval ratings, as research has found that “catch” questions do not improve data quality above and beyond recruiting MTurk workers with approval ratings above 95% (Peer, Vosgerau, & Acquisti, 2014). Worker specifications in the present study included restricting participation to MTurk workers who had approval ratings above 95%. Participants were required to provide electronic consent and there was no penalty for withdrawing from the study. Measures were presented in random order. On completion of the study, participants were debriefed and paid $1.50, an amount consistent with the compensation given to MTurk workers completing prior studies of similar length (Buhrmester, Kwang, & Gosling, 2011).

Data Analytic Strategy

Measurement Models. We examined four models using the total sample. The first model was a correlated five-factor model that consisted of six items (1, 7, 10, 19, 23, 28) with primary loadings on Factor I (Positive Beliefs about Worry), six items (2, 4, 9, 11, 15, 21) with primary loadings on Factor II (Negative Beliefs about Uncontrollability and Danger of Worry), six items (8, 14, 17, 24, 26, 29) with primary loadings on Factor III (Cognitive Confidence), six items (6, 13, 20, 22, 25, 27) with primary loadings on Factor IV (Need for Control), and six items (3, 5, 12, 16, 18, 30) with primary loadings on Factor V (Cognitive Self-Consciousness). No secondary loadings were modeled, but the factors were allowed to intercorrelate. In the second model, each of the MCQ-30 items loaded onto one factor (i.e., one-factor model). The third model was a second-order model, in which the correlations among the first-order factors from the correlated five-factor model were removed and direct pathways were modeled from a second-order factor to each first-order factor. The fourth model was a bifactor model in which all 30 items freely loaded on a general factor, as well as their respective domain-specific factor. The covariances of all factors were fixed to zero in the bifactor model (Brown, 2015).

Measurement Invariance. Tests of measurement invariance were examined through a multiple-groups confirmatory factor analysis framework (Brown, 2015). The best fitting model from the total sample was tested separately among
men and women to ensure that the same measurement model was supported in each group. Next, restrictive models were used to test for (a) configural invariance (equal form) and (b) metric/scalar invariance (equal factor loadings, indicator thresholds). When testing configural invariance, we simultaneously examined the adequacy of the MCQ-30 factor structure in the two groups. When testing metric/scalar invariance, we constrained factor loadings and indicator thresholds to equality. Metric/scalar invariance were simultaneously modeled because MCQ-30 item responses result in ordered-categorical data (e.g., Brown, 2015; Ebesutani et al., 2014; Olatunji et al., 2016).

**Structural Regression Model.** A structural regression model was used to examine if the domain-specific factors of the MCQ-30 related to health anxiety when holding the general metacognition factor constant. The structural regression model consisted of simultaneously modeling the bifactor model of the MCQ-30 described above and a bifactor model of the WI-6 (Fergus et al., 2016). Pathway coefficients were freely estimated from the general factor of the MCQ-30 to the general health anxiety factor and the two domain-specific health anxiety factors, as well as from the five domain-specific factors of the MCQ-30 to the general health anxiety factor and the two domain-specific health anxiety factors.

**Model Estimation and Evaluation.** All models were tested using Mplus 7.4 (Muthén & Muthén, 2015) and, given that the MCQ-30 item responses are ordered-categories, mean-and variance-adjusted weighted least squares (WLSMV) estimation was used for all models (Asparouhov, 2005). Use of WLSMV estimation mirrors the estimation used in extant studies examining the structure of the MCQ-30 (Grøtte et al., 2016). Three commonly-recommended (Brown, 2015; Hu & Bentler, 1999; Kline, 2011) fit statistics were used to evaluate the models: the comparative fit index (CFI), Tucker–Lewis fit index (TLI), and root mean square error of approximation (RMSEA). The following guidelines were used to evaluate fit: CFI and TLI should be close to .95, RMSEA should be close to .06, and the upper limit of the 90% RMSEA confidence interval should not exceed .10 (Hu & Bentler, 1999; Kline, 2011).

In addition to these fit statistics, model comparisons were evaluated as follows. First, the chi-square difference test was used (using the DIFFTEST function in Mplus; Muthén & Muthén, 2015). A significant difference between two comparable models indicates a significant decrement in model fit. However, because the difference test is affected by sample size, model testing might result in significant difference tests when differences in parameter estimates are trivial in magnitude (Brown, 2015; Kline, 2011). As such, and following the recommendations of Brown (2015) and Kline (2011), we also used alternative tests for comparing models (i.e., examining RMSEA 90% confidence intervals [CIs] and change in CFI [ΔCFI]). Differences in model fit are considered nonsignificant if models have overlapping 90% RMSEA CIs (Wang & Russell, 2005). Meade, Johnson, and Braddy (2008) identified a ΔCFI value of less than or equal to .002 and Cheung and Rensvold (2002) identified a ΔCFI value of less than or equal to .01 as representing functionally trivial differences in parameter estimates among models. Six statistical indices put forward by Rodriguez et al. (2016) were used to further evaluate the bifactor model.

**Results**

**MCQ-30 Measurement Models**

Goodness-of-fit statistics from the measurement models are presented in Table 1. In the total sample, the correlated five-factor model of the MCQ-30 generally provided adequate model fit. The CFI, TLI, and RMSEA were all close to the specified guidelines, whereas the upper limit of the RMSEA 90% CI did not exceed .10. A one-factor model was examined next for comparative purposes, as a one-factor model is nested in a correlated five-factor model (Brown, 2015). The one-factor model did not provide adequate model fit, as none of the fit statistics met the specified guidelines. The significant Δχ², nonoverlapping RMSEA 90% CIs, and ΔCFI > .01 all further supported the correlated five-factor model as providing a better fit to the data than the one-factor model. A second-order model was fit to the data next and evidenced similar model fit as the correlated five-factor model. The CFI, TLI, and RMSEA of the second-order model were all close to the specified guidelines, whereas the upper limit of the RMSEA 90% CI did not exceed .10. When a second-order model is overidentified, it can be directly compared with a first-order model using the Δχ² test (Brown, 2015). There was a significant Δχ² between the second-order and correlated five-factor models, although the overlapping RMSEA 90% CIs suggested equivalency. The ΔCFI was .002, indicating a significant decrement in model fit according to Meade et al.’s (2008) guidelines but not a significant decrement in model fit using Cheung and Rensvold’s (2002) guidelines. As such, the difference in fit between the second-order and correlated five-factor models was largely equivocal.

Finally, a bifactor model was fit to the data. With the exception of a slightly elevated RMSEA, all of the fit statistics met the specified guidelines. As noted, second-order models are nested within bifactor models (Brown, 2015). The significant Δχ², nonoverlapping RMSEA 90% CIs, and ΔCFI > .01 all supported the bifactor model as providing a better fit to the data than the second-order model. The bifactor model produced the most favorable fit statistics among the models and it demonstrated a better fit to the data relative to its nested model. The bifactor model was retained as...
the best fitting measurement model. Factor loadings from the bifactor model are presented in Table 2. With the exception of items from the Need for Control domain-specific factor, the MCQ-30 items continued to robustly load on their respective domain-specific factors after controlling for the general factor. The loadings on the domain-specific factors were almost all lower relative to the loadings on the general factor.

### Evaluation of Bifactor Model

Six indices were used to further evaluate the bifactor model (Rodriguez et al., 2016). Omega_\text{H} (\omega_H) reflects the proportion of variance in MCQ-30 scores attributable to the general factor and was .85 (error variance is 3%). The overwhelming amount of the reliable variance in MCQ-30 scores was attributable to the general factor (i.e., 88%). Omega_\text{HIS} reflects the proportion of variance in scores attributable to each domain-specific factor independent of the general factor. The \omega_{HIS} values were as follows: Factor I (Positive Beliefs about Worry) = .42 (error variance 6%), Factor II (Negative Beliefs about Uncontrollability and Danger of Worry) = .27 (error variance 5%), Factor III (Cognitive Confidence) = .50 (error variance 5%), Factor IV (Need for Control) = .15 (error variance 13%), and Factor V (Cognitive Self-Consciousness) = .54 (error variance 11%). The reliable variance in the domain-specific factors independent of the general factor was: Factor I = 45%, Factor II = 28%, Factor III = 53%, Factor IV = 17%, and Factor V = 61%. As such, a majority of reliable variance in two subscale scores (Cognitive Confidence, Cognitive Self-Consciousness) was independent of the general factor.

Explained common variance (ECV) quantifies the amount of common variance attributable to general and domain-specific factors (Rodriguez et al., 2016). The ECV value was .59 in this study, indicating that the general factor accounts for 59% of common variance and 41% of common variance is spread across the domain-specific factors. ECV values >.85 are highly suggestive that a set of items are sufficiently unidimensional (Stucky & Edelen, 2015). Item explained common variance (I-ECV) quantifies the amount of common variance for each MCQ-30 item attributable to the general factor. I-ECV values >.85 are indicative that the items contribute more to the general factor than the domain-specific factor (Stucky & Edelen, 2015). The average I-ECV value was .59 (range .11-.95), with only three items, all from the Need for Control domain-specific factor, evidencing I-ECV values >.85.

Percentage of uncontaminated correlations (PUC) characterizes the percentage of MCQ-30 item correlations contaminated by variance attributed to the general and domain-specific factor (Rodriguez et al., 2016). The PUC value for the MCQ-30 is .83, indicating that the overwhelming majority of item correlations of the MCQ-30 are attributable to the general factor. When ECV and PUC values are both greater than .70, which both values were not in the present study (i.e., ECV was <.70), relative parameter bias is expected to be slight. Relative bias quantifies the bias across parameters if items are forced into a unidimensional, versus multidimensional, structure. The average relative bias across MCQ-30 items in the present study was 17%, which is at a level suggesting multidimensionality warrants modeling (an average bias less than 10% to 15% indicates an acceptable amount of bias to model multidimensional data unidimensionally; Rodriguez et al., 2016).

### Table 1. Goodness-of-Fit Statistics for Tested Models.

<table>
<thead>
<tr>
<th>Model</th>
<th>(\chi^2)</th>
<th>df</th>
<th>(\Delta\chi^2)</th>
<th>RMSEA</th>
<th>RMSEA 90% CI</th>
<th>CFI</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (N = 785)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifactor</td>
<td>1580.61</td>
<td>375</td>
<td>—</td>
<td>.064</td>
<td>.061-.067</td>
<td>.962</td>
<td>.956</td>
</tr>
<tr>
<td>Second-order</td>
<td>2099.42</td>
<td>400</td>
<td>419.90</td>
<td>.074</td>
<td>.070-.077</td>
<td>.946</td>
<td>.942</td>
</tr>
<tr>
<td>Correlated 5-Factor</td>
<td>2053.00</td>
<td>395</td>
<td>79.95</td>
<td>.073</td>
<td>.070-.076</td>
<td>.948</td>
<td>.942</td>
</tr>
<tr>
<td>I-Factor</td>
<td>6689.27</td>
<td>405</td>
<td>1146.46</td>
<td>.141</td>
<td>.138-.144</td>
<td>.802</td>
<td>.787</td>
</tr>
<tr>
<td>Measurement invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifactor men (n = 372)</td>
<td>1025.90</td>
<td>375</td>
<td>—</td>
<td>.069</td>
<td>.064-.074</td>
<td>.958</td>
<td>.952</td>
</tr>
<tr>
<td>Bifactor women (n = 413)</td>
<td>895.90</td>
<td>375</td>
<td>—</td>
<td>.058</td>
<td>.053-.062</td>
<td>.969</td>
<td>.964</td>
</tr>
<tr>
<td>Configural</td>
<td>1921.80</td>
<td>750</td>
<td>—</td>
<td>.063</td>
<td>.060-.067</td>
<td>.964</td>
<td>.958</td>
</tr>
<tr>
<td>Metric/scalar</td>
<td>1979.41</td>
<td>858</td>
<td>151.55</td>
<td>.056</td>
<td>.053-.060</td>
<td>.968</td>
<td>.967</td>
</tr>
</tbody>
</table>

Note. RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; TLI = Tucker–Lewis fit index; UL = upper limit; LL = lower limit. Models computed using mean- and variance-adjusted weighted least squares (WLSMV) estimation. \(\Delta\chi^2\) computed using Mplus 7.4 DIFFTEST function.

*\(\Delta\chi^2\) comparing I-factor and correlated 5-factor models. *\(\Delta\chi^2\) comparing second-order and correlated 5-factors. *\(\Delta\chi^2\) comparing second-order and bifactor models. *\(\Delta\chi^2\) comparing metric/scalar and configural models.
Measurement Invariance of MCQ-30

Results from tests of measurement invariance of the bifactor model are presented in Table 1. As a precursor to tests of measurement invariance, we initially examined the fit of the bifactor model separately in the two groups (Brown, 2015). The bifactor model showed adequate model fit among men and women. With the exception of a slight elevation in the RMSEA among men, all of the fit statistics of the bifactor model met the specified guidelines in both groups. The model with constraints testing for equal form (configural invariance) suggested equivalence between men and women. With the exception of a slight elevation of the RMSEA, all of the fit indices met specified guidelines. The model with constraints testing for equal factor loadings (metric invariance) and equal indicator thresholds (scalar invariance) suggested equivalence between the two groups. All of the fit indices met specified guidelines in that model. Of note, the fit statistics for the metric/scalar invariance model were more favorable than the configural invariance model. This pattern of findings is consistent with prior measurement invariance analyses of bifactor models (e.g., Ebesutani et al., 2014; Olatunji et al., 2016). Brown (2015) described the possibility of a more parsimonious model that constrains previously freed parameters to equality, coupled with trivial changes in $\chi^2$, resulting in what appears to be improved model fit. Although there was a significant $\Delta\chi^2$, the RMSEA 90% CIs were overlapping and there was no decrement in CFI between the configural and metric/scalar invariance models. As such, the analyses supported the presence of measurement invariance.

Structural Regression Model

The incremental validity of the MCQ-30 domain-specific factors above the general factor was examined using a
structural regression model that simultaneously modeled the bifactor model of the MCQ-30 and a bifactor model of the WI-6. The structural regression model provided an adequate fit to the data ($\chi^2 = 1288.54$, $df = 540$; RMSEA = .042 [90 CI% = .039–.045]; CFI = .945; TLI = .936), with all fit indices close to or exceeding specified guidelines. Standardized beta weights from the structural regression model are presented in Table 3. The Negative Beliefs about Uncontrollability and Danger of Worry domain-specific factor accounted for unique variance in the general health anxiety factor and the Health Worry domain-specific factor independent of the general metacognition factor. Unexpectedly, the Need for Control domain-specific factor shared a negative association with the general health anxiety factor while controlling for the general metacognition factor. Interestingly, the Cognitive Self-Consciousness domain-specific factor accounted for unique variance in only the Health Worry domain-specific factor independent of the general metacognition factor.

**Discussion**

In the present study, we examined the structure and measurement invariance of the MCQ-30 (Wells & Cartwright-Hatton, 2004) among men and women. We moved beyond the measurement models examined in previous studies of the factor structure of the MCQ-30 by also examining the tenability of second-order and bifactor models of the MCQ-30. The adequacy of the correlated five-factor solution found in prior studies was replicated in this sample of community adults. Importantly, the present results support a bifactor model of the MCQ-30, consisting of a general metacognition factor and five domain-specific metacognitive beliefs. Statistical indices from the bifactor model indicated that there are nonnegligible amounts of multidimensionality that can be attributed to the five distinct MCQ-30 domain-specific factors.

The general metacognition factor accounted for the overwhelming amount of variance in MCQ-30 scores. At the same time, our results suggest that the multidimensionality of MCQ-30 item scores is substantive enough to warrant the use of subscale scores. For example, the majority of reliable variance in two MCQ-30 subscale scores was independent of the general factor and nearly all of the MCQ-30 items had I-ECV values that indicated a greater contribution of the items to the respective domain-specific factors than to the general factor. Furthermore, the ECV value did not support considering the MCQ-30 items as unidimensional and the bias across parameters, if the MCQ-30 items were forced into a unidimensional structure, was found to be above an acceptable range. Two MCQ-30 domain-specifics factors were positively related to a criterion variable independent of the effects of the general factor, thus suggesting incremental utility. Collectively, study findings generally indicate that the MCQ-30 subscales evidence meaningful independence from the general factor. As such, continued use of the MCQ-30 subscales appears warranted and researchers should consider the use of multidimensional latent variable model specifications that account for the general and domain-specific factors (e.g., bifactor models) when examining structural models that include the MCQ-30.

The Negative Beliefs about Uncontrollability and Danger of Worry domain-specific factor was associated with a general health anxiety factor independent of the effects of the general metacognition factor. Negative Beliefs about Uncontrollability and Danger of Worry also related to the Health Worry domain-specific factor independent of the general metacognition factor. This pattern of findings is consistent with Bailey and Wells (2013) who found that Negative Beliefs about Uncontrollability and Danger of Worry shared the most robust association with health anxiety. Other studies have similarly found that Negative Beliefs about Uncontrollability and Danger of Worry is the metacognitive belief domain that shares the most robust association with health anxiety. Other studies have similarly found that Negative Beliefs about Uncontrollability and Danger of Worry is the metacognitive belief domain that shares the most robust association with health anxiety. Other studies have similarly found that Negative Beliefs about Uncontrollability and Danger of Worry is the metacognitive belief domain that shares the most robust association with health anxiety. Other studies have similarly found that Negative Beliefs about Uncontrollability and Danger of Worry is the metacognitive belief domain that shares the most robust association with health anxiety.

### Table 3. Standardized Beta Weights From Structural Regression Model.

<table>
<thead>
<tr>
<th>MCQ-30 factor</th>
<th>General Factor</th>
<th>Health Worry</th>
<th>Somatic/Bodily Preoccupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GF</td>
<td>.64**</td>
<td>.02</td>
<td>.07</td>
</tr>
<tr>
<td>POS</td>
<td>-.03</td>
<td>.09</td>
<td>.15</td>
</tr>
<tr>
<td>NEG</td>
<td>.26**</td>
<td>.22**</td>
<td>.09</td>
</tr>
<tr>
<td>CC</td>
<td>.08</td>
<td>-.15</td>
<td>-.06</td>
</tr>
<tr>
<td>NC</td>
<td>-.18*</td>
<td>.08</td>
<td>-.03</td>
</tr>
<tr>
<td>CSC</td>
<td>-.07</td>
<td>.19*</td>
<td>.05</td>
</tr>
</tbody>
</table>

Note. MCQ-30 = Metacognitions Questionnaire–30; WI-6 = Whiteley Index-6; GF = General Factor; POS = Positive Beliefs about Worry; NEG = Negative Beliefs about Uncontrollability and Danger of Worry; CC = Cognitive Confidence; NC = Need for Control; CSC = Cognitive Self-Consciousness.

**p < .01, *p < .05 (2-tailed).
Health Worry domain-specific factor independent of the general metacognition factor. This finding suggests that the tendency to focus on thought processes may be particularly relevant to aspects of health worry. Researchers propose that heightened cognitive self-consciousness makes individuals more likely to monitor their thoughts, thus becoming more aware that they lack control of thought processes (Palmier-Claus, Dunn, Taylor, & Morrison, 2013). This proposal may extend to health worry. The general metacognition factor was most relevant to the general health anxiety factor, with the domain-specific metacognition factors providing a more fine-grained understanding of health worry.

The Need for Control domain-specific factor unexpectedly shared a negative association with the general health anxiety factor in the incremental validity analyses while controlling for the general metacognition factor. This finding stands in contrast to Bailey and Wells’s (2013) finding that Need for Control shared a unique positive association with health anxiety. An important difference across studies is that the present analyses controlled for the effects of the general metacognition factor. Bailey and Wells’s (2013) use of scaled scores did not allow for the simultaneous inclusion of a total score with subscale scores. The Need for Control domain-specific factor exhibited the smallest amount of reliable variance independent of the general factor (17%) in the present study among the MCQ-30 domain-specific factors. Half of the items comprising the Need for Control domain had I-ECV values suggesting that those items contribute more to the general factor than the domain-specific factor. Although requiring further empirical attention, it is possible that the unique contribution of the Need for Control domain-specific factor does not relate to criterion variables in theoretically expected ways independent of the general metacognition factor.

Results from the present study extend the findings of Ramos-Cejudo et al. (2013), who found measurement invariance of a correlated five-factor model of a Spanish translation of the MCQ-30, by suggesting measurement invariance of a bifactor model of the English version of the MCQ-30 among men and women. These findings suggest that the MCQ-30 is most accurately operationalized as consisting of a general factor and five domain-specific factors among both men and women. Additionally, higher or lower scores on the MCQ-30 (both general and domain-specific factors) can be considered to reflect true group differences in the construct rather than the MCQ-30 having items that preclude men or women from responding in comparable ways (Brown, 2015). The measurement invariance found in this study suggests that the importance of multidimensionality found using the total sample cuts across gender, indicating that the MCQ-30 subscales warrant continued use among men and women.

As such, the MCQ-30 holds promise for shedding further light onto potential gender differences in self-regulation. For example, O’Carroll and Fisher (2013) found that women scored significantly higher in negative beliefs about uncontrollability and danger of worry than men. This subscale of the MCQ-30 is highly relevant to negatively valenced thinking in general (McEvoy et al., 2010) and women are more likely to report ruminating than men (Johnson & Whisman, 2013). Additionally, Spada, Mohiyeddini, et al. (2008) found that men scored significantly higher in beliefs related to need for control than women. Need to control thoughts is highly relevant to substance use (e.g., Spada, Caselli, & Wells, 2013), with males possibly being more likely to engage in impulsive, reward-seeking behaviors to self-regulate when they perceived themselves to lack control over thought processes (Nolen-Hoeksema, 2012). Whereas such noted gender differences in self-regulation may be related to gender differences in metacognitive beliefs, it is important to note that mean differences in MCQ-30 subscales among men and women have not always been found (Grotte et al., 2016; O’Carroll & Fisher, 2013; Spada, Mohiyeddini, et al., 2008; Wells & Cartwright-Hatton, 2004). Thus, mean-level differences in MCQ-30 scores may only be modest in magnitude.

Study limitations must be acknowledged. Consistent with prior examinations of the psychometric properties of the MCQ-30 (e.g., Wells & Cartwright-Hatton, 2004; Spada, Mohiyeddini, et al., 2008), a sample of community adults was used in the present study. Research supports MTurk as a viable method for data collection (Chandler & Shapiro, 2016). Nevertheless, MTurk samples should not be considered representative of the general population (Paolacci & Chandler, 2014). For example, the generalizability of the findings would be increased through replication using respondents who more consistently endorse heightened metacognitive beliefs (e.g., individuals with emotional disorders; Wells, & Carter, 2001). The present results are further limited by having only one external criterion variable. It is possible that other MCQ-30 subscales (e.g., cognitive confidence) may uniquely contribute to criterion variables not assessed in the present study (e.g., chronic fatigue, Maher-Edwards, Fernie, Murphy, Wells, & Spada, 2011; procrastination, Spada, Hiou, & Nikčević, 2006; psychotic-like experience, Jones & Fernyhough, 2006). Moreover, the criterion variable in the present study was another self-report measure that was completed concurrently with the MCQ-30. Future research examining how the general and domain-specific MCQ-30 factors predict behavioral data and prospective changes in symptoms represents important extensions of the present results.

The present study is further limited by restricting measurement invariance analyses to gender. Age has sometimes been found to correlate with MCQ-30 scores (Grotte et al., 2016; Spada, Mohiyeddini, et al., 2008), with younger age often relating to greater MCQ-30 scores. Women were older than men in the present study. The age difference was only modest in magnitude, so the invariance of the MCQ-30...
across gender was unlikely a result of the age difference between men and women. That said, the invariance of the MCQ-30 as a function of age could not have been fully examined in the context of the present study aims because the multiple-groups confirmatory factor analysis approach used necessitates imposing categorical cutoffs (Brown, 2015). Age is a continuous variable and MIMIC (“multiple indicators, multiple causes”) models accommodate regressing item indicators on continuous covariates. MIMIC modeling does not allow for an examination of either equal form or equal factor loadings (Brown, 2015), which are important qualities to examine in the context of bifactor models (e.g., Esbesutani et al., 2014; Olatunji et al., 2016). Additionally, future samples with greater racial/ethnic diversity are needed to examine if a bifactor model of the MCQ-30 evidences ethnoracial invariance.

Limitations notwithstanding, the present results provide support for a bifactor model of the MCQ-30. The multidimensionality of the MCQ-30 scores appears substantive enough that the continued use of MCQ-30 subscale scores is recommended. The gender invariance of the MCQ-30 highlights the need for continued research and theory that explains gender differences in metacognitive beliefs.

**Declaration of Conflicting Interests**

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

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**


Irak, M., & Tosun, A. (2008). Exploring the role of metacognition in obsessive-compulsive and anxiety symptoms. *Journal...


