Measuring Mindfulness in Black Americans: A Psychometric Validation of the Five Facet Mindfulness Questionnaire

Gold Nneka Okafor (gokafor@berkeley.edu)
University of California Berkeley https://orcid.org/0000-0001-9460-3598

Brett Q. Ford
University of Toronto

Stephen Antonoplis
University of California Berkeley

Anita Reina
University of Georgia

Samina Lutfeali
Stanford University

Amanda Shallcross
New York University

Research Article

Keywords: Mindfulness, Five Facet Mindfulness Questionnaire, construct validity, psychometrics, Black Americans.

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1325638/v1

Abstract

Objectives: Black Americans disproportionately experience higher levels of chronic stress. Mindfulness is a promising, cost-efficient treatment option for reducing stress and related mental health outcomes such as depression and anxiety. The Five Facet Mindfulness Questionnaire (FFMQ) is one of the most widely used tools to measure mindfulness; however, Black American samples have been underrepresented in validation studies of the FFMQ. Consequently, the validity of the FFMQ within Black Americans is unknown. The present study assessed the psychometric properties and nomological network of the original 39-item FFMQ (FFMQ-39) and the short form 15-item FFMQ (FFMQ-15) among a non-clinical, Black American sample in the United States.

Methods: In a longitudinal study, 586 Black Americans completed either the FFMQ-39 or the FFMQ-15 at two time points one month apart.

Results: Exploratory and confirmatory factor analyses supported a five-factor structure in both questionnaires. Both questionnaires had good fit indices (*RMSEA* > .05, *SRMR* > .05, *CFI* > .92, *TFI* > .92) and demonstrated strong test-retest reliability, expected associations with nomological network variables, and invariance across gender, mindfulness meditation experience, depression level, everyday discrimination, lifetime discrimination, household income, ethnic heritage, and skin tone.

Conclusion: The results indicate that both the FFMQ-39 and the FFMQ-15 can validly and reliably measure mindfulness in a non-clinical, Black American sample. These findings contribute to cultural generalizability and mindfulness assessment within underrepresented populations.

Key Words. Mindfulness, Five Facet Mindfulness Questionnaire, construct validity, psychometrics, Black Americans.

Introduction

In the past decade, there has been an exponential growth of research on mindfulness — a practice rooted in Buddhist principles that fosters awareness and acceptance of thoughts, emotions, and physical sensations to decrease suffering (Bodhi, 2011). Research indicates that mindfulness predicts lower anxiety, depression, (Webb et al., 2019), rumination (Blanke et al., 2020), and stress (Chiesa & Serretti, 2009); and is associated with greater psychological well-being (Baer et al., 2008), satisfaction with life (Rogge & Daks, 2020), and self-compassion (Svendsen et al., 2017). Despite these promising outcomes, research on mindfulness has focused on predominately White samples. This leaves open important questions about whether the effects of mindfulness are generalizable to individuals from other races and ethnicities, particularly to Black Americans who experience disproportionately higher levels of chronic and acute stress, and associated adverse mental health outcomes (Roberts et al., 2020). In a systematic review of 12,265 mindfulness studies from 1990 to 2016, only 24 studies (0.2%) focused on minority representation, cultural adaptations for interventions, or ethnoracial group comparisons (DeLuca et al., 2018). Moreover, only 11 of these (0.1%) included predominantly Black American samples. Similarly, a review of mindfulness-related therapies (i.e., mind-body therapies) for cardiometabolic diseases found only five out of 425 trials (1%) targeted Black populations (Johnson et al., 2018). In the present research, we begin to rectify this issue by completing a necessary first step in the study of mindfulness in Black Americans: validating a widely used measure of mindfulness for Black Americans.

The most widely used tool to measure mindfulness is the 39-item Five Facet Mindfulness

Questionnaire (FFMQ-39;Baer et al., 2006). The abbreviated version of the FFMQ-39, the 15-item FFMQ (FFMQ-15;Baer et al., 2012), has been used less frequently; however, it is a valuable option for assessing mindfulness in time-limited settings (e.g., large scale interventions; hospital settings). Both the FFMQ-39 and FFMQ-15 have been validated in predominantly White samples (e.g., Baer et al., 2008; Christopher et al., 2012; 2012; Gu et al., 2016). The only study to validate the 39-item FFMQ within a Black American sample used a low income sample of Black Americans with recent suicidal ideation who were recruited from clinical settings (Watson-Singleton et al., 2018). To date, neither of these FFMQ questionnaires have been validated with a non-clinical, Black American sample. Given that mindfulness is a skill that can be useful regardless of underlying psychopathology, it is necessary to validate all available versions of mindfulness measures (e.g., long and short) within normative community samples to optimize generalizability.

Five Facet Mindfulness Questionnaire

Mindfulness is a complex and multifaceted construct that has sparked continued debate about its various components. Within this discourse, five mindfulness factors (or "facets") have received the most empirical attention — *acting with awareness (awareness)*, *describing (describe)*, *non-judging of experience (nonjudgment)*, *nonreactivity to inner experience* (*nonreactivity*), and *observing* (*observe)* (Baer et al., 2006; Baer et al., 2008; Christopher et al., 2012; Gu et al., 2016; Sweeney et al., 2021*)*. The focus on these five factors is largely due to the development and subsequent widespread implementation of the FFMQ which assesses the five factors. *Awareness* involves maintaining a focus on actions, without distraction. *Describe* involves describing or labeling in words internal experiences such as beliefs, opinions, emotions, and expectations. *Nonjudgement* involves being open-minded and curious about one's internal

experience. *Nonreactivity* involves experiencing emotions from a meta-cognitive or decentered perspective. *Observe* involves attending to internal experiences and cognitions; and external stimuli, such as smells, sounds, and sights. Prior evidence from confirmatory factor analyses indicate that the five-factor model of the FFMQ-39 and FFMQ-15 fit the data well in largely White samples (*CFI* > .90, *TLI* > .90, *RMSEA* > .06, and *SRMR* < .09; Baer et al., 2006; Baer et al., 2008; Christopher et al., 2012; Gu et al., 2016; Sweeney et al., 2021). The factors in both questionnaires have displayed good internal consistency. The FFMQ-39 alphas range from .75 to .91 (Baer et al., 2006) and the FFMQ-15 alphas range from .64 to .80 (Gu et al., 2016). The test-retest stability of the English FFMQ-39 and FFMQ-15 has only been reported in one study, which found weak stability over a 6-week period (i.e., correlations for each factor ranging from .22 to .54; (Watson-Singleton et al., 2018)). However, this study also included an intervention during that 6-week period which may have interfered with assessing the stability of the measure; thus, the temporal stability of both the FFMQ-39 and FFMQ-15 has been largely unexamined.

In addition to these analyses of factor structure and reliability, the previously established nomological network of the FFMQ-39 and the FFMQ-15 demonstrate the factors correlate with conceptually related constructs, with the exception of the *observe* factor. For instance, each factor has correlated with lower anxiety, depression (Webb et al., 2019), rumination (Svendsen et al., 2017), and greater levels of life satisfaction, psychological well-being (Christopher et al., 2012), and mindfulness experience (Baer et al., 2008). The *observe* factor has had less success, however, showing inconsistent associations with outcomes such as rumination and psychological well-being (Baer, et al., 2008; Gu et al., 2016; Sweeney et al., 2021). Finally, previous invariance analyses on the FFMQ analyses have been limited. To our knowledge, no previous FFMQ invariance analyses have included predominantly Black samples.

Overall, the 39-item and 15-item FFMQs have demonstrated good internal consistency, validity, and invariance at least in predominantly White samples. However, the questionnaires' psychometric performance in Black Americans is unknown. Given that prior work has empirically questioned the validity of the five-factor interpretation of the FFMQ across non-White cultures (Karl et al., 2020), determining its validity for Black Americans remains an open question.

Considering why the FFMQ may perform differently for Black samples

There are reasons to question the performance of the FFMQ within Black Americans. To begin, the language of the FFMQ items is not inclusive of all Black Americans. For example, an item in the *observe* factor states, "I pay attention to sensations, such as the wind in my hair or sun on my face." Several Black American hair styles including afros, braids, and locs do not blow in the wind, which suggests that the inclusion of Black Americans may not have been prioritized when the FFMQ items were developed. Since this item references an unrelatable experience for many Black Americans, this item may lead to a variance within Black American responses. It is also possible that Black Americans may feel alienated by this phrasing which could be an unintended negative mood induction, thus influencing responses to other items in the questionnaire.

In addition to the ethnocentric phrasing in the FFMQ, there is reason to believe that the factors of the FFMQ may not adequately capture the most important factors of mindfulness within Black Americans. For instance, the *awareness* and *observe* factors measure attention skills. An example *awareness* item is "I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted." Another example *observe* item is "I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing." Previous research has

found that due to exposure to stress and trauma, Black Americans have compromised auditory attention skills compared to White Americans (Beydoun et al., 2018). Thus, it is possible that the *awareness* and *observe* facets might not emerge as prominent facets in an exploratory factor analysis among Black American participants; potentially leading to an alternate FFMQ model structure for Black Americans.

Heterogeneity within Black Americans

With the limited amount of mindfulness research within Black Americans, there is no prior understanding of how mindfulness interacts with the significant heterogeneity *within* Black American samples. The analysis of heterogeneity within Black Americans speaks to the within-group generalizability of research that is often overlooked, especially in studies where Black people are marginally represented (Buchanan et al., 2021). Four important social experiences that may shape conceptualizations of mindfulness within a Black American sample are ethnic heritage, perceived discrimination, household income, and skin tone (Earl et al., 2011; Oh et al., 2021). For instance, ethnic heritage is a key characteristic within Black Americans and includes African Americans who have descended from slavery, immigrants from the Caribbeans, and immigrants from Africa. These communities vary widely on perceptions of mental health (e.g., awareness and stigma) (Adewale et al., 2016) and also have different historic origins of trauma and lived experiences (e.g., slavery, colonialism, and immigration). Thus, it is possible that mindfulness (a construct related to mental health) may also be understood differently across these groups. Therefore, the FFMQ questionnaires may lack measurement invariance across Black American heritage groups.

In addition to heritage, demographics related to perceived stress have the potential to cause an invariance of the FFMQ models. Higher levels of perceived stress have been associated

with lower levels of the FFMQ factors (Carmody & Baer, 2008). It is possible that people who experience more stress are less efficient in describing, observing, or paying attention to their emotional experiences — core aspects of the *describe*, *observe*, and *awareness* factors in the FFMQ. Experiencing discrimination, having a household income level below the poverty line, and having darker skin complexion have been independently associated with higher stress levels (Williams et al., 1997; Oh et al., 2021). Thus, it is also important to question the invariance of the FFMQ within perceived discrimination, household income, and skin tone.

Current Investigation

The psychometric properties of the FFMQ-39 and FFMQ-15 have been well-established in studies that include nearly exclusively White samples. Given that mindfulness predicts important psychological health outcomes, it is important to understand whether this commonly used measure of mindfulness is valid in Black Americans, who experience disproportionately higher stress-related mental and physical health outcomes (Roberts et al., 2020). The current research is the first study to analyze the psychometric properties of the FFMQ-39 and the FFMQ-15 within a non-clinical, Black sample in the United States. We also describe the generalizability of our results by reporting extensive demographics (e.g., gender, mindfulness meditation experience, depression level, everyday discrimination, lifetime discrimination, household income, ethnic heritage, and skin tone) and analyzing the measurement invariance within each of these groups.

In our three-part study, participants completed either the FFMQ-39 or the FFMQ-15 at two time points, separated by one month, and were followed up with at a third time point about 2.5 years later to collect additional socio-demographic information. Pre-registered psychometric analyses included: (1) exploratory factor analyses (EFAs) to identify the factor structure of each

questionnaire; (2) confirmatory factor analyses (CFAs) to confirm strength and reliability of the factors; (3) Cronbach's alphas to confirm the internal reliability; (4) test-retest analyses to assess temporal reliability; (5) nomological network examination to describe the relationship between the observed mindfulness construct and its correlation with other domains including anxiety, depression level, rumination, psychological well-being, satisfaction with life, self-compassion, and mindfulness meditation experience; (6) invariance analyses to test the generalizability of the observed FFMQ models across various identities and experiences of the Black community including gender, mindfulness meditation experience, depression level, everyday discrimination, lifetime discrimination, household income, ethnic heritage, and skin tone. As the first study to validate the FFMQ-39 and the FFMQ-15 with a non-clinical, Black sample, we do not have predictions for each of the psychometric analyses. Rather, our goal is to explore the psychometric properties of the FFMQ to determine whether the FFMQ can be used to accurately measure levels of mindfulness in Black Americans.

Methods

Participants

A pool of 17,465 U.S. participants recruited from Amazon Mechanical Turk Prime completed a brief screening survey where they reported their racial identity, along with other demographics. Only participants who self-identified as Black (n = 1,963) were invited to participate in the present study and of those, 619 participants enrolled in the Time 1 assessment; 541 participants began the Time 2 assessment (which took place one month later); and 263 began the Time 3 assessment (which took place approximately two and a half years later). All participants lived in the United States and will be referred to as "Black Americans." Participants who failed attention checks were not invited to participate at future time points. Several

participants were removed from analyses based on the following pre-registered exclusion criteria: not identifying as Black at Time 1 (*n* = 14), not correctly answering the attention checks at Time 1 (*n* = 20), Time 2 (*n* = 12, excluded only from Time 2 analyses), or at Time 3 (*n* = 8, excluded only from Time 3 analyses). Some of these exclusion criteria overlapped within participants, resulting in a final Time 1 sample size of 586, a Time 2 sample size of 520 (89% retention), and a Time 3 sample size of 251 (43% retention). Table 1 shows the socio-demographic information for the participants.

Table 1

Socio-Demographic percentages of participants

Socio-Demographic	Percentages
Gender	
Female	73.7%
Male	23.9%
Gender variant / non-binary	0.2%
Declined to answer	2.2%
Age	
18 - 30	39.2%
31 - 40	33.7%
41 - 50	14.8%
51 - 60	6.4%
61 +	4.2%
Declined to answer	1.7%
Mindfulness Meditation Experience	
No experience	46%
Any level of experience	52.1%
Declined to answer	1.9%
Household Income	
$19,999 or below	12.5%
$20,000 - $39,999	28.8%
$40,000 - $59,999	25.1%
$60,000 - $79,999	15.2%

$80,000 - $99,999	7.1%
$100,000 or more	10%
Declined to answer	1.3%
Ethnic Heritage*	
Black American Native (African American descendants of the enslaved)	85.4%
African immigrant or Descendant of African immigrant	7.1%
Afro-Caribbean	7.5%
Skin Tone*	
Light complexion	22.9%
Medium complexion	47.4%
Dark complexion	29.7%

Note. The asterisked variables were collected at Time 3 (which had a final sample size of N = 251). All other variables were collected in the screener survey or Time 1 (which had a final sample size of n = 586).

Procedures

This multi-part longitudinal study was distributed online via Amazon's Mechanical Turk. A brief screener survey recruited U.S.-based participants to determine if the participants identified as Black or African American (e.g., African American, Jamaican, Nigerian, Ethiopian). The screener also collected the age and gender of the participants. The race inclusion criterion was not advertised in the screener survey. At Time 1, half of the participants were randomly assigned to complete the 39-item FFMQ, while the other half were randomly assigned to complete the 15-item FFMQ. Time 2 was scheduled to take place approximately one month after Time 1 (average = 32.6 days, range = 23.7 - 66.9 days). At Time 2, all participants completed the same version of the FFMQ they completed at Time 1. The FFMQ was assessed before any other measures at both Time 1 and Time 2 (except for several demographic questions assessed first in Time 1). Anxiety, depression level, rumination, psychological well-being, satisfaction with life, self-compassion, and mindfulness meditation experience were collected at Time 1. Lifetime discrimination and everyday discrimination were collected at Time 2.

Approximately two and a half years later (average = 31.4 months, range = 30.3-32.1 months), we invited participants to complete Time 3, which was conducted to collect additional demographic information not assessed at Time 1, including ethnic heritage and skin tone. Collectively, participants had the opportunity to earn a total of $13.05 for completing the study ($0.05 for the screener, $3 for Time 1, $7 for Time 2, and $3 for Time 3; or ~$12USD/hour).

Measures

Table 2 includes the mean, standard deviation, alpha, omega total, and omega hierarchical of all of the measures in the present study at Time 1 (excluding everyday discrimination and lifetime discrimination which were collected at time 2). The FFMQ questionnaires collected at Time 2 do not appear here because they were solely used for pre-registered test-retest analyses.

Table 2

Measure Descriptives

Measure	Present Study (*N* = 586)						
	Mean	*SD*	Min	Max	Alpha	$Omega_t$	$Omega_h$
FFMQ-39							
Awareness	28.55	6.34	9	40	.90	.93	.69
Describe	28.70	6.13	9	40	.90	.92	.75
Nonjudgment	25.64	7.18	9	40	.90	.92	.84
Nonreactivity	21.92	4.94	8	34	.83	.88	.76
Observe	27.75	5.39	9	40	.81	.86	.66
FFMQ-15							
Awareness	10.65	2.49	3	15	.71	.78	--
Describe	10.84	2.61	3	15	.81	.82	--
Nonjudgment	10.57	2.74	3	15	.81	.82	--
Nonreactivity	10.20	2.50	3	15	.76	.77	--

Measure	Present Study (N = 586)						
	Mean	*SD*	Min	Max	Alpha	$Omega_t$	$Omega_h$
Observe	11.08	2.42	3	15	.65	.67	--
Anxiety	5.28	3.49	0	15	.82	.88	.54
Depression	13.72	10.99	0	57	.93	.94	.83
Rumination	3.40	0.93	1	5	.94	.95	.82
Psychological Well-being	78.24	13.81	33	108	.85	.88	.61
Satisfaction with Life	18.26	7.82	5	35	.90	.92	.85
Self-Compassion	36.63	9.73	12	60	.88	.91	.68
Mindfulness Meditation Experience	1.94	1.18	1	5	--	--	--
Everyday Discrimination*	1.50	1.06	0	5	.93	.96	.85
Lifetime Discrimination*	1.94	1.83	0	9	.64	.72	.57

Note. Asterisked measures represent the averaged data prior to binary transformation. $Omega_t$ represents the omega total. $Omega_h$ represents the omega hierarchical. See Table 1 for descriptive statistics for gender, skin tone, heritage and household income.

Five Facet Mindfulness Questionnaire

Five Facet Mindfulness Questionnaire - Long form (FFMQ-39). Half of the participants completed the long-form, original 39-item FFMQ (Baer et al., 2006). This questionnaire assesses the five facets of mindfulness: acting with awareness *(awareness)*, describing *(describe)*, non-judging of experience *(nonjudgment)*, nonreactivity to inner experience (*nonreactivity*), and observing (*observe)*. The items were measured on a 5-point Likert-type scale from 1 (*never or rarely true*) to 5 (*very often or always true*). Example items include: “I do jobs or tasks automatically without being aware of what I’m doing” (reverse-keyed; *awareness*), “I’m good at finding words to describe my feelings" (*describe*), “I tell myself I shouldn’t be feeling the way I’m feeling” (reverse-keyed; *nonjudgment*), “When I have distressing thoughts or images, I just notice them and let them go” (*nonreactivity*) and “I pay attention to sensations, such as the wind in my hair or sun on my face” (*observe*). Each factor was summed for a total factor score. The factor scores ranged from 8–40 for all factors except for

the nonreactivity factor which ranges from 7–35.

Five Facet Mindfulness Questionnaire-15 (FFMQ-15). The other half of the participants completed the 15-item short-form version of the Five Facet Mindfulness Questionnaire (FFMQ-15) (Baer et al., 2012). Each factor contained three items. The items were chosen to balance content validity and item factor loadings. The factor scores ranged from 3-15.

Nomological Network Measures

The nomological network examines the relationship between the observed FFMQ questionnaires and other constructs that are known to be correlated with mindfulness. This analysis converges with the mindfulness nomological network found in previous research in order to provide construct validity for the FFMQ-39 and FFMQ-15. The descriptives for the nomological network are in Table 2. The scores were averaged or summed based on the conventional use of the measures in previous research.

Anxiety. Anxiety was measured with the seven-item anxiety subscale of the Hospital Anxiety and Depression Scale (HADS-A) (Zigmond & Snaith, 1983). This subscale identifies anxiety severity in a non-psychiatric population. The items were measured on a 4-point Likert-type from 0 (*not at all*) to 3 (*very often*). The responses were averaged to create a composite score. Higher scores indicate a higher level of anxiety.

Depression level. The Beck Depression Inventory (BDI-II) (Beck, Steer, & Brown, 1996) measures the severity of depression symptoms experienced during the past two weeks. This inventory contains 21-items rated on a scale of 0 (e.g. *I do not feel sad*) to 3 (e.g. *I am so sad or unhappy I cannot stand it*) and was summed to create a composite score. Due to Institutional Review Board concerns, the current study removed one item referencing suicide ideation.

Rumination. Rumination was measured with the Rumination-Reflection Questionnaire (RRQ) - rumination subscale (Trapnell & Campbell, 1999). The RRQ - rumination subscale consists of 12-items measured on a 5-point Likert-type scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The scores were averaged to form a composite score.

Psychological well-being. Psychological well-being was measured with the abbreviated version of the Ryff Psychological Well-being Scale (Ryff et al., 1995). The scale consists of 18-items with a six-point Likert-scale from 1 (*strongly disagree*) to 5 (*strongly agree)*. The scores were summed to form a composite score.

Satisfaction with life. The Satisfaction with Life (SWL) measure (Diener et al., 1985) was a 5-item scale that assesses positive evaluations of one's life. Items were measured on a 7-point, Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). The scores were summed to form a composite score.

Self-compassion. The Self-Compassion Scale-short form (SCS-SF) (Raes et al., 2011), was used to measure self-compassion. Originally derived from the Neff (2003) Self-Compassion Scale (SCS), the SCS-SF is a 12-item scale measuring global self-compassion on a 5-point Likert-type scale from 1 (*almost never*) to 5 (*almost always*). The items were summed to create a composite score.

Mindfulness meditation experience. Mindfulness meditation experience was assessed with one item that asked about the frequency of mindfulness meditation on a one to five response scale: 1 (*I have no experience*), 2 (*I tried it once or twice*), 3 (*I practice mindfulness meditation several times per year*), 4 (*I practice mindfulness meditation several times per month*), and 5 (*I practice mindfulness meditation once a week or more*).

Invariance Measures

To test the consistency of the FFMQ-39 and the FFMQ-15 within groups in our sample, measurement invariance was assessed across the following variables: gender, mindfulness meditation experience, depression level, everyday discrimination, lifetime discrimination, household income, ethnic heritage, and skin tone. Although several of these measures are continuous, measurement invariance analyses require continuous measures to be segmented into meaningful categories. Below, we describe how we determined these categories.

Mindfulness meditation experience. The mindfulness meditation experience was transformed into a binary variable for the measurement invariance analyses. Participants who indicated having no experience with mindfulness were scored as 0 (n = 273) and participants who indicated any level of mindfulness meditation experience were scored as 1 (n = 309).

Depression level. BDI-II scores were transformed into a binary variable based off previous research where 0 (n = 352) represents minimal depression (0 -13 BDI-II score) and 1 (n = 233) represents mild to severe depression (14 - 60 BDI-II score) (Beck et al., 1996).

Everyday Discrimination. Everyday discrimination was measured with a modified version of the Williams et al. (1997) Everyday Discrimination Scale. This nine-item questionnaire measures perceived unfair treatment in everyday life, modified in the present study to include experiences due to a participant's race. The responses were scored on a six-point Likert scale that ranged from 0 (*never*) to 5 (*Almost everyday)*. Following previously established procedures (Lewis et al., 2010), the responses were averaged into a composite. The composite was transformed into a binary variable where 0 (n = 201) indicated a composite score of less than one – representing a participant who reported experiencing discrimination *less* than once on a daily basis, and 1 (n = 318) indicated a composite score of one or *more* – representing a

participant who reported experiencing discrimination more than once on a daily basis.

Lifetime Discrimination. The Lifetime Discrimination Scale (Kessler et al., 1999) measures the perceived experience of discrimination across nine domains such as education, employment, and housing over the course of one's life. Items are rated 1 (*yes*) or 0 (*no*). The responses were summed into a single composite. The present study adapted the items of this scale to measure lifetime experiences of racial discrimination, in particular. Participants who did not report any lifetime discrimination were scored as 0 (n = 141) and participants who reported experiencing any amount of lifetime discrimination were scored as 1 (n = 379).

Household Income. Participants indicated their household income on scale ranging from $19,000 and below to $200,000 and above. The scale increments were separated by $20,000. We examined invariance between participants below versus above the United States average household poverty line. Thus, household income was dichotomized such that 1 (n = 74) indicated participants whose household income was $19,000 or below and 0 ($n$ = 511) indicated participants whose household income was $20,000 and above.

Ethnic Heritage. Participants' ethnic heritage was derived from their parents' ethnic heritage collected at Time 1. Participant parents' ethnic heritage was indicated by selecting if one or more parents identified as an "African American", "African Caribbean / Afro-Caribbean", "Hispanic or Latino, including Mexican American, Central American", "African", "Mixed", or "Other". For invariance analysis, we were primarily interested in contrasting participants who descended from American slavery vs. not, which we indirectly assessed via parents' ethnic heritage. Specifically, we dichotomized the parents' ethnic heritage such that if one or more parent was identified as African American, the ethnic heritage was scored as 1 (i.e., likely descendant of American slavery; n = 489); and if both parents were not identified as African

American, the ethnic heritage was scored as 0 (i.e., likely not descendant of American slavery; *n* = 104). We derived ethnic heritage from the parents' ethnic heritage because the sample size of the more direct measure of participants' ethnic heritage collected at Time 3 was too small for an invariance analysis. Specifically, at TIme 3, 212 participants self-identified as a Black American Native/African American descendant of the enslaved, and 36 participants self-identified as an African immigrant, descendant of African immigration, and/or African Caribbean/Afro Caribbean. Since the latter subgroup sample size was less than 50, we were not able to run an invariance analysis on this variable (Tanaka, 1987). Confirming that parents'ethnic heritage is indeed a valid marker of participants' ethnic heritage, we found that the parents'ethnic heritage significantly predicted the participants' ethnic heritage such that if one or more parent was identified as "African American" as opposed to "African Caribbean / Afro-Caribbean", "Hispanic or Latino, including Mexican American, Central American", "African", "Mixed", or "Other", the odds of a participant identifying as a descendant of the American enslaved increased by 16.55 times ($p < .001$). This simple model predicted participants' ethnic heritage with an accuracy of 88% in cross-validation analyses.

Skin tone. The complexion of the participants' skin tone was collected following the method of Bond and Cash (1992). Participants were presented with nine colored squares. Each color represented a different skin tone, including three light, three medium, and three dark skin tones. Participants were asked to select the square with the color that most resembled the complexion on their face. The responses were transformed into a binary variable where 0 (*n* = 57) indicated light complexions and 1 (*n* = 192) indicated medium and dark complexions.

Analytic Plan

The present study followed a pre-registered analytic plan to assess the validity of the

FFMQ-39 and the FFMQ-15. All pre-registered tests are reported. Any deviations from the pre-registration are explicitly noted. All analyses were conducted using R, version 4.1.0 (R Core Team, 2020), and the *psych* (Revelle, 2021) and *lavaan* (Rosseel, 2012) packages. The de-identified data will be available to the public via osf.io. The following FFMQ analyses were conducted on the data collected at Time 1. The FFMQ data from Time 2 were only used for the test-retest analyses.

Results

Factor Structure

Suitability

We conducted two tests to confirm the suitability for factor analysis of the FFMQ-39 and the FFMQ-15. First, we used Barlett's test of sphericity to test whether item responses were orthogonal and, thus, not able to be compressed. Orthogonality was rejected for both questionnaires, p's $< .001$, indicating that item responses correlated significantly differently from the scenario where they did not correlate at all. Thus, both questionnaires have the potential to produce meaningful factors. Second, we used the Kaiser-Meyer-Olkin (KMO) test to measure the average proportion of variance between responses to two items that was general or related to other item responses. KMO values greater than .60 are considered adequate (Tabachnick et al., 2019). The KMO criterion for the FFMQ-39 and FFMQ-15 were .91 and .77, indicating suitability for factor analysis.

Factor Extraction

We used eigenvalues greater than one, scree plot tests, and parallel factor analyses to determine the number of factors to extract from each questionnaire. For the 39-item FFMQ, we observed seven eigenvalues greater than one; however, the scree plot test and the parallel factor

analysis indicated five factors. The 15-item FFMQ had four eigenvalues greater than one, while the scree plot and the parallel factor analyses indicated five factors. See supplementary material to view scree plots. Floyd and Widaman (1995) argue that a scree plot test is more reliable than selecting factors based on eigenvalues greater than one because eigenvalues are dependent on the number of items. The parallel factor analyses were included for additional support for the factor number selection and were not pre-registered. In the current analyses, scree plots and the parallel factor analyses supported five factors for each version of the FFMQ.

Exploratory Factor Analysis

Following Baer et al. (2006), factor loadings were examined using an EFA with principal axis factoring, promax rotation, and five extracted factors. Items loaded onto their respective factors as expected. In the original construction of the FFMQ, Baer et al. (2006) removed items with factor loadings less than .40 and/or items with cross-loading differences of less than .20 from the highest factor loading in the item. When this method was applied to the measures of the current study, the 39-item FFMQ was reduced to 36 items and the 15-item FFMQ was unaffected. Given that the primary goal of this study is to perform psychometric tests on the FFMQ-39 and the FFMQ-15 versions, the fit indices, internal consistencies, test-retest reliabilities, and nomological network of the 36-item FFMQ are in the supplemental material. Also presented in supplementary material are pre-registered analyses that followed criteria that differ from Baer et al. (2006), and that are outlined by Watson-Singleton et al. (2018) – the only other investigation that has examined the FFMQ in a Black sample.

The five factors observed in the FFMQ-39 and the FFMQ-15 measures aligned with the five factors identified in previous research: *awareness*, *describe*, *nonjudgment*, *nonreactivity*, and *observe*. In the FFMQ-39, the five factors accounted for 51% of the total variance. Of this

percentage, *awareness* accounted for 12%, *describe* accounted for 11%, *nonjudgment* accounted for 12%, *nonreactivity* accounted for 8%, and *observe* accounted for 8%. In the FFMQ-15, the five factors accounted for 56% of the variance. *Awareness* explained 11%, *describe* explained 12%, *nonjudgment* explained 13%, *nonreactivity* explained 11%, and *observe* explained 9%.

Table 3 reports how the five factors correlate with each other, for both the FFMQ-39 and the FFMQ-15. Although most factors were positively related, they remained distinct, ranging from -.17 to .54. Correlations between factors were similar across the FFMQ-39 and FFMQ-15, although generally smaller in magnitude for the latter, likely due to the lower number of items. Importantly, these factor correlations were similar to those reported by Baer et al., (2006). For the FFMQ-39, two (of ten) correlations differed: (1) the previously demonstrated null association between *observe* and *nonjudgment* was significantly negative in the current study (r = -.17, p = .003); and (2) the previously demonstrated positive correlation between *awareness* and *observe* was not significant (r = .10, p = .07). For the FFMQ-15, three (of ten) correlations differed such that previously-demonstrated positive correlations were not significant in the present study: (1) *awareness* and *nonreactivity* (r = .10, p = .09), (2) *awareness* and *observe* (r = .02, p = .80), and (3) *nonjudgment* and *nonreactivity* (r = .04, p = .51).

Table 3

Intercorrelations and Test-Retest Reliabilities of the FFMQ-39 and the FFMQ-15

	Awareness	Describe	Nonjudgment	Nonreactivity	Observe
FFMQ-39					
Awareness	.85*	--	--	--	--
Describe	.48*	.82*	--	--	--
Nonjudgment	.54*	.30*	.84*	--	--
Nonreactivity	**.29***	**.41***	**.24***	.80*	--
Observe	.10	.30*	**-.17***	**.47***	.79*

FFMQ-15					
Awareness	.70*	--	--	--	--
Describe	.39*	.71*	--	--	--
Nonjudgment	.52*	.28*	.71*	--	--
Nonreactivity	**.10**	**.24***	**.04**	.52*	--
Observe	.02	.25*	**.02**	**.23***	.70*

Note. * $p < .01$. Bold intercorrelations significantly differ between the two different versions of the FFMQ at the 95% significance level. Test-retest reliabilities over one month are given on the diagonal. Reliabilities are not bolded based on whether they significantly differed between FFMQ versions.

To examine whether the strength of the associations between the five factors differed in the FFMQ-39 compared to the FFMQ-15, we conducted a William's correlation test, with a 95% significance level. These analyses indicated that the correlation between the *nonjudgment* and *observe* factors was significantly different between the FFMQ-39 and the FFMQ-15. Additionally, all intercorrelations between the *nonreactivity* factor and the other factors in the FFMQ-39 were significantly stronger compared to the parallel intercorrelations in the FFMQ-15. This suggests that the internal relationships between factors within the FFMQ-39 differ somewhat from the internal relationships between the factors within the FFMQ-15 such that some intercorrelations (e.g., the *nonjudgment* and *observe* correlation) were stronger in the longer (vs. shorter) version of the FFMQ. We ran a second William's correlation test on the latent variable correlations of the FFMQ-39 and FFMQ-15 to correct for unreliability and the results were consistent with the original William's correlation test.

Confirmatory Factor Analysis

A confirmatory factor analysis (CFA) was conducted on the five-factor structure of the FFMQ-39 and the FFMQ-15. Following procedures from prior FFMQ validation studies, both CFAs were applied using structural equation modeling with a robust maximum likelihood

estimator. Full information maximum likelihood was applied to estimate the missing values. There are mixed approaches for modeling the FFMQ scales as either a five-factor model or a five-factor higher-order model (e.g., Baer et al. 2004; Watson-Singleton et al., 2018). We compared both models and found that the five-factor model outperformed the five-factor higher-order model for both the FFMQ-39 and the FFMQ-15 (see supplemental material). Thus, we decided to move forward with the five-factor model. The analyses were run at the item level (as opposed to parcels) in order to observe the individual item properties and relationships to their factors (Christopher et al., 2012). The item-level factor loadings for the FFMQ-39 and FFMQ-15 are in the supplemental material. For the FFMQ-39, the average factor loading was .74 for *awareness*, .74 for *describe*, .75 for *nonjudgment*, .65 for *nonreactivity*, and .59 for *observe*. For the FFMQ-15, the average factor loading was .73 for *awareness*, .78 for *describe*, .78 for *nonjudgment*, .73 for *nonreactivity*, and .64 for *observe*.

Table 4 details the fit statistics for the FFMQ-39 and FFMQ-15. Overall, the five-factor model fits well for both the FFMQ-39 and FFMQ-15. Examining the chi-square test and the chi-square over degrees of freedom ratio (χ^2/df), the chi-square test was significant. The chi-square ratios were small (< 3; Schermelleh-Engel et al., 2003), indicating that the model fit the data well for the number of parameters in the model. Next, we examined the absolute fit of the data. Both the root mean squared error of approximation (RMSEA) and the standardized root mean square residual (SRMR) indicated fair or acceptable fit of the model (between .05 and .08 for RMSEA, Browne & Cudeck, 1993; <.10 for SRMR, Hu & Bentler, 1999). Finally, we examined incremental fit indices. For the FFMQ-15, both the comparative fit index (CFI) and the Tucker-Lewis index (TLI) indicated good fit (both >.90; Hu & Bentler, 1999). For the FFMQ-39, however, the CFI (.86) and the TLI (.85) both indicated inadequate fit. An exploratory inspection

of modification indices indicated that the largest contributors to misfit were four pairs of similarly worded items from the *awareness* factor (e.g., "I am easily distracted" and "When I do things, my mind wanders off and I'm easily distracted"). After allowing residuals to correlate for these four pairs as well as accounting for acquiescence (Aichholzer, 2014), the CFI and TLI both increased to .92.

Table 4

Fit Indices of FFMQ-39 and FFMQ-15

Questionnaire	*n*	χ^2	*df*	$\chi 2/df$	RMSEA	SRMR	CFI	TLI	BIC
FFMQ-39	289	1379.8	692	1.99	0.06	0.08	0.86	0.85	28664.49
FFMQ-39[a]	289	1081.97	687	1.57	0.05	0.07	0.92	0.92	28351.9
FFMQ-15	297	153.02	80	0.92	0.06	0.05	0.94	0.92	11904.82

Note. FFMQ-39[a] = model after allowing for the residuals of similarly worded items to correlate and acquiescence. χ^2 = chi-square, *df* = degrees of freedom, χ^2/df = chi-square over degrees of freedom ratio, RMSEA = root mean squared error of approximation, SRMR = standardized root mean square residual, CFI = comparative fit index, TLI = Tucker-Lewis index, BIC = Bayesian information criterion.

Reliability

Internal Consistency

The internal consistency of the FFMQ-39 and the FFMQ-15 were assessed with Cronbach's coefficient alpha, McDonald's omega total ($omega_t$), and McDonald's omega hierarchical ($omega_h$). The inclusion of the omega analyses was not pre-registered. As shown in Table 2, all internal reliability indices were acceptable for the FFMQ-39 (all alphas > .81, all $omega_t$'s > .86 , $omega_h$'s > .66) and the FFMQ-15 (all alphas > .65, all $omega_t$'s > .67). The FFMQ-15 does not have an omega hierarchical because the omega hierarchical cannot be calculated for factors of three or fewer items.

Test-Retest Reliability

In order to assess the temporal stability of the FFMQ-39 and the FFMQ-15, the five factors of each measure obtained at Time 1 were correlated with the parallel factors measured one month later at Time 2. As shown in Table 3, all test-retest correlations were strong and significant at the 99% significance level. The supplemental material contains a correlation table of all FFMQ-39 and FFMQ-15 factors across time.

Nomological Network: Predicting Psychological Experience

Table 5 demonstrates the nomological network between the FFMQ factors and measures of anxiety, depression level, rumination, psychological well-being, satisfaction with life, self-compassion, and mindfulness meditation experience. To the extent that the FFMQ-39 and FFMQ-15 accurately measures mindfulness in Black Americans, Black Americans' scores on the five factors should correlate with these constructs in a pattern that parallels the correlations of previous research. Our nomological network parallelled previous research, with the exceptions of the correlations with the *observe* factor and mindfulness meditation experience.

Table 5

Nomological Network of FFMQ-39 and FFMQ-15

	Awareness	Describe	Nonjudgment	Nonreactivity	Observe
Associations with FFMQ-39 Factors					
Anxiety	**-.44**	**-.31**	**-.53**	**-.32**	*.10*
Depression level	**-.48**	**-.31**	**-.50**	**-.33**	*.03*
Rumination	**-.55**	**-.31**	**-.66**	**-.38**	.05
Psychological Well-being	**.55**	**.53**	**.46**	**.45**	***.20***
Satisfaction with Life	**.32**	**.28**	**.28**	**.31**	.08
Self-compassion	**.57**	**.49**	**.56**	**.54**	**.15**
Mindfulness Meditation Experience	.05	**.15**	*-.01*	**.14**	**.25**
Associations with FFMQ-15 Factors					

Anxiety	**-.37**	**-.21**	**-.38**	**-.22**	*.03*
Depression level	**-.29**	**-.28**	**-.41**	**-.22**	*.00*
Rumination	**-.45**	**-.26**	**-.43**	**-.27**	.02
Psychological Well-being	**.36**	**.43**	**.45**	**.31**	.09
Satisfaction with Life	**.15**	**.21**	**.30**	**.24**	*.04*
Self-compassion	**.41**	**.37**	**.45**	**.35**	**.14**
Mindfulness Meditation Experience	-.01	*.08*	*.05*	*.10*	**.19**

Note. Bold correlations indicate $p < .05$. Italics indicates that the pattern of significance for the correlation observed in the current research differed from previous research.

Invariance

Conceptually, invariance analyses test the goodness of fit of the observed models within subgroups of the sample. This investigation addresses the generalizability of the five-factor structure to diverse subgroups within the Black community. For instance, *full* measurement invariance observed between genders would indicate that the five-factor structure would be consistent in both Black men and Black women. The current study specifically examined the subgroups of gender (men vs. women), mindfulness meditation experience (no experience vs. any level of experience), depression levels (minimal levels vs. at least mild levels), everyday discrimination (reporting experiencing discrimination less than once on a daily basis vs. reporting experiencing discrimination more than once on a daily basis), lifetime discrimination (did not report experiencing discrimination vs. reported experiencing discrimination), household income ($19,999 or below vs. $20,000 or higher), ethnic heritage (having one or more parent identified as African American vs. not identifying either parent as African American), and skin tone (light complexions vs. medium and dark complexions).

Following previous methodology (Hirschfeld & Brachel, 2014), each subgroup was tested with four nested models of invariance: configural, weak, strong, and strict. Configural invariance ensures that the number of latent variables and pattern of loadings of the latent

variables are similar across the subgroups. Weak invariance adds to configural invariance to ensure that the size of these loadings is similar across groups. Strong invariance adds to weak invariance to ensure that the intercepts of each factor are similar across groups. Finally, strict invariance adds to strong invariance to ensure that the residual variances and residual covariances of the factor loadings are similar between groups.

A type I ANOVA was applied to the four models to test the change in model fit between each level of invariance. *Full* invariance was achieved when a non-significant chi-square *p*-value was observed and when the difference between the model CFI's (ΔCFI) was less than .01 (Hirschfeld & Brachel, 2014). This criterion indicates that the model did not significantly differ from the previous model. Given that fit indices typically worsen with smaller sample sizes, we anticipated that fit indices might not meet conventional cutoff criteria and planned to focus on changes in indices across levels of invariance rather than the absolute level of fit indices. If we did not observe *full* invariance, the model was further investigated to identify the specific indices (i.e. factor loadings, intercepts, residuals, and residual covariances) that significantly differed from the same index in the prior model at a 95% significance level. Once the significant indices were identified, the model was re-analyzed while the identified indices were allowed to vary between groups. If invariance was achieved with the modified model, this was referred to as *partial* invariance. If the modified model did not achieve invariance, this would mean that no invariance was found within the model.

Table 6 displays results from the invariance analyses for the FFMQ-39 and the FFMQ-15. Following recommendations from Tanaka (1987), invariance analyses were applied to all measurements with subgroups of at least 50 participants within each FFMQ questionnaire. Within our large number of total participants, half of the participants completed the FFMQ-39

and half of the participants completed the FFMQ-15. The subgroups of skin tone and household income within each FFMQ questionnaire did not meet the 50 participant criteria. In order to complete the invariance analyses on these measurements, we increased the subgroup sample size by extracting the FFMQ-15 items from the FFMQ-39 and added these responses to the original FFMQ-15 responses. The extraction of the FFMQ-15 items from the FFMQ-39 items has been supported with previous research (Gu et al., 2016).

We observed *full* or *partial* invariance in all models, across all subgroups. Collectively, the five-factor structure was reliable between gender, depression level, mindfulness meditation experience, lifetime discrimination, everyday discrimination, household income, ethnic heritage, and skin tone.

Table 6

Measurement Invariance

	FFMQ-39					FFMQ-15				
	df	CFI	ΔCFI	RMSEA	Invariance	*df*	CFI	ΔCFI	RMSEA	Invariance
Gender										
Configural	1384	.783	--	.08	Full	160	.933	--	.06	Full
Weak	1418	.782	.001	.08	Full	170	.931	.002	.06	Full
Strong	1452	.781	.001	.08	Full	176	.926	.005	.06	Partial
Strict	1491	.782	.001	.08	Full	189	.924	.002	.06	Partial
Mindfulness Meditation Experience										
Configural	1384	.837	--	.07	Full	160	.912	--	.07	Full
Weak	1410	.837	0	.07	Partial	170	.912	0	.07	Full
Strong	1437	.837	0	.07	Partial	180	.913	.001	.07	Full
Strict	1476	.835	.002	.07	Partial	195	.910	.003	.07	Full
Depression level										
Configural	1384	.789	--	.07	Full	160	.91	--	0.07	Full
Weak	1418	.798	.001	.07	Full	170	.915	.005	0.07	Full
Strong	1447	.796	.002	.07	Partial	176	.916	.001	0.07	Partial
Strict	1484	.794	.002	.07	Partial	191	.913	.003	0.06	Partial
Everyday Discrimination										

Configural	1384	.791	--	.08	Full	160	.939	--	.06	Full
Weak	1418	.792	.001	.08	Full	170	.940	.001	.06	Full
Strong	1447	.793	.002	.08	Partial	180	.938	.002	.06	Full
Strict	1486	.794	.002	.08	Partial	191	.936	.002	.06	Partial
Lifetime Discrimination										
Configural	1384	.770	--	.09	Full	160	.934	--	.06	Full
Weak	1418	.770	0	.08	Full	170	.934	0	.06	Full
Strong	1452	.770	0	.08	Full	180	.932	.002	.06	Full
Strict	1491	.769	.001	.08	Full	195	.923	.009	.06	Full
Household income										
Configural	---	---	---	---	---	160	.936	--	.06	Full
Weak	---	---	---	---	---	170	.934	.002	.06	Full
Strong	---	---	---	---	---	176	.936	.002	.06	Partial
Strict	---	---	---	---	---	191	.931	.005	.06	Partial
Ethnic Heritage										
Configural	---	---	---	---	---	160	.946	--	.07	Full
Weak	---	---	---	---	---	170	.944	.002	.07	Full
Strong	---	---	---	---	---	176	.946	.002	.06	Full
Strict	---	---	---	---	---	191	.947	.001	.06	Full
Skin tone										
Configural	---	---	---	---	---	60	.939	--	.07	Full
Weak	---	---	---	---	---	170	.941	.002	.07	Full
Strong	---	---	---	---	---	180	.943	.002	.06	Full
Strict	---	---	---	---	---	195	.940	.003	.06	Full

Note. df = degrees of freedom, CFI = comparative fit index, ΔCFI = difference between prior model CFI, RMSEA = root mean squared error of approximation.

Discussion

Rigorous longitudinal and intervention studies have shown mindfulness to be associated with improved mental and physical health outcomes (Palta et al., 2012; Webb et al., 2019). This study contributes to the generalizability of mindfulness research by validating the psychometric properties of the widely used FFMQ-39 and the FFMQ-15 within a non-clinical, Black American

sample. Validating these two FFMQ questionnaires within a Black American sample is important for several reasons. First, this study is the first to validate both the FFMQ-39 and the FFMQ-15 within a non-clinical Black sample. By extension this research advances the scientific understanding, application, and measurement of mindfulness within a sample of Americans that have been previously excluded from mindfulness research studies but who represent a significant proportion of the U.S. population. Second, these findings support the use of the FFMQ questionnaires as measures of individual differences in trait mindfulness that can be used to predict important health outcomes within Black Americans. Third, it advances our ability to verify the efficacy of mindfulness interventions within the Black American community. Promising research has begun to apply mindfulness interventions to reduce stress, emotion regulation difficulties, and blood pressure within Black Americans (Palta et al., 2012; Watson-Singleton et al., 2021) but without clear assessments of mindfulness, it is unclear whether such interventions are effectively targeting mindfulness skills. The present research suggests we can be confident in our ability to assess whether such mindfulness interventions are indeed targeting mindfulness factors in Black Americans. And finally, this research not only validates the original 39-item FFMQ, but also provides evidence supporting the use of the short-form (FFMQ-15) for studies of Black Americans that require briefer assessments (e.g., large scale interventions; hospital settings), thus increasing opportunities to measure mindfulness within Black communities in the U.S.

Taken together, this study conducted a comprehensive range of psychometric analyses that investigated the usability of the FFMQ-39 and FFMQ-15 within Black Americans. Our examination included analyses of EFAs, CFAs, internal consistencies, test-retest correlations, nomological networks, and measurement invariance. The rigorous application of these methods

allows us to confidently support the use of both the FFMQ-39 and the FFMQ-15 within a Black American sample.

As demonstrated, the EFAs for both questionnaires identified the same five-factor structure observed for the original FFMQ questionnaires (Baer et al., 2006; Gu et al., 2016). CFAs indicated that the five-factor structure was a good model fit for both the FFMQ-39 and the FFMQ-15. The FFMQ-39 model required correlating the residuals of similarly worded items within the questionnaire (e.g., "It seems I am 'running on automatic' without much awareness of what I'm doing." and "I do jobs or tasks automatically without being aware of what I'm doing."). We recommend that future researchers apply these constraints when testing the model fit for the FFMQ-39. Moreover, this observation encourages future researchers to be more vigilant about item similarity during scale development and validation. Given the shorter nature of the FFMQ-15, future research can confidently apply the FFMQ-15 in order to save time while maintaining the integrity of the questionnaire.

Next, the test-retest analyses indicated a strong temporal consistency between Time 1 and Time 2 in both the FFMQ-39 and the FFMQ-15, making the current study the first FFMQ analysis to report consistent test-retest reliabilities in English. This novel finding demonstrates that the observed FFMQ factors are consistent over time in a Black American sample and supports the use of the FFMQ to measure mindfulness in longitudinal studies with Black American samples.

To further test the validity of the observed questionnaires, we established a nomological network that describes the relationship of FFMQ factors measured among Black Americans and constructs that are theoretically related to mindfulness. Our nomological network was largely consistent with previous research, with the exceptions of the associations with the *observe* factor

and mindfulness meditation experience. The inconsistent *observe* associations are understandable given that the *observe* factor has a history of varying associations with psychological assessments (Baer et al., 2008; Christopher et al., 2012; Gu et al., 2016; Rogge & Daks, 2021; Sweeney et al., 2021). This inconsistency might stem from the fact that all of the nomological network constructs refer to internal sensations (e.g. thoughts, feelings, and cognitions). While the *observe* factor is the only factor that measures both internal sensations and external sensations (e.g. smells, sights, and sounds). Thus, it is possible that the items that measure external sensations within the *observe* factor might weaken the relationship between the factor and theoretically related constructs.

Another interesting observation within the nomological network is the inconsistent relationship between mindfulness meditation experience. Specifically, in the FFMQ-39 we saw an unexpected null association between mindfulness meditation experience and *nonjudgment*. In the FFMQ-15 we saw unexpected non-associations between mindfulness meditation experience and the factors *describe, nonjudgment,* and *nonreactivity*. This difference might be explained by the low average and variance of mindfulness meditation experience in the sample. The average participant in our sample meditated less than once or twice in their lifetime. It is possible that a wider range of mindfulness meditation experience is required to accurately observe the relationship between mindfulness meditation experience and the FFMQ factors. Given that mindfulness meditation is underutilized within Black Americans (Biggers et al., 2020), future research should target a diverse representation of mindfulness meditation experience within Black Americans to better understand the relationship between mindfulness meditation experience and the FFMQ factors.

Lastly, the measurement invariance analysis is one of this study's strongest contributions.

We observed *full* and *partial* invariance between gender, mindfulness meditation experience, depression level, everyday discrimination, lifetime discrimination, household income, ethnic heritage, and skin tone within Black Americans. This contribution is groundbreaking in several ways: (1) this is the first study to assess the invariance of any mindfulness measure within Black Americans, (2) this is the first study to analyze the invariance of any FFMQ questionnaire within Black Americans; and (3) this is the first FFMQ invariance analysis to include perceived discrimination, skin tone, and household income. Furthermore, the invariance analysis advances the general understanding of diversity *within* Black American samples – an area that is vastly under researched within the fields of psychology, mental health, and medicine (Buchanan et al., 2021). Overall, the invariance analysis supports the generalizability of applying the FFMQ questionnaires across Black Americans.

The results of this study should be interpreted in light of several limitations. First, our comprehensive nomological network supported the construct validity of the FFMQ in a Black sample using theoretically related constructs. However, due to the underrepresentation of Black people within psychology research, most of the measures in the nomological network have not been validated within a Black sample. The BDI-II (Dutton et al., 2004) and the Self-Compassion Scale (Zhang et al., 2019) are the only nomological network measures that have been validated within a Black sample. Future research should validate the remaining psychological measures to ensure their applicability to a Black sample. Secondly, although we observed invariance of the FFMQ between Black Americans who descended from the enslaved and Black Americans who immigrated from Africa or the Caribbean region, the ethnic heritage variable used for these analyses was derived from the ethnicity of the participants' parents, as opposed to a self-reported ethnic heritage variable. Parents' ethnicity did predict participants' ethnic heritage with very high

accuracy (88%). Still, future research should collect self-reported ethnic heritage during initial research stages. Finally, our detailed inclusion of heterogeneous demographics within our Black American sample is a key asset to this study because it supported the generalizability and invariance of the observed FFMQ model structure. Notably, subsamples of household income, ethnic heritage, and skin tone were not large enough to perform invariance analyses on both the FFMQ-39 and the FFMQ-15. Fortunately, this issue was ameliorated in the invariance analysis by extracting the FFMQ-15 items from the FFMQ-39 responses to adequately power this test. To avoid this issue in the future, researchers should employ methods to ensure socioeconomic diversity within the sample and collect comprehensive within group demographics during initial stages of data collection.

Taken together, the results of this research are foundational to understanding and measuring mindfulness within a non-clinical Black sample. The psychometric properties of the FFMQ-39 and the FFMQ-15 indicate that both questionnaires can validly measure mindfulness within a Black sample. Given that mindfulness is a valuable practice and an effective treatment to improve mental and physical health, it is essential that mindfulness can be validly measured in all populations — particularly disenfranchised communities of color who experience disproportionately poorer health outcomes.

References

Adewale, V., Ritchie, D., & Skeels, S. E. (2016). African-American and African perspectives on mental health: A pilot study of the pre and post colonial and slavery influences and their implications on mental health. *Journal of Communication in Healthcare*, *9*(2), 78–89. https://doi.org/10.1080/17538068.2016.1170316

Aichholzer, J. (2014). Random intercept EFA of personality scales. *Journal of Research in Personality*, *53*, 1–4. https://doi.org/10.1016/j.jrp.2014.07.001

Baer, R. A., Carmody, J., & Hunsinger, M. (2012). Weekly Change in Mindfulness and Perceived Stress in a Mindfulness-Based Stress Reduction Program. *Journal of Clinical Psychology*, *68*(7), 755–765. https://doi.org/10.1002/jclp.21865

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*(1), 27–45. https://doi.org/10.1177/1073191105283504

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct Validity of the Five Facet Mindfulness Questionnaire in Meditating and Nonmeditating Samples. *Assessment*, *15*(3), 329–342. https://doi.org/10.1177/1073191107313003

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory (BDI-II)* (Vol. 10, p. s15327752jpa6703_13). Pearson.

Beydoun, M. A., Dore, G. A., Canas, J.-A., Liang, H., Beydoun, H. A., Evans, M. K., & Zonderman, A. B. (2018). Systemic inflammation is associated with longitudinal changes in cognitive performance among urban adults. *Frontiers in Aging Neuroscience*, *10*. http://dx.doi.org/10.3389/fnagi.2018.00313

Biggers, A., Spears, C. A., Sanders, K., Ong, J., Sharp, L. K., & Gerber, B. S. (2020). Promoting Mindfulness in African American Communities. *Mindfulness*, *11*(10), 2274–2282. https://doi.org/10.1007/s12671-020-01480-w

Blanke, E. S., Schmidt, M. J., Riediger, M., & Brose, A. (2020). Thinking mindfully: How mindfulness relates to rumination and reflection in daily life. *Emotion*, *20*(8), 1369–1381. https://doi.org/10.1037/emo0000659

Bodhi, B. (2011). What does mindfulness really mean? A canonical perspective. *Contemporary Buddhism*, *12*(1), 19–39. https://doi.org/10.1080/14639947.2011.564813

Bond, S., & Cash, T. F. (1992). Black Beauty: Skin Color and Body Images among African-American College Women1. *Journal of Applied Social Psychology*, *22*(11), 874–888. https://doi.org/10.1111/j.1559-1816.1992.tb00930.x

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Buchanan, N. T., Perez, M., Prinstein, M. J., & Thurston, I. (2020). *Upending Racism in Psychological Science: Strategies to Change How Our Science is Conducted, Reported, Reviewed & Disseminated*. PsyArXiv. https://doi.org/10.31234/osf.io/6nk4x

Carmody, J., & Baer, R. A. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. Journal of Behavioral Medicine, 31(1), 23–33. https://doi.org/10.1007/s10865-007-9130-7

Chiesa, A., & Serretti, A. (2009). Mindfulness-Based Stress Reduction for Stress Management in Healthy People: A Review and Meta-Analysis. *The Journal of Alternative and Complementary Medicine*, *15*(5), 593–600. https://doi.org/10.1089/acm.2008.0495

Christopher, M. S., Neuser, N. J., Michael, P. G., & Baitmangalkar, A. (2012). Exploring the Psychometric Properties of the Five Facet Mindfulness Questionnaire. *Mindfulness*, *3*(2), 124–131. https://doi.org/10.1007/s12671-011-0086-x

Cortina, J. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*. /doiLanding?doi=10.1037%2F0021-9010.78.1.98

DeLuca, S. M., Kelman, A. R., & Waelde, L. C. (2018). A Systematic Review of Ethnoracial Representation and Cultural Adaptation of Mindfulness- and Meditation-Based Interventions. *Psychological Studies*, *63*(2), 117–129. https://doi.org/10.1007/s12646-018-0452-z

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Dutton, G. R., Grothe, K. B., Jones, G. N., Whitehead, D., Kendra, K., & Brantley, P. J. (2004). Use of the Beck Depression Inventory-II with African American primary care patients. *General Hospital Psychiatry*, *26*(6), 437–442. https://doi.org/10.1016/j.genhosppsych.2004.06.002

Earl, T., Williams, D., & Anglade, S. (2011). An Update on the Mental Health of Black Americans: Puzzling Dilemmas and Needed Research. *Journal of Black Psychology*, *37*(4), 485–498. https://doi.org/10.1177/0095798410396077

Floyd, F. J., & Widaman, K. F. (1995). Factor Analysis In the Development and Refinement of Clinical Assessment Instruments. *Factor Analysis*, 14.

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*(7), 791–802. https://doi.org/10.1037/pas0000263

Hirschfeld, G., & Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R – A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment Resaerch and Evaluation*, *19*, 1–12.

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

Johnson, C. C., Sheffield, K. M., & Brown, R. E. (2018). Mind-Body Therapies for African American Women at Risk for Cardiometabolic Disease: A Systematic Review. *Evidence Based Complementary and Alternative Medicine*, *2018*, 1–11. https://doi.org/10.1155/2018/5123217

Karl, J. A., Prado, S. M. M., Gračanin, A., Verhaeghen, P., Ramos, A., Mandal, S. P., Michalak, J., Zhang, C.-Q., Schmidt, C., Tran, U. S., Druica, E., Solem, S., Astani, A., Liu, X., Luciano, J. V., Tkalčić, M., Lilja, J. L., Dundas, I., Wong, S. Y. S., & Fischer, R. (2020). The Cross-cultural Validity of the Five-Facet Mindfulness Questionnaire Across 16 Countries. *Mindfulness*, *11*(5), 1226–1237. https://doi.org/10.1007/s12671-020-01333-6

Kessler, R. C., Mickelson, K. D., & Williams, D. R. (1999). The Prevalence, Distribution, and Mental Health Correlates of Perceived Discrimination in the United States. *Journal of Health and Social Behavior*, *40*(3), 208–230. https://doi.org/10.2307/2676349

Lewis, T. T., Aiello, A. E., Leurgans, S., Kelly, J., & Barnes, L. L. (2010). Self-reported Experiences of Everyday Discrimination are associated with Elevated C-Reactive Protein

levels in older African-American Adults. *Brain, Behavior, and Immunity*, *24*(3), 438–443. https://doi.org/10.1016/j.bbi.2009.11.011

Neff, K. D. (2003). The Development and Validation of a Scale to Measure Self-Compassion. *Self and Identity*, *2*(3), 223–250. https://doi.org/10.1080/15298860309027

Oh, H., Lincoln, K., & Waldman, K. (2021). Perceived colorism and lifetime psychiatric disorders among Black American adults: Findings from the National Survey of American Life. *Social Psychiatry and Psychiatric Epidemiology*. https://doi.org/10.1007/s00127-02102102-z

Palta, P., Page, G., Piferi, R. L., Gill, J. M., Hayat, M. J., Connolly, A. B., & Szanton, S. L. (2012). Evaluation of a Mindfulness-Based Intervention Program to Decrease Blood Pressure in Low-Income African-American Older Adults. *Journal of Urban Health*, *89*(2), 308–316. https://doi.org/10.1007/s11524-011-9654-6

R Core Team. (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna Austria. Retrieved from http://www.R-project.org/

Revelle W (2021). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2.1.9, https://CRAN.R-project.org/package=psych

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*(3), 250–255. https://doi.org/10.1002/cpp.702

Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial Inequality in Psychological Research: Trends of the Past and Recommendations for the Future. *Perspectives on Psychological Science*, *15*(6), 1295–1309.

https://doi.org/10.1177/1745691620927709

Rogge, R. D., & Daks, J. S. (2021). Embracing the Intricacies of the Path Toward Mindfulness: Broadening Our Conceptualization of the Process of Cultivating Mindfulness in Day-to-Day Life by Developing the Unified Flexibility and Mindfulness Model. *Mindfulness*, *12*(3), 701–721. https://doi.org/10.1007/s12671-020-01537-w

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*. Retrieved from http://www.doaj.org/doaj?func=fulltext&aId=1325391

Ryff, C. D., Lee, C., & Keyes, M. (1995). The structure of psychological well-being revisited. *Journal of Personality & Social Psychology*, 719–727.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness of fit measures. *Methods of Psychological Research*, 23–74.

Svendsen, J. L., Kvernenes, K. V., Wiker, A. S., & Dundas, I. (2017). Mechanisms of mindfulness: Rumination and self-compassion. *Nordic Psychology*, *69*(2), 71–82. https://doi.org/10.1080/19012276.2016.1171730

Sweeney, T., Morriss, R., Nixon, E., Guo, B., & Callaghan, P. (2021). Psychometric Properties of the Five Facets Mindfulness Questionnaire in Moderate-to-Severe, Persistent Depression. *Mindfulness*, *12*(4), 1009–1021. https://doi.org/10.1007/s12671-020-01573-6

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics* (Seventh edition). Pearson.

Tanaka, J. S. (1987). “How Big Is Big Enough?”: Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables. *Child Development*, *58*(1), 134–146. https://doi.org/10.2307/1130296

Trapnell, P., & Campbell, J. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, *76*, 284–304. https://doi.org/10.1037//0022-3514.76.2.284

Watson-Singleton, N. N., Pennefather, J., & Trusty, T. (2021). Can a culturally-responsive Mobile health (mHealth) application reduce African Americans' stress?: A pilot feasibility study. *Current Psychology*. https://doi.org/10.1007/s12144-021-01534-9

Watson-Singleton, N. N., Walker, J. H., LoParo, D., Mack, S. A., & Kaslow, N. J. (2018). Psychometric Evaluation of the Five Facet Mindfulness Questionnaire in a Clinical Sample of African Americans. *Mindfulness*, *9*(1), 312–324. https://doi.org/10.1007/s12671-017-0776-0

Webb, C. A., Beard, C., Forgeard, M., & Björgvinsson, T. (2019). Facets of Mindfulness Predict Depressive and Anxiety Symptom Improvement Above CBT Skills. *Mindfulness*, *10*(3), 559–570. https://doi.org/10.1007/s12671-018-1005-1

Williams, D. R., Yu, Y., Jackson, J. S., & Anderson, N. B. (1997). Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. *Journal of Health Psychology*, *2*(3), 335–351. https://doi.org/10.1177/135910539700200305

Zhang, H., Dong, L., Watson-Singleton, N. N., Tarantino, N., Carr, E. R., Niles-Carnes, L. V., Patterson, B., & Kaslow, N. J. (2019). Psychometric Properties of the Self-Compassion Scale (SCS) in an African American Clinical Sample. *Mindfulness*, *10*(7), 1395–1405. https://doi.org/10.1007/s12671-019-01099-6

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FFMQBlackValidationSupplementalMaterial.pdf
- DataDepositionInformationItem.pdf