Evaluating a short-form Five Facet Mindfulness Questionnaire in adolescents: Evidence for a four-factor structure and invariance by time, age, and gender

Hiba Abujaradeh
University of Pittsburgh
Pittsburgh, PA
ORCID ID: 0000-0003-1061-1481

Blake A. Colaianne, MEd
Penn State University
University Park, PA
ORCID ID: 0000-0001-5111-5527

Robert W. Roeser, PhD
Penn State University
University Park, PA
ORCID ID: 0000-0003-3326-5921

Eli Tsukayama
University of Pittsburgh
Pittsburgh, PA
ORCID ID: 0000-0002-0468-971X

Brian M. Galla, PhD
University of Pittsburgh
Pittsburgh, PA
ORCID ID: 0000-0003-4895-6311

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by University of Pittsburgh School of Education departmental funds to Brian M. Galla, by the Bennett Pierce Chair in Caring and Compassion at Pennsylvania State University to Robert W. Roeser, and by the Compassion Graduate Student Fellowship from the Prevention Research Center at The Pennsylvania State University to Blake A. Colaianne. This project was also supported by the Institute of Education Sciences Grant R305B150033 to Blake A. Colaianne at The Pennsylvania State University. The views expressed in this article are the authors and do not necessarily represent views of the granting agency.

Acknowledgments: The authors are grateful to the students who participated in this research and to the school administrators, teachers, and staff who helped to facilitate it. They are also grateful to Hannah M. Fiore for her help with study coordination and data collection.

Full Citation: Abujaradeh, H., Colaianne, B. A., Roeser, R. W., Tsukayama, E., & Galla, B. M. (2020). Evaluating a short-form Five Facet Mindfulness Questionnaire in adolescents: Evidence for a four-factor structure and invariance by time, age, and gender. *International Journal of Behavioral Development*, *44*(1), 20-30. doi: https://doi.org/10.1177/016502541987303

Evaluating a short-form Five Facet Mindfulness Questionnaire in adolescents: Evidence for a four-factor structure and invariance by time, age, and gender

Abstract

Little is known about whether a widely used mindfulness measure in adults—the Five Facet Mindfulness Questionnaire (FFMQ)—is also reliable and valid in adolescents. The current study evaluated the psychometric properties of a 20-item short-form FFMQ in a sample of 599 high school students (M_{age} = 16.3 years; 49% female) living in the U.S. Students completed the FFMQ and a battery of self-report questionnaires assessing aspects of psychological well-being and social skills 3 times over the course of one academic year. Confirmatory factor analysis indicated that a modified four-factor hierarchical model (excluding the Observe subscale and 1 item from the Describe subscale) best fit the data. This four-factor, hierarchical FFMQ demonstrated evidence of measurement invariance across time, gender, and grade level. Reliabilities for the FFMQ total score and its subscales ranged from .61 to .88. The FFMQ total score, and its subscales (excluding Observe), demonstrated evidence of convergent (e.g., with self-compassion) and discriminant (e.g., with social perspective taking skills) validity. Finally, the FFMQ total score and Act with Awareness, Nonjudgment, and Nonreactivity subscales demonstrated evidence of incremental predictive validity for cross-time changes in psychological well-being outcomes (e.g., perceived stress). Overall, results provide preliminary support for the reliability and validity of a short-form FFMQ for use in high-school-age adolescents.

The topic of mindfulness with adolescents has blossomed in recent years. Mindfulness training is now delivered to numerous adolescents in schools, afterschool programs, and clinics across the U.S. and beyond. Despite enthusiasm for mindfulness training, advances in the assessment of mindfulness in adolescents have lagged somewhat (Goodman, Madni, & Semple, 2017; Pallozzi, Wertheim, Paxton, & Ong, 2017). The development of reliable and valid assessments of mindfulness in adolescents is important for at least three reasons. First, it is important to know whether current conceptualizations of mindfulness—which are rooted in Buddhist philosophy (Analayo, 2003) and clinical science (Baer, 2003) and are geared toward adults—track the expression of mindfulness in adolescents. Second, mindfulness is hypothesized to be a primary mechanism through which mindfulness training is linked to beneficial changes in psychological well-being and health outcomes in children and adolescents (Burke, 2010; Galla, 2016). Thus, measures that examine mediators of treatment effects are critical for testing core theoretical arguments in the field. Third, there is growing interest in understanding naturalistic changes in mindfulness across adolescence and how such changes may impact positive developmental outcomes (Roeser & Eccles, 2015). It is therefore imperative to develop measures that can be used for tracking developmental change across adolescence.

To advance measurement of mindfulness in adolescents, we evaluated the psychometric properties of a short-form version of the Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) in a longitudinal study of 599 high school adolescents. The FFMQ is among the most widely studied and widely used measures of mindfulness in adults (Goldberg et al., 2016). However, far less work has examined whether this measure is reliable and valid in adolescents. The aims of the current study were to test the factor

structure of the FFMQ; to examine measurement invariance across time, gender, and grade level; to assess reliability; and to test evidence for convergent, discriminant, and predictive validity.

What is Mindfulness and How Is It Measured?

Mindfulness can be defined as a sustained and receptive awareness of the present moment (Analayo, 2003). In the psychological literature, it is often conceptualized as involving two core dimensions: self-regulation of attention to present-moment experience and an attitude of nonjudgmental acceptance (Bishop et al., 2004). The first dimension involves directing and sustaining attention to present-moment subjective experience, allowing for increased awareness and recognition of ongoing thoughts, feelings, and bodily sensations. Nonjudgmental acceptance refers to the curious and nonreactive orientation taken toward whatever arises in conscious awareness. Mindfulness is characterized as both a trainable mental quality (Shapiro, Carlson, Astin, & Freedman, 2006) and a relatively stable disposition whose expression naturally varies across and within individuals (Brown & Ryan, 2003).

With some exceptions (Levinson, Stoll, Kindy, Merry, & Davidson, 2014), mindfulness is most commonly measured through self-report questionnaires (Bergomi, Tschacher, & Kupper, 2013; Sauer et al., 2013). These instruments differ primarily in their theoretical and philosophical foundations and in the number of mindfulness dimensions they emphasize (Bergomi et al., 2013; Van Dam et al., 2018). Some measures attempt to capture a single dimension of the construct. For example, the Mindful Attention Awareness Scale (Brown & Ryan, 2003), which is rooted in Buddhist philosophy (Bodhi, 2011) and Self-Determination Theory (Ryan & Deci, 2000) assesses the degree to which individuals are attentive and aware of their experience in daily life. Other measures capture a multidimensional characterization of mindfulness. For example, the Kentucky Inventory of Mindfulness Skills (Baer, Smith, & Allen, 2004), which

derives mainly from clinical science, assesses four interrelated but separable dimensions of mindfulness, including some that may not directly overlap with the proposed operationalization by Bishop et al. (2004) (e.g., describing experience with words).

FFMQ

The diverse array of mindfulness measures prompted Baer, Smith, Hopkins, Krietemeyer, and Toney (2006) to examine the factor structure of existing self-report measures. Their aim was to understand whether mindfulness—as it was being measured at the time—was better characterized as a unidimensional or multidimensional construct, and if the latter, the number of its constituent dimensions. Factor analysis of a combined pool of 112 items across five different self-report scales yielded five distinguishable, yet correlated dimensions of mindfulness: Observing present-moment experience, Describing experience with language, Acting with Awareness, Nonjudgment of experience, and Nonreactivity to inner experience. Four of the five dimensions (Describe, Acting with Awareness, Nonjudgment, and Nonreactivity) appeared to capture an overall mindfulness construct, especially among non-meditating adults. The Observe scale (assessed via items such as, “I pay attention to sensations, such as the wind in my hair or sun on my face”) did not load onto an overall mindfulness factor and was unexpectedly positively correlated with psychological symptoms and other measures of distress. This suggested that in the absence of formal mindfulness training, the Observe subscale may capture maladaptive forms of self-focused attention (e.g., self-consciousness).

Several studies have since replicated the original finding that a four-factor hierarchical model, excluding Observe, provides optimal fit for the data among non-meditators (Curtiss & Klemanski, 2014; Gu et al., 2016; Williams, Dalgleish, Karl, & Kuyken, 2014). On the other hand, studies have also found support for a five-factor correlated and a five-factor hierarchical

model as providing the best fit in individuals with no meditation experience (Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011; Veehof, ten Klooster, Taal, Westerhof, & Bohlmeijer, 2011). In addition, Medvedev, Siegert, Kersten, and Krägeloh (2017) suggested that certain items (e.g., Item 32 of the Describe subscale, "My natural tendency is to put my experiences into words") may not function particularly well psychometrically, perhaps due to vague wording of the statements. These authors suggest further that removing such items improves the model fit.

Beyond evaluating its factor structure, studies have also examined reliability and validity of the FFMQ. Studies find evidence for reliabilities > .60 on each of the five subscales (Christopher, Neuser, Michael, & Baitmangalkar, 2012) and test–retest reliability (Veehof et al., 2011). Likewise, studies document evidence for convergent validity with conceptually more closely related constructs (Hollis-Walker & Colosimo, 2011), discriminant validity from less closely related constructs (Gu et al., 2016; Veehof et al., 2011), and incremental predictive validity (Christopher et al., 2012; Veehof et al., 2011). The dimensions of Acting with Awareness, Nonjudgment, and Nonreactivity have repeatedly demonstrated incremental predictive validity for psychological symptoms and well-being (Desrosiers, Klemanski, & Nolen-Hoeksema, 2013; Medvedev, Norden, Krägeloh, & Siegert, 2018; van Son et al., 2014; Veehof et al., 2011).

The original version of the FFMQ is 39 items, which can raise concerns of survey fatigue for large assessment batteries. Attempts to shorten the FFMQ have been developed, with varying numbers of items: 24 items (Bohlmeijer et al., 2011), 20 items (Hou, Wong, Lo, Mak, & Ma, 2013; Tran, Glück, & Nader, 2013), and 15 items (Gu et al., 2016). These scales also demonstrate initial evidence of construct validity (Bohlmeijer et al., 2011), and lower, but adequate reliability (Gu et al., 2016; Tran et al., 2013). Of course, reductions in scale reliabilities

are often associated with use of fewer items (Ziegler, Kemper, & Kruyen, 2014). The current study sought to evaluate the 20-item short form developed by Tran, Glück, and Nader (2013).[1] We selected this version of the scale based on the fact that it reduced the FFMQ by about 50%, was validated in nonclinical adult samples (compared to the other short forms that were validated in clinical or subclinical samples and in intervention trials), and retained more than 3 items per facet (thus enabling comparison of model fit statistics).

Is the FFMQ Reliable and Valid In Adolescents?

Although the FFMQ has been used in prior studies with adolescent samples (Calvete, Gámez-Guadix, & Cortazar, 2017; Ciesla, Reilly, Dickson, Emanuel, & Updegraff, 2012; Galla, 2016; Royuela-Colomer & Calvete, 2016), to our knowledge only one study has explicitly sought to evaluate the psychometric properties of the FFMQ (Royuela-Colomer & Calvete, 2016). In this study, 520 Spanish adolescents (age range = 13–19 years) completed a translated version of the FFMQ. A five-factor, correlated model provided the best fit to the data in this Spanish-language version of the FFMQ (αs ranged from .65 to .83). A random subsample of 247 adolescents completed the same FFMQ 4 months after the first assessment. Test–retest reliability in this subsample was generally adequate (intraclass correlations between .39 and .63). Longitudinal analysis also revealed that Nonreactivity and Acting with Awareness subscales predicted reductions in depressive symptoms 4 months later, controlling for baseline levels. Overall, this study offered an important initial evaluation of the psychometric properties of the FFMQ in adolescents. However, a comparable evaluation has not yet occurred in adolescents from North America, so it remains unknown whether these results would replicate in other samples. Moreover, the study did not evaluate whether the factor structure of the FFMQ was

invariant across time, gender, and grade level—a foundational requirement for tracking longitudinal change and making group comparisons (Widaman, Ferrer, & Conger, 2010).

The Current Study

We examined the reliability and validity of a short-form version of the FFMQ in a longitudinal study of 599 high-school-age adolescents in the U.S. The aims of this study were fourfold: First, we sought to evaluate the factor structure of the 20-item FFMQ. Based on prior research, we anticipated that a four-factor model (excluding Observe) would fit the data better than a five-factor model. Second, we tested measurement invariance in the factor structure across time, gender, and grade level. Third, we examined evidence of convergent validity in terms of closely related psychological constructs (e.g., self-compassion) and discriminant validity in terms of less closely related social skills (e.g., social perspective taking). Finally, we tested incremental predictive validity of mindfulness with regard to cross-time changes in psychological well-being (e.g., life satisfaction).

Method

Participants

The sample included 599 students (M_{age} = 16.27 years, SD = 1.15, range = 13.92–19.67) attending a large suburban public high school in the Northeastern U.S. Students were recruited through a random selection of teachers in each grade level. The majority of students (80%) self-identified as Caucasian, and 49% identified as female, which is representative of the school's population (88% White, 49% female; National Center for Education Statistics, 2018). Data on socioeconomic status were not collected, but according to the National Center for Education Statistics (2018), approximately 12% of the school's students qualify for free or reduced price lunches. The analytic sample represented about one third of the school population.

Approximately 19% of the sample were freshmen, 24% were sophomores, 28% were juniors, and 29% were seniors.[2]

Procedure

The study was approved by the University of Pittsburgh Institutional Review Board. Data collection spanned nearly an entire academic calendar year (beginning in September and ending in April). The school sent an informational letter about the study and an opt-out permission form to parents. Students also completed assent forms during the first assessment. Students who were not available during the first assessment were given one more opportunity to provide assent during the second assessment. This means that some students did not provide data for the study until later in the academic year. Students were included in the study as long as they provided data during at least the first or the second assessment wave (and thus provided assent).

Students completed self-report measures assessing study constructs 3 times during the academic year. The three assessment waves, henceforth referred to as T1, T2, and T3, were spaced approximately 3 months apart (September, January, April). All measures were completed using Qualtrics Survey System on school computers during regular school hours. Students' responses to a single attention check embedded in each survey ("For this question, select 'rarely true'") suggest that they were mostly attentive when completing the survey (percent correct responses: T1 = 90%, T2 = 85%, and T3 = 83%).

Approximately 66% (n = 395) of students took all three surveys, 27% (n = 163) took two surveys, and 7% (n = 41) took one survey. Survey completion rates did not differ by gender (girls versus boys), $\chi^2(2) = 1.37$, $p = .503$, or race (Caucasian versus other races/ethnicities), $\chi^2(2) = 2.78$, $p = .249$, but they did differ by grade level (lower level [freshmen, sophomores] versus

upper level [juniors, seniors] students), $\chi^2(2) = 24.28$, $p < .001$, with older students being more likely to have taken fewer surveys.

Measures

Mindfulness. Students completed a 20-item version of the FFMQ taken from prior research (Tran et al., 2013). The FFMQ assesses individual differences in five facets of mindfulness, including Acting with Awareness (Items 5, 8, 13, and 18; see Baer et al., 2006), Describe (Items 16, 22, 32, and 37), Observe (Items 15, 20, 26, and 31), Nonjudgment (Items 14, 25, 30, and 35), and Nonreactivity (Items 9, 19, 21, and 24). Items were rated from 1 = *never or very rarely true* to 5 = *very often or always true*. The Online Supplementary material provides content for each FFMQ item used.

Convergent and discriminant validity measures

Self-compassion. Participants completed the 12-item Self-Compassion Scale, Short Form (Raes, Pommier, Neff, & Van Gucht, 2011). This scale taps facets related to self-compassion, including self-kindness, self-judgment, common humanity, balanced awareness, isolation, and overidentification. Items were endorsed from 1 = *strongly disagree* to 6 = *strongly agree* (αs = .86, .88, .86 across three time points, respectively).

Social perspective taking. Students also reported their tendency to adopt the point of view of others using 4 items taken from the Interpersonal Reactivity Index (Davis, 1980). Items (e.g., "Before criticizing somebody, I try to imagine how I would feel if I were in their place") were rated from 1 = *does not describe me well* to 5 = *describes me very well* (αs = .80, .82, .84).

Empathic concern. Students reported their tendency to experience feelings of sympathy and compassion for individuals who are less fortunate using 3 items taken from the Interpersonal Reactivity Index (Davis, 1980). Items (e.g., "I often have tender, concerned feelings for people

less fortunate than me") were rated from 1 = *does not describe me well* to 5 = *describes me very well* (αs = .76, .79, .82).

Prosocial behavior

Students reported how often they engage in prosocial interpersonal behaviors using items adapted from prior research (Caprara, Steca, Zelli, & Capanna, 2005; Segal, Cimino, Gerdes, Harmon, & Wagaman, 2013). Items (e.g., "I try to help others who are in need," "I help others even if it does not personally benefit me") were rated from 1 = *never/almost never true* to 5 = *almost always/always true* (αs = .87, .87, .89).

Psychological well-being outcomes

Satisfaction with life. Students reported on their global cognitive judgments of life satisfaction using the 5-item Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). Items (e.g., "In most ways my life is close to my ideal") were rated from 1 = *strongly disagree* to 7 = *strongly agree* (αs = .86, .88, .86).

Perceived stress. Students reported on the degree to which they have recently felt their life was stressful, unpredictable, and uncontrollable using the 4-item Perceived Stress Scale (e.g., During the past month "…how often have you felt that you were unable to control the important things in your life?"; Cohen, Kamarck, & Mermelstein, 1983). Items were endorsed using a 5-point scale from 1 = *never* to 5 = *very often* (αs = .71, .68, .69).

Positive and negative affect. Students reported their positive and negative affect in the past month using the 10-item Positive and Negative Affectivity Schedule, Short Form (Mackinnon et al., 1999). Five items were used to capture positive affect (i.e., alert, excited, enthusiastic, inspired, and determined; αs = .76, .79, .79) and 5 items were used to capture negative affect

(i.e., distressed, upset, scared, nervous, and afraid; αs = .85, .85, .87). Items were endorsed from 1 = *not at all* to 5 = *extremely*.

Rumination. Students completed 4 items adapted from the Multidimensional Measure of Academic Coping (Skinner, Pitzer, & Steele, 2013). The rumination subscale assesses the tendency to dwell on negative or stressful life events. Items (e.g., When something bad or stressful happens to me, "I keep thinking about it over and over") were endorsed from 1 = *not at all true for me* to 4 = *very true for me* (αs = .95, .95, .95).

Demographic covariates

Because research has shown gender and age differences in traits related to mindfulness (e.g., self-compassion, Bluth, Campo, Futch, & Gaylord, 2017), we included students' self-reported gender (0 = *male*, 1 = *female*) and grade level as covariates in analyses. In the current study, grade level was treated as a dichotomous variable (lower level [freshmen, sophomores] vs. upper level [juniors, seniors] students) to provide adequate sample size for multigroup measurement invariance tests.

Data analysis

Descriptive statistics were conducted using SPSS v24 (IBM Corp, 2016). All other analyses were completed in Mplus v7.2 (Muthén & Muthén, 2012). Students with missing data were included in all models using full information maximum likelihood (FIML), which produces less biased and more reliable results than listwise or pairwise deletion (Baraldi & Enders, 2010; Schafer & Graham, 2002). FIML was enabled by treating demographic characteristics as missing data correlates (i.e., auxiliary variables). All analyses used MLR estimation (maximum likelihood estimation with robust standard errors).

Model fit was assessed using standard indices and their corresponding cutoffs. Values of .90 or higher for the comparative fit index (CFI) indicate acceptable fit to the data, and values of .95 or higher indicate excellent fit (Bentler & Bonett, 1980; Hu & Bentler, 1999; Schumacker & Lomax, 2010). Root mean square error of approximation (RMSEA) values of .08 or less indicate acceptable fit, and values of .05 or less indicate excellent fit (Browne & Cudeck, 1993; Schumacker & Lomax, 2010). We also used Bayesian information criteria (BIC) values to compare models, where lower values indicate better fit.

To test for differences in fit across nested models, we followed recommendations where a change in model fit of <.010 in CFI and a change of <.015 in RMSEA would indicate that the more restrictive model does not fit worse than the unrestricted model (Chen, 2007). We supplemented these indices with χ^2 difference tests using MLR correction (Satorra & Bentler, 2001).

Results

Structural Analysis of the FFMQ

We examined the factor structure of the FFMQ through a series of confirmatory factor analyses using T1 data. Following prior research (Baer et al., 2006; Gu et al., 2016; Williams et al., 2014), we tested five measurement models: (1) a single-factor model in which all items served as indicators of an overall, latent mindfulness factor; (2) a five-factor correlated model in which items served as indicators of five separate but correlated factors; (3) a five-factor hierarchical model in which items served as indicators of five factors that in turn served as indicators for an overall, higher order mindfulness factor; (4) a four-factor correlated model in which items served as indicators of four separate but correlated factors (excluding Observe subscale); and (5) a four-factor hierarchical model in which items served as indicators of four

factors that in turn served as indicators for an overall, higher order mindfulness factor (excluding Observe subscale).

Table 1 presents fit indices for all models testing the factor structure of the 20-item FFMQ. The single-factor model demonstrated poor fit to the data. The five-factor correlated and five-factor hierarchical models demonstrated adequate fit to the data (see Online Supplementary material for standardized loadings in the five-factor hierarchical model). In the five-factor correlated model, however, Observe demonstrated either negative ($r_{\text{Acting with Awareness}} = -.18$, $p = .007$; $r_{\text{Nonjudgment}} = -.17$, $p = .007$) or nonsignificant ($r_{\text{Describe}} = -.02$, $p = .790$; $r_{\text{Nonreactivity}} = .10$, $p = .227$) correlations with the remaining four factors. Likewise, in the five-factor hierarchical model, Observe loaded negatively onto the second-order mindfulness factor (standardized loading = −.17, $SE = .08$, $p = .039$). Excluding the Observe factor qualitatively improved model fit for both the four-factor correlated and four-factor hierarchical models, although doing so still did not result in excellent fit to the data.

To explore the factor structure further, we examined the standardized factor loadings for each indicator in the four-factor hierarchical model. All items, except for Item 32 on the Describe subscale, showed moderate-to-strong loadings on their respective factors (standardized loadings = .37 to .86, ps < .001). By contrast, the standardized factor loading for Item 32 was considerably smaller (standardized loading = .18, $p = .002$). Moreover, it was the only item whose latent factor did not account for statistically significant variance ($R^2 = .03$, $SE = .02$, $p = .125$).

Based on these results, we refitted the above set of measurement models, but excluded the Observe factor and Item 32 in the Describe factor. This resulted in 15 items in total. Both the four-factor correlated and four-factor hierarchical models provided excellent fit to the data. We retained the 15-item four-factor hierarchical model in all remaining analyses, since it fit the data

as well as the four-factor correlated model, $\Delta\chi^2(2) = 4.33$, $p = .115$, and represented a theorized conceptualization of the FFMQ.

Measurement Invariance

We next examined measurement invariance of the 15-item, four-factor hierarchical model.[3] We tested configural, metric, and scalar invariance across (1) time (assessment wave), (2) gender (girls [$n = 293$] vs. boys [$n = 306$]), and (3) grade level (lower level [$n = 260$] vs. upper level [$n = 339$] students). Measurement invariance tests for gender and grade level were conducted using T1 data. Following prior research (Dimitrov, 2010; Rudnev, Lytkina, Davidov, Schmidt, & Zick, 2018), we fit a series of five models: (1) *No invariance* (Model 1): This is the baseline model in which no invariance is assumed (i.e., all model parameters are freely estimated; (2) *Invariant first-order loadings* (Model 2): Model 2 is obtained from Model 1 by adding equality constraints to all first-order factor loadings across groups; (3) *Invariant first-order and second-order factor loadings* (Model 3): Model 3 is obtained from Model 2 by adding equality constraints to all second-order factor loadings across groups; (4) *Invariant first-order and second-order factor loadings and item intercepts* (Model 4): Model 4 is created from Model 3 by adding equality constraints to all item intercepts across groups; and (5) *Invariant first-order and second-order factor loadings, item intercepts, and first-order factor intercepts* (Model 5): Model 5 is created from Model 4 by adding equality constraints to all first-order factor intercepts across groups. Full results of these models are summarized in Table 2.

Time (assessment wave)

All models provided excellent fit to the data. The four-factor hierarchical model demonstrated evidence of both metric (i.e., factor loadings equal) invariance (Model 1 vs. Model 2: ΔCFI = 0; ΔRMSEA = 0; Model 2 vs. Model 3: ΔCFI = 0; ΔRMSEA = 0) and scalar (i.e.,

intercepts equal) invariance (Model 3 vs. Model 4: ΔCFI = −.003; ΔRMSEA = .001; Model 4 vs. Model 5: ΔCFI = −.001; ΔRMSEA = .001). Thus, the factor structure of the FFMQ was equivalent across the September, January, and April assessment waves.

Gender

All models provided adequate-to-excellent fit to the data. The data showed that metric invariance (Model 1 vs. Model 2: ΔCFI = .002; ΔRMSEA = −.003; Model 2 vs. Model 3: ΔCFI = 0; ΔRMSEA = 0), but not scalar invariance (Model 3 vs. Model 4: ΔCFI = −.011; ΔRMSEA = .003; Model 4 vs. Model 5: ΔCFI = −.005; ΔRMSEA = .002), could be established. Thus, the data support metric invariance for the FFMQ across girls and boys.

Grade level

The configural, metric, and scalar invariance models all provided excellent fit to the data. The four-factor hierarchical model demonstrated evidence of both metric (Model 1 vs. Model 2: ΔCFI = −.002; ΔRMSEA = −.001; Model 2 vs. Model 3: ΔCFI = .001; ΔRMSEA = −.001) and scalar (Model 3 vs. Model 4: ΔCFI = .003; ΔRMSEA = −.002; Model 4 vs. Model 5: ΔCFI = −.002; ΔRMSEA = 0) invariance. Thus, the factor structure of the FFMQ was equivalent for lower level (freshmen, sophomores) and upper level (juniors, seniors) high school students.

Descriptive Statistics and Bivariate Correlations

Table 3 provides descriptive statistics for the four subscale scores and the overall FFMQ score at each assessment wave. Acting with Awareness, Nonjudgment, and the overall score showed higher internal reliability consistency estimates, with αs ranging from .81 to .88. Estimates of internal reliability were lower for Nonreactivity and Describe, with αs ranging from .61 to .69.

Table 4 presents the bivariate correlations among the four FFMQ subscales within and across each assessment wave. All correlations were in the expected direction and were statistically significant at $p < .01$. However, the magnitude of the correlations within assessment wave varied considerably, ranging from $r = .16$ to .43.

Evidence of Convergent and Discriminant Validity

Table S2 in the Online Supplementary material shows bivariate correlations between the FFMQ and self-compassion, social perspective taking, empathic concern, and prosocial behavior. As expected, the FFMQ subscales and total score demonstrated convergent validity with self-compassion (rs = .25 to .68) and discriminant validity from less closely related constructs (perspective taking, rs = −.14 to .28; empathic concern, rs = −.11 to .11; prosocial behavior, rs = −.07 to .18).

Evidence of Incremental Predictive Validity

Finally, we tested incremental predictive validity of FFMQ subscales for each psychological well-being outcome. We fit a series of multiple regression analyses in which each T2 psychological well-being outcome was simultaneously regressed on T1 FFMQ subscales, demographic covariates (gender, grade level), and T1 psychological well-being. As presented in Table 5, FFMQ subscales were differentially associated with changes in each outcome. Acting with Awareness predicted significant decreases in perceived stress and negative affect. Nonjudgment predicted significant decreases in rumination, perceived stress, and negative affect. Nonreactivity predicted significant decreases in perceived stress and increases in positive affect. Describe did not demonstrate evidence of incremental predictive validity for any outcome.

We then reran the analyses using the FFMQ total score as a predictor of outcomes. Controlling for T1 levels of the outcome and demographic covariates, the FFMQ total score

demonstrated evidence of incremental predictive validity for all five outcomes: life satisfaction (B = .31, 95% CI = [.11, .50], p = .003, β = .13); perceived stress (B = −.30, 95% CI = [−.41, −.18], p < .001, β = −.25); positive affect (B = .13, 95% CI = [.02, .24], p = .024, β = .10); negative affect (B = −.17, 95% CI = [−.30, −.04], p = .010, β = −.11); and rumination (B = −.19, 95% CI = [−.30, −.09], p < .001, β = −.12).

Discussion

In a study of 599 high school students, we evaluated the psychometric properties of a short-form version of the Five Factor Mindfulness Questionnaire. Results revealed that a modified four-factor hierarchical model, excluding the Observe subscale and 1 item from the Describe subscale, best fit the data. This 15-item four-factor model demonstrated evidence of configural, metric, and scalar invariance across time and grade level, and configural and metric invariance across gender. Reliabilities for the FFMQ total score and Nonjudgment and Acting with Awareness subscales were higher (.81 to .88), while the reliabilities for Nonreactivity and Describe subscales were lower (between .61 and .69). These four subscales were positively correlated with one another within and across assessment waves. Likewise, the FFMQ total score and the four subscales demonstrated evidence of convergent validity with conceptually more closely related constructs (self-compassion) and discriminant validity from conceptually less closely related constructs (e.g., social perspective taking, empathic concern).[4] Finally, Acting with Awareness, Nonjudgment, and Nonreactivity subscales showed evidence of incremental predictive validity for 3-month changes in psychological well-being outcomes, above and beyond demographic covariates and baseline levels of psychological well-being.

The current results suggest that a mindfulness measure originally intended for use with adults can also be reliable and valid in adolescent samples that share demographic characteristics

with this sample. In particular, the four-factor hierarchical structure of the FFMQ in our adolescent sample replicates findings from numerous other studies with adult samples (Baer et al., 2006; Curtiss & Klemanski, 2014; Gu et al., 2016; Williams et al., 2014). Likewise, reliability estimates and correlations among mindfulness facets are consistent with prior validation studies in adults (Gu et al., 2016; Tran et al., 2013). Interestingly, our results also align with other studies showing that certain items (Item 32; Medvedev et al., 2017), perhaps due to ambiguous wording, may not load properly onto the factor scores. Importantly, our results also replicate prior work showing that Acting with Awareness, Nonjudgment, and Nonreactivity incrementally predict psychological well-being outcomes (Bohlmeijer et al., 2011; Veehof et al., 2011).

Our results did not fully replicate those of Royuela-Colomer and Calvete (2016), who undertook the only other formal evaluation (at the time of this study) of the FFMQ in adolescents. Despite the samples being roughly equivalent in terms of age and gender, they found that a five-factor correlated model best fit the data. Another key difference is that our results showed that Acting with Awareness, Describe, Nonjudgment, and Nonreactivity were all positively and significantly correlated, whereas they found that Nonreactivity was uncorrelated with Acting with Awareness and Nonjudgment. In both studies, Nonreactivity and Acting with Awareness showed incremental predictive validity for changes in psychological well-being; we found associations with positive and negative affect and perceived stress and they found associations with depression symptoms. However, we found additional evidence for the predictive validity of Nonjudgment on changes in negative affect, perceived stress, and rumination. Replication studies are needed to parse these results, but it is possible that cultural differences can explain some of the discrepant findings. Our sample was drawn from a high

school in the U.S. composed primarily of Caucasian students from middle and higher income households, whereas Royuela-Colomer and Calvete sampled high school students from Spain who were more socioeconomically diverse.

Theoretical and Practical Implications

What are the scientific implications of this study? The current investigation contributes to evidence that the FFMQ reflects a hierarchical model of mindfulness in adolescents. This is notable because the original derivation of the FFMQ was based on adult samples and conceptualizations of mindfulness rooted in both clinical science and Buddhism (Baer et al., 2006). Replicating the factor structure in adolescents was therefore not a forgone conclusion. Future research should study whether a five-factor model, with Observe, fits the data better following mindfulness training programs, as has been found in adult samples.

This study also contributes to evidence that mindfulness, as assessed through the FFMQ, predicts psychological well-being in adolescents. This is one of a growing number of studies in adolescents to show that different dimensions of mindfulness (Acting with Awareness, Nonreactivity, Nonjudgment) prospectively predict incremental changes in various aspects of psychological well-being. Adolescents with higher scores on Acting with Awareness, Nonjudgment, and Nonreactivity all reported reductions in perceived stress 3 months later. Our data showed a pattern of differential predictive validity for other outcomes: Acting with Awareness and Nonjudgment incrementally predicted reductions in negative affect, but only Nonjudgment predicted reductions in rumination, and only Nonreactivity predicted increases in positive affect. These results affirm the value of considering how specific dimensions of mindfulness may be more predictive of different aspects of well-being.

This research makes an important advance by showing that the FFMQ is invariant across time and grade level. Our data suggest that the FFMQ reflects the same underlying construct regardless of whether it is administered in fall, winter, or spring and regardless of whether participants are in lower level (9th, 10th) or upper level (11th, 12th) grades in high school. Thus, this scale may be useful for researchers interested in charting change in mindfulness across time and age (Roeser & Eccles, 2015). At the same time, when looking across gender, our data revealed equivalent item loadings (metric invariance) but not intercepts (scalar invariance). This suggests that it is possible to compare the magnitude of correlations between the latent FFMQ factors and other outcomes (e.g., perceived stress) across boys and girls, but that comparing latent means of the FFMQ across boys and girls (e.g., do boys have higher mindfulness than girls) may not be yet warranted with this scale.

What are the practical implications of the current findings? The results indicate that the Observe subscale may not be a valid indicator of mindfulness in adolescents. It did not load onto the hierarchical mindfulness factor, and it was negatively or nonsignificantly correlated with the remaining four factors. Consistent with prior work in adults (Baer et al., 2006; Veehof et al., 2011; Williams et al., 2014), this suggests that even during adolescence, individuals may interpret Observe items as being more reflective of self-conscious attention rather than dispassionate awareness of ongoing perceptual experience. We therefore caution against its use in either the total mindfulness score or as a separate facet of mindfulness.

The lower reliabilities among Nonreactivity and Describe suggest room for improving the content of these scales. With short-form scales, lower reliabilities are expected and may be an acceptable trade-off for other advantages they bring to basic research (e.g., efficiency of measurement) (Ziegler et al., 2014). We are quick to note too, that despite the low reliabilities,

both the Nonreactivity and Describe scales did load significantly onto a hierarchical mindfulness factor (suggesting evidence of construct validity) and also demonstrated evidence of convergent and discriminant validity. Describe did not demonstrate evidence of incremental predictive validity for changes in psychological well-being outcomes. Of course, this may be due to the outcomes assessed, but overall, the utility of Describe for predicting life outcomes remains to be seen with this short-form scale.

While our analysis shows the theoretical value of considering FFMQ scales separately for testing specific mechanisms of mindfulness, researchers may also use the FFMQ total score. It too demonstrated excellent reliability and showed evidence for convergent, discriminant, and incremental predictive validity. Moreover, confirmatory factor analysis indicated that the subscales do form a hierarchical "mindfulness" factor. In situations where researchers are not interested in testing incremental predictive validity of specific subscales, it appears that the FFMQ total score may be substituted.

Limitations

This study has several limitations that suggest useful directions for future research. Time constraints for school testing did not permit inclusion of other mindfulness measures (e.g., Child and Adolescent Mindfulness Measure; Greco, Baer, & Smith, 2011), so future research should examine the strength of relationships with this short-form FFMQ. Future studies should also evaluate whether the psychometric properties of this short-form FFMQ replicate across more diverse samples of adolescents. Although the FFMQ demonstrated initial evidence of measurement invariance across time, grade level, and gender, more research is required to determine whether it is invariant before and after mindfulness training. Additionally, tests of incremental predictive validity relied on self-report questionnaires of psychological well-being.

Future research should include behavioral measures, informant ratings, and experience-sampling methods to further assess the predictive validity of the FFMQ and its factors.

Conclusion

The current study provides preliminary evidence that a short-form adaptation of the FFMQ may be reliable and valid in adolescents. A four-factor hierarchical model, excluding the Observe subscale and one of the Describe subscale items, demonstrated excellent fit to the data. This factor structure was invariant across time, grade level, and (partially) gender, making it a potentially useful tool for longitudinal studies charting developmental change, and for studies interested in making comparisons across grade level and developmental time. This multidimensional scale also enables researchers to test theoretical questions to understand which aspects of mindfulness are associated with improving psychological and health outcomes.

References

Analayo B. (2003). *Satipaṭṭhāna: The direct path to realization*. Birmingham, UK: Windhorse.

Baer R. A., Smith G. T., Allen K. B. (2004). Assessment of mindfulness by self-report: The Kentucky inventory of mindfulness skills. *Assessment*, 11, 191–206. doi:10.1177/1073191104268029

Baer R. A., Smith G. T., Hopkins J., Krietemeyer J., Toney L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13, 27–45. doi:10.1177/1073191105283504

Baraldi A. N., Enders C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, 48, 5–37. doi:10.1016/j.jsp.2009.10.001

Bentler P. M., Bonett D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606. doi:10.1037/0033-2909.88.3.588

Bergomi C., Tschacher W., Kupper Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. *Mindfulness*, 4, 191–202. doi:10.1007/s12671-012-0110-9

Berry D. R., Cairo A. H., Goodman R. J., Quaglia J. T., Green J. D., Brown K. W. (2018). Mindfulness increases prosocial responses toward ostracized strangers through empathic concern. *Journal of Experimental Psychology: General*, 147, 93–112. doi:10.1037/xge0000392

Bishop S. R., Lau M., Shapiro S. L., Carlson L. E., Anderson N. D., Carmody J., …Devins G. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice*, 11, 230–241. doi:10.1037/0033-2909.88.3.588

Bluth K., Campo R. A., Futch W. S., Gaylord S. A. (2017). Age and gender differences in the associations of self-compassion and emotional well-being in a large adolescent sample. *Journal of Youth and Adolescence*, 46, 840–853. doi:10.1007/s10964-016-0567-2

Bodhi B. (2011). What does mindfulness really mean? A canonical perspective. *Contemporary Buddhism*, 12, 19–39. doi:10.1080/14639947.2011.564813

Bohlmeijer E., ten Klooster P. M., Fledderus M., Veehof M., Baer R. (2011). Psychometric properties of the Five Facet Mindfulness Questionnaire in depressed adults and development of a short form. *Assessment*, 18, 308–320. doi:10.1177/1073191111408231

Brown K. W., Ryan R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84, 822–848. doi:10.1037/0022-3514.84.4.822

Browne M. W., Cudeck R. (1993). Alternative ways of assessing model fit. In Bollen K. A., Long J. S. (Eds.), *Testing structural models* (pp. 136–162). Newbury Park, CA: Sage.

Burke C. A. (2010). Mindfulness-based approaches with children and adolescents: A preliminary review of current research in an emergent field. *Journal of Child and Family Studies*, 19, 133–144. doi:10.1007/s10826-009-9282-x

Calvete E., Gámez-Guadix M., Cortazar N. (2017). Mindfulness facets and problematic Internet use: A six-month longitudinal study. *Addictive Behaviors*, 72, 57–63. doi:10.1016/j.addbeh.2017.03.018

Caprara G. V., Steca P., Zelli A., Capanna C. (2005). A new scale for measuring adults' prosocialness. *European Journal of Psychological Assessment*, 21, 77–89. doi:10.1027/1015-5759.21.2.77

Chen F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. doi:10.1080/10705510701301834

Christopher M., Neuser N., Michael P., Baitmangalkar A. (2012). Exploring the psychometric properties of the Five Facet Mindfulness Questionnaire. *Mindfulness*, 3, 124–131. doi:10.1007/s12671-011-0086-x

Ciesla J. A., Reilly L. C., Dickson K. S., Emanuel A. S., Updegraff J. A. (2012). Dispositional mindfulness moderates the effects of stress among adolescents: Rumination as a mediator. *Journal of Clinical Child and Adolescent Psychology*, 41, 760–770. doi:10.1080/15374416.2012.698724

Cohen S., Kamarck T., Mermelstein R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385–396. doi:10.2307/2136404

Condon P., Desbordes G., Miller W. B., DeSteno D. (2013). Meditation increases compassionate responses to suffering. *Psychological Science*, 24, 2125–2127. doi:10.1177/0956797613485603

Curtiss J., Klemanski D. H. (2014). Factor analysis of the Five Facet Mindfulness Questionnaire in a heterogeneous clinical sample. *Journal of Psychopathology and Behavioral Assessment*, 36, 683–694. doi:10.1007/s10862-014-9429-y

Davis M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.

Desrosiers A., Klemanski D. H., Nolen-Hoeksema S. (2013). Mapping mindfulness facets onto dimensions of anxiety and depression. *Behavior Therapy*, 44, 373–384. doi:10.1016/j.beth.2013.02.001

Diener E., Emmons R. A., Larsen R. J., Griffin S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71–75.

Dimitrov D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43, 121–149. doi:10.1177/0748175610373459

Galla B. M. (2016). Within-person changes in mindfulness and self-compassion predict enhanced emotional well-being in healthy, but stressed adolescents. *Journal of Adolescence*, 49, 204–217. doi:10.1016/j.adolescence.2016.03.016

Goldberg S. B., Wielgosz J., Dahl C., Schuyler B., MacCoon D. S., Rosenkranz M., …Davidson R. J. (2016). Does the Five Facet Mindfulness Questionnaire measure what we think it does? Construct validity evidence from an active controlled randomized clinical trial. *Psychological Assessment*, 28, 1009–1014. doi:10.1037/pas0000233

Goodman M. S., Madni L. A., Semple R. J. (2017). Measuring mindfulness in youth: Review of current assessments, challenges, and future directions. *Mindfulness*, 8, 1409–1420. doi:10.1007/s12671-017-0719-9

Greco L. A., Baer R. A., Smith G. T. (2011). Assessing mindfulness in children and adolescents: Development and validation of the Child and Adolescent Mindfulness Measure (CAMM). *Psychological Assessment*, 23, 606–614. doi:10.1037/a0022819

Gu J., Strauss C., Crane C., Barnhofer T., Karl A., Cavanagh K., Kuyken W. (2016). Examining the factor structure of the 39-item and 15-item versions of the Five Facet Mindfulness Questionnaire before and after mindfulness-based cognitive therapy for people with recurrent depression. *Psychological Assessment*, 28, 791–802. doi:10.1037/pas0000263

Hollis-Walker L., Colosimo K. (2011). Mindfulness, self-compassion, and happiness in non-meditators: A theoretical and empirical examination. *Personality and Individual Differences*, 50, 222–227. doi:10.1016/j.paid.2010.09.033

Hou J., Wong S. Y. S., Lo H. H. M., Mak W. W. S., Ma H. S. W. (2013). Validation of a Chinese version of the Five Facet Mindfulness Questionnaire in Hong Kong and development of a short form. *Assessment*, 21, 363–371. doi:10.1177/1073191113485121

Hu L., Bentler P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. doi:10.1080/10705519909540118

IBM Corp. (2016). *IBM SPSS statistics for Windows (version 24)*. Armonk, NY: Author.

Levinson D. B., Stoll E. L., Kindy S. D., Merry H. L., Davidson R. J. (2014). A mind you can count on: Validating breath counting as a behavioral measure of mindfulness. *Frontiers in Psychology*, 5, 1–10. doi:10.3389/fpsyg.2014.01202

Mackinnon A., Jorm A. F., Christensen H., Korten A. E., Jacomb P. A., Rodgers B. (1999). A short form of the Positive and Negative Affect Schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, 27, 405–416. doi:10.1016/S0191-8869(98)00251-7

Medvedev O. N., Norden P. A., Krägeloh C. U., Siegert R. J. (2018). Investigating unique contributions of dispositional mindfulness facets to depression, anxiety, and stress in general and student populations. *Mindfulness*, 9, 1757–1767. doi:10.1007/s12671-018-0917-0

Medvedev O. N., Siegert R. J., Kersten P., Krägeloh C. U. (2017). Improving the precision of the Five Facet Mindfulness Questionnaire using a Rasch approach. *Mindfulness*, 8, 995–1008. doi:10.1007/s12671-016-0676-8

Muthén L. K., Muthén B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Center for Education Statistics. (2018). *Common Core of Data (CCD): Public school data 2015-16, 2017-18 school years*. Retrieved from https://nces.ed.gov/ccd/schoolsearch/index.asp

Pallozzi R., Wertheim E., Paxton S., Ong B. (2017). Trait mindfulness measures for use with adolescents: A systematic review. *Mindfulness*, 8, 110–125. doi:10.1007/s12671-016-0567-z

Raes F., Pommier E., Neff K. D., Van Gucht D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*, 18, 250–255. doi:10.1002/cpp.702

Roeser R. W., Colaianne B. A., Greenberg M. T. (2018). Compassion and human development: Current approaches and future directions. *Research in Human Development*, 15, 238–251. doi:10.1080/15427609.2018.1495002

Roeser R. W., Eccles J. S. (2015). Mindfulness and compassion in human development: Introduction to the special section. *Developmental Psychology*, 51, 1–6. doi:10.1037/a0038453

Royuela-Colomer E., Calvete E. (2016). Mindfulness facets and depression in adolescents: Rumination as a mediator. *Mindfulness*, 7, 1092–1102. doi:10.1007/s12671-016-0547-3

Rudnev M., Lytkina E., Davidov E., Schmidt P., Zick A. (2018). Testing measurement invariance for a second-order factor. A cross-national test of the Alienation Scale. *Methods, Data, Analyses*, 12, 47–76. doi:10.12758/mda.2017.11

Ryan R. M., Deci E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78. doi:10.1037110003-066X.55.1.68

Satorra A., Bentler P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507–514. doi:10.1007/BF02296192

Sauer S., Walach H., Schmidt S., Hinterberger T., Lynch S., Büssing A., Kohls N. (2013). Assessment of mindfulness: Review on state of the art. *Mindfulness*, 4, 3–17. doi:10.1007/s12671-012-0122-5

Schafer J. L., Graham J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177. doi:10.1037//1082-989X.7.2.147

Schumacker R. E., Lomax R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.

Segal E. A., Cimino A. N., Gerdes K. E., Harmon J. K., Wagaman M. A. (2013). A confirmatory factor analysis of the interpersonal and social empathy index. *Journal of the Society for Social Work and Research*, 4, 131–153. doi:10.5243/jsswr.2013.9

Shapiro S. L., Carlson L. E., Astin J. A., Freedman B. (2006). Mechanisms of mindfulness. *Journal of Clinical Psychology*, 62, 373–386. doi:10.1002/jclp.20237

Skinner E., Pitzer J., Steele J. (2013). Coping as part of motivational resilience in school: A multidimensional measure of families, allocations, and profiles of academic

coping. *Journal of Educational and Psychological Measurement*, 73, 803–835. doi:10.1177/0013164413485241

Tran U. S., Glück T. M., Nader I. W. (2013). Investigating the Five Facet Mindfulness Questionnaire (FFMQ): Construction of a short form and evidence of a two-factor higher order structure of mindfulness. *Journal of Clinical Psychology*, 69, 951–965. doi:10.1002/jclp.21996

Van Dam N. T., van Vugt M. K., Vago D. R., Schmalzl L., Saron C. D., Olendzki A., …Meyer D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, 13, 36–61. doi:10.1177/1745691617709589

van Son J., Nyklíček I., Nefs G., Speight J., Pop V. J., Pouwer F. (2014). The association between mindfulness and emotional distress in adults with diabetes: Could mindfulness serve as a buffer? Results from Diabetes MILES: The Netherlands. *Journal of Behavioral Medicine*, 38, 251–260. doi:10.1007/s10865-014-9592-3

Veehof M. M., ten Klooster P. M., Taal E., Westerhof G. J., Bohlmeijer E. T. (2011). Psychometric properties of the Dutch Five Facet Mindfulness Questionnaire (FFMQ) in patients with fibromyalgia. *Clinical Rheumatology*, 30, 1045–1054. doi:10.1007/s10067-011-1690-9

Widaman K. F., Ferrer E., Conger R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, 4, 10–18. doi:10.1111/j.1750-8606.2009.00110.x

Williams M. J., Dalgleish T., Karl A., Kuyken W. (2014). Examining the factor structures of the Five Facet Mindfulness Questionnaire and the Self-Compassion Scale. *Psychological Assessment*, 26, 407–418. doi:10.1037/a0035566

Ziegler M., Kemper C. J., Kruyen P. (2014). Short scales–Five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35, 185–189. doi:10.1027/1614-0001/a000148

Table 1. Testing Alternate Factor Structures for the FFMQ

Model	χ^2	*df*	MLR scaling correction factor	CFI	RMSEA	SRMR	BIC
FFMQ 20 items							
One factor	1,544.10	170	1.161	.468	.116	.121	32,490.57
Five factor	411.14	160	1.157	.903	.051	.068	31,237.06
Five factor, hierarchical	429.28	165	1.160	.898	.052	.072	31,227.28
Four factor (no Observe)	249.37	98	1.176	.930	.051	.056	24,997.48
Four factor, hierarchical (no Observe)	252.80	100	1.177	.930	.051	.057	24,989.03
FFMQ 19 items (excluding Item 32)							
One factor	1,434.38	152	1.164	.484	.119	.120	30,884.61
Five factor	315.08	142	1.159	.930	.045	.059	29,644.20
Five factor, hierarchical	333.26	147	1.161	.925	.046	.064	29,633.85
Four factor (no Observe)	170.33	84	1.178	.959	.041	.045	23,405.93
Four factor, hierarchical (no Observe)	174.67	86	1.179	.958	.041	.047	23,398.41

Note. N = 599. FFMQ = Five Facet Mindfulness Questionnaire; χ^2 = chi-square; *df* = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; BIC = Bayesian information criterion. Gender (0 = *male*; 1 = *female*) and grade level (0 = *lower level [freshmen and sophomores] grades*; 1 = *higher level [juniors, seniors] grades*) were included as auxiliary variables (missing data correlates). Item 32 ("My natural tendency is to put my experiences into words") is from the Describe subscale.

Table 2. Tests of Measurement Invariance Across Time, Gender, and Grade Level.

	Model fit indices					Model fit comparisons					
Model	**χ^2**	***df***	**MLR scaling correction factor**	**CFI**	**RMSEA**	**Model Comparisons**	**$\Delta\chi^2$**	**Δdf**	***p***	**ΔCFI**	**ΔRMSEA**
Model 1	1,152.91	873	1.118	.968	.023						
Model 2	1,178.53	895	1.115	.968	.023	M2 versus M1	25.24	22	.286	.000	.000
Model 3	1,181.24	901	1.116	.968	.023	M3 versus M2	2.97	6	.812	.000	.000
Model 4	1,227.58	923	1.113	.965	.023	M4 versus M3	48.36	22	.001	−.003	.000
Model 5	1,243.98	929	1.112	.964	.024	M5 versus M4	17.59	6	.007	−.001	.001
Girls versus boys											
Model 1	290.77	172	1.140	.946	.048						
Model 2	296.20	183	1.142	.948	.045	M2 versus M1	5.67	11	.894	.002	−.003
Model 3	299.35	186	1.144	.948	.045	M3 versus M2	3.33	3	.343	.000	.000
Model 4	333.59	197	1.135	.937	.048	M4 versus M3	36.81	11	<.001	−.011	.003
Model 5	348.60	200	1.133	.932	.050	M5 versus M4	16.23	3	.001	−.005	.002

	Model fit indices					Model fit comparisons					
Model	**χ^2**	***df***	**MLR scaling correction factor**	**CFI**	**RMSEA**	**Model Comparisons**	**$\Delta\chi^2$**	**Δdf**	***p***	**ΔCFI**	**ΔRMSEA**
Lower level (freshmen, sophomores) versus upper level (juniors, seniors) students											
Model 1	265.73	172	1.143	.956	.043						
Model 2	279.57	183	1.144	.954	.042	M2 versus M1	13.89	11	.239	−.002	−.001
Model 3	280.71	186	1.148	.955	.041	M3 versus M2	1.68	3	.642	.001	−.001
Model 4	286.40	197	1.140	.958	.039	M4 versus M3	4.32	11	.960	.003	−.002
Model 5	292.33	200	1.139	.956	.039	M5 versus M4	6.06	3	.109	−.002	.000

Note. $N = 599$. χ^2 = chi-square; *df* = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; Δ = change in parameter; χ^2 and *p*-values for model fit comparisons are based on Satorra and Bentler's (2001) scaled χ^2 difference test for MLR estimation. Gender (0 = male; 1 = female) and grade level (0 = *lower level [freshmen and sophomores] grades*; 1 = *higher level [juniors, seniors] grades*) were included as auxiliary variables (missing data correlates). Model 1 (M1) = baseline model (without invariance); Model 2 (M2) = invariant first-order factor loadings; Model 3 (M3) = invariant first-order and second-order factor loadings; Model 4 (M4) = invariant first-order and second-order factor loadings and item intercepts; and Model 5 (M5) = invariant first-order and second-order factor loadings, item intercepts, and first-order factor intercepts.

Table 3. Descriptive Statistics for the FFMQ Across Three Assessment Waves.

Variable	*n*	*M*	*SD*	Range	α
Act Aware T1	532	3.06	.91	1.0–5.0	.88
Nonjudge T1	532	3.67	.88	1.0–5.0	.81
Nonreact T1	532	3.17	.74	1.0–5.0	.62
Describe T1	532	3.38	.83	1.0–5.0	.66
Full FFMQ T1	532	3.32	.57	1.6–5.0	.82
Act Aware T2	538	2.96	.88	1.0–5.0	.86
Nonjudge T2	536	3.57	.93	1.0–5.0	.85
Nonreact T2	537	3.12	.71	1.0–5.0	.61
Describe T2	537	3.38	.80	1.0–5.0	.67
Full FFMQ T2	539	3.25	.58	1.6–5.0	.83
Act Aware T3	469	2.95	.87	1.0–5.0	.85
Nonjudge T3	468	3.58	.93	1.0–5.0	.88
Nonreact T3	468	3.18	.72	1.0–5.0	.68
Describe T3	469	3.33	.78	1.0–5.0	.69
Full FFMQ T3	469	3.26	.59	1.1–4.9	.85

Note. Total $N = 599$. T1 = September assessment; T2 = January assessment; T3 = April assessment; FFMQ: Five Facet Mindfulness Questionnaire.

Table 4. Bivariate Correlations Between Mindfulness Subscales.

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Describe T1											
2. Acting with Awareness T1	.29										
3. Nonjudgment T1	.39	.34									
4. Nonreactivity T1	.16	.21	.21								
5. Describe T2	.54	.31	.36	.23							
6. Acting with Awareness T2	.17	.68	.32	.19	.37						
7. Nonjudgment T2	.23	.24	.62	.14	.35	.40					
8. Nonreactivity T2	.18	.25	.21	.55	.22	.26	.18				
9. Describe T3	.54	.30	.30	.18	.64	.29	.27	.20			
10. Acting with Awareness T3	.19	.65	.30	.23	.32	.66	.32	.21	.38		
11. Nonjudgment T3	.31	.27	.62	.15	.30	.33	.65	.14	.39	.43	
12. Nonreactivity T3	.19	.21	.19	.44	.21	.14	.17	.59	.32	.22	.24

Note. N = 599. T1 = September assessment; T2 = January assessment; T3 = April assessment. All correlations are significant at $p < .01$. Gender (0 = *male*; 1 = *female*) and grade level (0 = *lower level [freshmen and sophomores] grades*; 1 = *higher level [juniors, seniors] grades*) were included as auxiliary variables (missing data correlates).

Table 5. Incremental Predictive Validity Tests: Results of Multiple Regressions Predicting 3-Month Changes in Psychological Well-Being From Mindfulness Subscales.

	Life satisfaction T2				Perceived stress T2				Positive affect T2				Negative affect T2				Rumination T2			
Predictors	***B***	**LCI**	**UCI**	**β**	***B***	**LCI**	**UCI**	**β**	***B***	**LCI**	**UCI**	**β**	***B***	**LCI**	**UCI**	**β**	***B***	**LCI**	**UCI**	**β**
Outcome T1	.58***	.49	.66	.56	.45***	.36	.55	.46	.46***	.38	.55	.46	.59***	.49	.69	.58	.59***	.51	.67	.60
Describe T1	.11	−.02	.24	.07	−.03	−.10	.04	−.03	−.02	−.10	.06	−.02	.03	−.05	.11	.03	.02	−.05	.10	.02
Act Aware T1	.01	−.13	.15	.01	−.09**	−.15	−.03	−.12	.08	.00	.16	.09	−.09*	−.16	−.01	−.09	−.03	−.10	.04	−.03
Nonjudge T1	.07	−.06	.20	.04	−.07*	−.14	−.01	−.09	−.04	−.12	.05	−.04	−.10*	−.20	−.01	−.10	−.13**	−.21	−.06	−.13
Nonreact T1	.15	−.02	.33	.08	−.12**	−.19	−.04	−.13	.15**	.05	.25	.14	.00	−.09	.09	.00	−.07	−.15	.01	−.06
Gender	.02	−.17	.21	.01	.04	−.05	.13	.03	.04	−.07	.15	.03	.10	−.03	.21	.05	.11	−.01	.22	.06
Grade level	.14	−.05	.33	.05	−.08	−.17	.01	−.06	.07	−.04	.18	.04	−.06	−.18	.06	−.03	.00	−.11	.11	.00