Mental Help Seeking Attitudes Scale (MHSAS): Development, Reliability, Validity, and Comparison with the ATSSPH-SF and IASMHS-PO

Note: This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

To obtain a copy of the MHSAS and/or access the Supplemental Material for the article, please visit:
http://drjosephhammer.com/research/mental-help-seeking-attitudes-scale-mhsas/

APA-style citation for this article:
Hammer, J. H., & Parent, M. C., & Spiker, D. A. (2018). Mental Help Seeking Attitudes Scale (MHSAS): Development, reliability, validity, and comparison with the ATSSPH-SF and IASMHS-PO. *Journal of Counseling Psychology, 65*, 74-85. doi: 10.1037/cou0000248

Authors:
Joseph H. Hammer, Ph.D.
243 Dickey Hall
Department of Educational, School, and Counseling Psychology
University of Kentucky
Lexington, KY 40506
jhhammer@gmail.com

Mike C. Parent, Ph.D.
Department of Psychological Sciences
Texas Tech University
Lubbock, TX 79409
michael.parent@ttu.edu

Douglas A. Spiker, M.A.
251 Dickey Hall
Department of Educational, School, and Counseling Psychology
University of Kentucky
Lexington, KY 40506

Abstract

Attitudes is a key help seeking construct that influences treatment seeking behavior via intention to seek help, per the Theory of Planned Behavior. This article presents the development and psychometric evaluation of the Mental Help Seeking Attitudes Scale (MHSAS), designed to measure respondents' overall evaluation (unfavorable vs. favorable) of their seeking help from a mental health professional. In Study 1 (N = 857 U.S. adults), exploratory and confirmatory factor analysis and Item Response Theory analysis were used to identify an optimal set of nine items that demonstrated initial evidence of internal consistency, unidimensionality, and strong measurement equivalence/invariance across gender, past help seeking experience, and psychological distress. Initial convergent evidence of validity was demonstrated via theoretically-anticipated relationships between the MHSAS and key variables in the help seeking nomological network (e.g., subjective norms, perceived behavioral control, intention, public stigma, self-stigma, anticipated risks and benefits, gender, previous help seeking). Initial incremental evidence of validity was demonstrated when the MHSAS demonstrated the ability to account for unique variance in help seeking intention, beyond that accounted for by the Attitudes Toward Seeking Professional Psychological Help – Short Form scale (ATSPPH-SF) and the Psychological Openness subscale of the Inventory of Attitudes Toward Seeking Mental Health Services (IASMHS-PO). Study 2 (N = 207 U.S. adults at Times 1 and 2) provided initial evidence of test-retest reliability over a three-week period. The MHSAS offers mental health professionals a new tool for measuring attitudes that may avoid limitations of current help seeking attitudes measures (e.g., construct-irrelevant variance).

Key words: scale development, help seeking, attitudes, validity, reliability

Public Significance Statement: This study developed a self-report instrument designed to measure people's attitudes toward seeking mental health services that avoids some limitations of existing instruments. Measuring help seeking attitudes accurately is important because attitudes have been shown to be an important factor in whether a person will seek mental health services when they are having psychological difficulties.

Mental Help Seeking Attitudes Scale (MHSAS): Reliability, Validity, and Comparison with the ATSSPH-SF and IASMHS-PO

Only a third of people struggling with mental health concerns receive professional help each year (Wang et al., 2007). To understand and close this treatment gap, researchers have examined a variety of logistical, demographic, and psychological factors (Nam et al., 2013). One of the most cited factors thought to influence the help seeking process is people's help seeking attitudes. Help seeking attitudes are people's overall evaluation (i.e., good vs. bad) of the act of seeking help from a mental health professional. When applied to the help seeking context, the Theory of Planned Behavior (TPB; Ajzen, 1985), an extension of the Theory of Reasoned Action (Ajzen & Fishbein, 1980), states that help seeking attitudes, in combination with subjective norms (i.e., perceived social pressure from important others to seek or not seek help) and perceived behavior control (i.e., perceived efficacy and control regarding seeking help), predict people's behavioral intention to seek help. Specifically, favorable attitudes, favorable subjective norms, and greater perceived behavioral control lead to the formation of behavioral intention, the precursor to actual future help seeking behavior. In turn, the TPB states that people who intend to seek help will do so when the opportunity arises.

Help seeking attitudes is important to study because meta-analytic evidence suggests that attitudes is the strongest predictor of help seeking intention (Li, Dorstyn, & Denson, 2014). Furthermore, in studies where intention was not measured, attitudes directly accounted for variance in actual future help seeking behavior (e.g., Andrykowski & Burris, 2010). Thus, help seeking attitudes is a construct of central importance in help seeking research and practice. Recognizing this importance, researchers have extensively studied the relationship between help seeking attitudes and related constructs. Of note, meta-analyses suggest that help seeking attitudes are related to acculturation/enculturation (Sun, Hoyt, Brockberg, Lam, & Tiwari, 2016); gender (Nam et al., 2010); conformity to masculine norms (Wong, Moon-Ho, Wang, & Miller, 2017); anticipated benefits and risks of seeking help, self-stigma of seeking help, public stigma of seeking help or of mental illness, self-disclosure, self-concealment, social support, and

depression (Nam et al., 2013). In summary, because people's attitudes influence their behavior via intention (Armitage & Conner, 2001), counseling psychologists could increase help seeking behavior by implementing prevention and intervention programs that improve people's help seeking attitudes (Gulliver, Griffiths, Christensen, & Brewer, 2012; Hammer & Vogel, 2010).

Inaccurate measurement of psychological constructs such as help seeking attitudes can lead to incorrect inferences and misleading conclusions (Carmines & Zeller, 1979). Therefore, it is essential that counseling psychologists use psychometrically-sound instruments to measure help seeking attitudes. We argue that existing help seeking attitudes instruments have serious limitations that necessitate the development of a new instrument that meets modern psychometric standards: the Mental Help Seeking Attitudes Scale (MHSAS). The MHSAS can help researchers advance scholarship on what helps and stops people from seeking mental healthcare, and could help clinicians more accurately understand their clients' perceptions of seeking help.

Limitations of Existing Instruments

The 29-item Attitudes Toward Seeking Professional Psychological Help scale (ATSPPH; Fischer & Turner, 1970) was the first help seeking attitudes instrument to see wide use. The ATSPPHS was not grounded in a theoretical framework and psychometric problems led to the development of the 10-item short form (ATSPPH-SF; Fischer & Farina, 1995), designed to produce a single. However, evidence is mixed regarding the unidimensionality of the ATSPPH-SF (e.g., Elhai, Schweinle, & Anderson, 2008), including across cultural groups (e.g., Singapore residents; Picco et al., 2016; Taiwanese college students; Chang, 2007). Relatedly, the ATSPPH-SF score's internal consistency estimates sometimes fall below optimal levels (e.g., α = .35; Leach, Jana-Masri, & Priester, 2009). This is problematic, as instruments should consistently demonstrate the intended factor structure and adequate reliability to be considered psychometrically sound (DeVellis, 2016).

The *Standards for Educational and Psychological Testing* (*Standards*; American Education Research Association [AERA], American Psychological Association, & National Council on Measurement in Education, 2014) state that an instrument should only measure what

the instrument is intended to measure, not more (i.e., construct-irrelevant variance). Measuring construct-irrelevant variance reduces the validity of the instrument score (historically referred to as "content validity;" DeVellis, 2016). Certain ATSPPH-SF items appear to measure help seeking intention (e.g., "I would want to get psychological help if I were worried or upset for a long period of time") or mental health literacy (e.g., "Psychological problems, like many things, tend to work out by themselves"), instead of attitudes. Therefore, use of the ATSPPH-SF may invite tautological conclusions (e.g., correlating a thing with itself). For example, correlating the ATSPPH-SF's score (which partially measures intention) with the score of a help seeking intention instrument could result in a biased understanding of the true relationship between attitudes and intention. This can lead to inaccurate research and clinical conclusions.

To address limitations of the ATSPPH-SF, Mackenzie, Knox, Gekoski, and Macaulay (2004) developed the 24-item Inventory of Attitudes Toward Seeking Mental Health Services (IASMHS). The IASMHS was designed to be a "multifactorial, theoretically based attitude inventory" (p. 2414) grounded in the TPB. Factor analysis was used to identify 24 items measuring three factors: Psychological Openness (PO), Help-Seeking Propensity (HSP), and Indifference to Stigma (ITS). However, each factor appears to measure multiple constructs. For example, PO items seem to measure mental health literacy (e.g., "Psychological problems, like many things, tend to work out by themselves), mental illness self-disclosure (e.g., "There are certain problems which should not be discussed outside of one's immediate family") avoidance coping (e.g., "Keeping one's mind on a job is a good solution for avoiding personal worries and concerns"), or self-concept clarity (e.g., "It is probably best not to know everything about oneself"). Thus, in the words of the developers, "each factor represents a key construct of Ajzen's (1985) TPB, although not exclusively so" (Makenzie et al., 2004; p. 2427).

We believe these multiply-determined subscale scores are problematic for two reasons. First, the developers stated that the TPB guided the development of the IASMHS. This has led some researchers to treat the PO, HSP, and ITS subscale scores as if they are pure measures of attitudes, subjective norms, and perceived behavioral control, respectively (e.g., Hyland,

Boduszek, Dhingra, Shevlin, Maguire, & Morley, 2015). However, from the perspective of the TPB, these subscale scores contain construct-irrelevant variance and are invalid measures of these three constructs. Second, the multiple constructs tapped by each IASMHS subscale score have already been defined, operationalized, and studied as independent constructs. Repacking these constructs into a new global construct with a novel name (e.g., "psychological openness") may offer a fresh perspective, but may also fuel construct proliferation (Shaffer, DeGeest, & Li, 2016) and create confusion among scholars.

Exacerbating this confusion is that, while the TPB treats attitudes as a single construct to be measured with a unidimensional instrument that produces a single overall evaluation score (Ajzen, 2006), the IASMHS developers defined help seeking attitudes more broadly than the TPB, such that attitudes, subjective norms, and perceived behavioral control are all considered "help seeking attitudes." This has led to confusion: some researchers describe the IASMHS as a measure of different dimensions of help seeking attitudes (e.g., Dunford & Granger, 2017), but others describe the IASMHS as measuring the three independent constructs of attitudes, subjective norms, and perceived behavioral control (e.g., Hyland et al., 2015).

It is, of course, permissible to develop a help seeking attitudes instrument using a theoretical framework other than the TPB. However, the IASMHS developers specifically sought "to extend the ATSPPH to include new items according to Ajzen's (1985) TPB" (Mackenzie et al., 2004; p. 2414). It seems problematic that the IASMHS does not operationalize help seeking attitudes in line with the requirements of TPB, as the TPB is well-suited to helping social scientists understanding help seeking behaviors. Specifically, the TPB is parsimonious, enjoys strong empirical support both within (Vogel & Heath, 2015) and outside (Armitage & Conner, 2001) of help seeking contexts, and demonstrates cross-cultural utility (e.g., Kaiser & Gutscher, 2003; Mo & Mak, 2009; Nigg, Lippke, & Maddock, 2008). Given that the TPB is one of the most cited social scientific theory in help seeking research, it seems prudent to ground help seeking attitudes instruments in the TPB.

In addition to measuring unintended constructs, some IASMHS items may not clearly measure help seeking attitudes (i.e., construct underrepresentation). As noted previously, help seeking attitudes are people's overall evaluation (i.e., good vs. bad) of the act of seeking help from a mental health professional. Examination of the PO subscale, which is the only IASMHS subscale that partially measures what the TPB would define as attitudes, reveals several items do not clearly assess how good or bad the individual thinks seeking help is. For example, if person A thinks seeking help is bad and person B thinks seeking help is good, both may agree "it is probably best not to know everything about oneself." The question of how much personal insight a healthy human being should possess is separate from the question of whether seeking professional help is good or bad. Thus, while it may certainly correlate with other PO items, this item may not measure the evaluative component that defines help seeking attitudes. Furthermore, the IASMHS has received mixed support for its cross-cultural utility (e.g., bereaved U.S. adults, Drapeau, Cerel, & Moore, 2016; international and African American college students, Mesidor & Sly, 2014; Japanese undergraduates, Mojaverian, Hashimoto, & Kim, 2013). In summary, the ATSPPH-SF and IASMHS demonstrate limitations related to construct-irrelevant variance, construct underrepresentation, and internal structure instability across cultural contexts.

Nomological Network of Help Seeking Attitudes

Having established the limitations of existing instruments that necessitated the development of the MHSAS, we now describe the nomological network in which help seeking attitudes is embedded. This network was constructed by referencing theory and empirical findings. The network was used to articulate empirically-testable hypotheses about how the MHSAS score should relate to other constructs. Anticipated effect sizes were directly sourced from extant literature (see the Supplemental Material for references supporting the anticipated direction and strength of hypothesized effects). If the MHSAS score was found to relate to other constructs in the hypothesized manner, this would constitute initial evidence of validity for the MHSAS score (AERA et al., 2014).

Per the TPB, more favorable attitudes (i.e., higher MHSAS score) should be associated with more favorable subjective norms (Hypothesis 1; H1), greater perceived behavioral control (H2), greater intention (H3), and previous experience with seeking mental health treatment (H4; Ajzen, 1985). Per the Internalized Stigma Model (Lannin, Vogel, Brenner, & Tucker, 2015), more favorable attitudes should be inversely associated with both the public stigma (H5) and self-stigma of seeking help (H6). In other words, perceiving greater societal level stigma towards those who seek help, and stronger internalization of this societal prejudice, should relate to less favorable attitudes. Per the masculine role socialization paradigm (Addis and Mahalik, 2003), because men living in the U.S. tend to conform more strongly to traditional western masculine norms—such as self-reliance and emotional control—that discourage help seeking than do women, women should report more favorable attitudes than men (H7; Mahalik, Lagan, & Morrison, 2006). Per the Attachment Model of Help Seeking (Shaffer, Vogel, & Wei, 2006), "attitudes are formed through an evaluation and weighting of the anticipated outcomes (i.e., the benefits and risks) of seeking help" (p. 444). Thus, the anticipated risks (e.g., vulnerability in disclosing something very personal) and benefits (e.g., getting useful response to one's disclosures) of seeking professional help should be associated with more unfavorable (H8) and more favorable (H9) attitudes, respectively. In addition, because the instruments all purport to measure help seeking attitudes construct, the MHSAS score should correlate strongly with the ATSPPH-SF and IASMHS-PO scores (H10).

In addition to these hypothesized associations, we anticipated that self-identified gender, previous experience with seeking mental health treatment, and psychological distress may each influence how respondents interpret the meaning of the MHSAS items. First, gender role socialization systematically influences our perceptions of the world, and it is possible that men and women may conceptualize the utility of help seeking in different ways (Addis & Mahalik, 2003). Second, those without previous help seeking experience may hold qualitatively different views of what therapy is and does than those who have experienced treatment (Gulliver, Griffiths, & Christensen, 2010). Third, those individuals with little current psychological

distress who are asked their opinion about seeking professional help if they were to have a mental health concern in the future may formulate and express their attitude toward seeking help in a different (e.g., more dispassionate) manner than those whose who are in serious psychological distress (Komiya et al., 2000).

Before researchers can use a help seeking attitudes instrument to validly examine group differences on attitudes or compare the association between attitudes and other constructs of interest across groups, it is important to empirically establish that the instrument has the same theoretical structure and meaning for each group (Miller & Sheu, 2008). This is accomplished by testing an instrument for measurement equivalence/invariance (ME/I). In the absence of evidence of ME/I, group difference and comparative association results cannot be trusted, as they may merely be the product of differences in measurement, rather than true differences in the degree of the construct (Cheung & Lau, 2012). Existing measures have either demonstrated inconsistent evidence of ME/I across key groups (i.e., ATSPPH-SF; Chong, 2016; Nam & Lee, 2015) or have never been ME/I tested (e.g., IASMHS). Given that researchers and clinicians will be interested in using the MHSAS across various groups, we sought to provide evidence that the MHSAS demonstrates ME/I across men and women (H11), those who have previously sought help from a mental health professional and those who have not (H12), and those who are experiencing both low stress and moderate to high psychological distress (H13).

The Present Study

The present study sought to develop and examine the psychometric properties of the MHSAS. In addition to the 13 aforementioned hypotheses, we proposed four additional hypotheses regarding the psychometric performance of the MHSAS. The *Standards* state "analysis of the internal structure of a test can indicate the degree to which the relationships among test items and test components conform to the construct on which the proposed test score interpretations are based" (AERA et al., 2014, p. 16). The TPB suggests the construct of attitudes is internally consistent, relatively temporally stable, and unidimensional (Ajzen, 2001; Ajzen, 2006). Therefore, initial internal structure evidence of validity and reliability for the

MHSAS score would be provided if the MHSAS demonstrates adequate internal consistency (H14), temporal stability (i.e., test-retest reliability; H15), and a unidimensional factor structure (H16). Lastly, the utility of using the MHSAS in lieu of the ATSPPH-SF or IASMHS-PO would be demonstrated if the MHSAS was able to account for incremental variance in intention beyond the variance accounted for by these two instruments (H17).

Study 1

Study 1 involved the development and refinement of the item pool, factor and item selection, and examination of initial evidence for the reliability and validity of the MHSAS.

Method

Scale development. Using the TPB definition of attitudes (see Introduction) and Ajzen's (2006) guidelines for constructing a TPB attitudes instrument, we first created the instrument stem. The stem accounts for the TPB's principle of compatibility, which states that all TPB constructs (e.g., attitudes, intention) should be defined using the same elements of action, target, context, and time. In the case of the MHSAS, the stem, "If I had a mental health concern, seeking help from a mental health professional would be…" implies an action (i.e., seeking help), target (i.e., mental health professional), context (i.e., for assistance with a mental health concern), and time (i.e., upon the development of a mental health concern). We used the term "mental health professional" in lieu of "psychologist" or "professional" because many scholars seek to study attitudes toward seeking help from mental health professionals as a group (e.g., Komiya et al., 2000). Next, the seven-point semantic differential scale recommended by Ajzen (2006) was chosen for the MHSAS.

Ajzen (2006) recommends creating a large set (i.e., "20 to 30"; p. 5) of bipolar adjective pairs. As recommended by Ajzen (2006) and DeVellis (2016), we drew upon published adjective scale lists (Osgood, Suci, & Tannenbaum, 1957), empirical literature (e.g., Bayer & Peay, 1997; Hammer & Vogel, 2017), respondent suggestions (N = 20 convenience sample of university students in the authors' professional networks were asked to provide relevant adjectives with opposite meanings via a Qualtrics survey), and rational deduction to generate 46

adjective pairs (35 from the literature, 11 from respondent suggestions). We then sought three forms of feedback on these items.

First, 126 community adults (same inclusion criteria, recruitment process, and survey procedure as main Study 1 sample [see below]) indicated via free response that the MHSAS instructions were clear. Each item selected for the final version of the instrument (see below) was rated as clear by at least 94% of the respondents (i.e., binary response of "clear" vs. "unclear"). Second, we asked five experts with counseling psychology doctorate degrees who have published help seeking research to evaluate the clarity of the MHSAS' instructions and test content evidence of validity for the items. A Microsoft Word document containing the definition of the construct was provided to the experts, who were asked to rate each item on a scale ranging from 1 (*does not fit the construct at all*) to 5 (*fits the construct very well*). With the exception of unsatisfying/satisfying (M = 2.8), each item selected for the final version of the instrument achieved a mean test content evidence of validity score of at least 4.0. Experts also provided three new adjective pairs that were added to the item pool, resulting in 49-item pool.

Third, Ajzen (2006) and other attitude researchers recommend that semantic differential scales utilize adjective pairs that relate meaningfully to the context (Heise, 1970). Thus, we asked an additional 113 community adults to provide ratings of how meaningful and relevant each adjective pair was when it comes to evaluating the act of seeking help from a mental health professional, from 1 (*not meaningful and relevant at all)* to 3 (*somewhat meaningful and relevant*) to 5 (*very meaningful and relevant*). Each item selected for the final version of the instrument achieved a mean score of at least 2.9.

To ensure that all 49 candidate items were measuring the evaluation dimension (i.e., unfavorable vs. favorable) of attitudes per Ajzen's (2006) guidelines, we also included sixteen items measuring potency (e.g., powerful vs. powerless) or activity (e.g., fast vs. slow) dimensions (see Supplemental Material). Per the recommendations of Ajzen (2006), we balanced the valence of the item anchors to counteract possible response sets. Half of the items had the positively-valanced pole on the right side of the page (i.e., bad-good), whereas the other

half had the negatively-valenced pole on the right side of the page (i.e., healthy-unhealthy). Thus, prior to analyses, items were recoded so that a higher score indicated a more positive evaluation of help seeking. We also randomized the order of the items such that each participant would be presented with a randomized list of all items. The final pool consisted of 65 items.

Procedure. To avoid the limited age range of college student samples, a community adult Study 1 sample (N = 857) was recruited to guide the selection of items for, and the psychometric evaluation of, the instrument. The Study 1 sample was recruited via ResearchMatch, a national health volunteer registry created by several academic institutions and supported by the U.S. National Institutes of Health as part of the Clinical Translational Science Award program. ResearchMatch has a large population of volunteers who have consented to be contacted by researchers about health studies for which they may be eligible. Review and approval for this study and all procedures was obtained from the University of Kentucky Office of Research Integrity. Participants were contacted via the registry messaging system regarding the study, which was advertised as a therapy questionnaire. Inclusion criteria included minimum age of 18 and fluency in written English. Interested participants were directed to an online survey that began with an informed consent page, continued with the survey items (see Measures section; instrument order same for all participants, with the 65 MHSAS items being first), and ended with a conclusion page. Participants had the option of entering a drawing for one of several $25 Amazon.com gift cards. Median survey completion time was 16.9 minutes.

Participants. Study 1 participants were 857 (257 men, 592 women, 5 other gender identity, 3 preferred not to answer) community adults. Participants were 18 to 85 (M = 44.30, *SD* = 26.34) years of age. Approximately 67% of the sample identified as White, 11% as African-American/Black, 7% as Latino/a, 7% as Multiracial, 6% as Asian American/Pacific Islander, 1% Other, and 1% Native American/Alaskan Native. Approximately 1% reported having less than a high school diploma, 4% earned a high school diploma or GED, 9% earned an associate's degree or attended vocational school, 14% had some college, 35% earned a bachelor's, 38% earned a

graduate/professional degree, and 1% preferred not to answer. Approximately 27% reported they had never sought help from a mental health professional, while 72% reported they had.

Measures. Table 1 provides the correlations, means, standard deviations, and internal consistency estimates for all instruments for the total Study 1 sample. Validity and reliability information on each instrument, as well as a full copy of the MHSAS, is provided in the Supplemental Material. For all instruments, higher scores indicate more of that construct (e.g., more favorable attitudes, more stigma, stronger intention, more favorable subjective norms).

***Demographics*.** Gender, age, race, and education items were included.

Previous help seeking. Previous help-seeking was assessed with a single yes/no item: "Have you ever sought help from a mental health professional (e.g., psychologist, psychiatrist, social worker, or counselor)?

Psychological distress. The 6-item Kessler Psychological Distress Scale (K6; Kessler et al., 2002) assesses nonspecific depressive and anxiety symptoms in the past 30 days. An example items is "During the past 30 days, how often did you feel nervous?" Participants respond using a 5-point scale ranging from 1 (*None of the time*) to 5 (*All of the time*).

Public stigma of seeking help. The 5-item Social Stigma of Receiving Psychological Help scale (SSRPH; Komiya et al., 2000) assesses perceived awareness of the larger societal stigmas associated of seeking help. An example item is "People will see a person in a less favorable way if they come to know that he/she has seen a psychologist." Participants respond using a 4-point Likert scale from 1 (*strongly disagree*) to 4 (*strongly agree*).

Self-stigma of seeking help. The 10-item Self-stigma of Seeking Help scale (SSOSH; Vogel et al., 2006) assesses perceived self-stigma associated with seeking psychological help. An example item is "I would feel inadequate if I went to a therapist for psychological help." Participants respond using a 5-point scale 1 (*strongly disagree*) to 5 (*strongly agree*). Five items are reverse-coded.

Anticipated utility and anticipated risk. The anticipated utility and the anticipated risk of seeking help from a counselor were measured with the anticipated utility (example item: "Would

you feel better if you disclosed feelings or sadness or anxiety?") and anticipated risk (example item: "How risky would it feel to disclose your hidden feelings to a counselor") subscales, respectively, of the 8-item Disclosure Expectations Scale (DES; Vogel & Wester, 2003). Participants respond to the two 4-item subscales using a 5-point Likert-type scale ranging from 1 (*not at all*) to 5 (*very*).

MHSAS. The 65 MHSAS items assess respondents' overall evaluation (unfavorable vs. favorable) of their seeking help from a mental health professional if they found themselves to be dealing with a mental health concern. Participants respond using a 7-point semantic differential scale anchored by bipolar adjectives at either end (e.g., good vs. bad).

ATSPPH-SF. The 10-item Attitudes Towards Seeking Professional Psychological Help – Short Form scale (Fischer & Farina, 1995) was designed to assess attitudes toward seeking professional psychological help. An example item is "The idea of talking about problems with a psychologist strikes me as a poor way to get rid of emotional conflicts." Participants respond using a 4-point Likert Type scale ranging from 0 (*disagree*) to 3 (*agree*).

IASMHS-PO. The Psychological Openness subscale of the Inventory of Attitudes Toward Seeking Mental Health Services (Mackenzie et al., 2004) was designed to assess the degree to which an individual is open to acknowledging the presence of a psychological problem and to seek care for such a problem. An example item is "There are certain problems which should not be discussed outside of one's immediate family." Participants respond using a 5-point Likert Type scale ranging from 0 (*disagree*) to 4 (*agree*).

Subjective norms. Congruent with past help-seeking research utilizing the TPB (e.g., Hammer & Vogel, 2013; Hess & Tracey, 2013), we followed the recommendations of Ajzen (2006) for creating TPB-based subjective norms, perceived behavioral control, and intention instruments. Subjective norms was assessed with a 3-item help-seeking subjective norms instrument (e.g., "If I had a mental health concern, it would be expected of me that I seek help from a mental health professional;" rated from [1] *extremely unlikely* to [7] *extremely likely*).

***Perceived behavioral control*.** Perceived behavioral control will be assessed with a 4-item help-seeking perceived behavioral control instrument (e.g., "If I had a mental health concern, if I wanted to I could seek help from a mental health professional;" rated from [1] *definitely false* to [7] *definitely true*).

***Intention*.** Intention was assessed with a 3-item help-seeking intention instrument (e.g., "If I had a mental health concern, I would intend to seek help from a mental health professional" rated from [1] *extremely unlikely* to [7] *extremely likely*).

Analysis Plan. ***Internal structure evidence of validity*.** *Exploratory Factor Analysis (EFA)*. After data cleaning (see Supplemental Material), Study 1's sample was randomly split into exploratory (n = 490) and confirmatory (n = 367) subsamples. A series of exploratory factor analyses (EFA's) on the exploratory subsample was used to confirm that the item set primarily measured the general evaluative help seeking attitudes factor of interest (see Supplemental Material). This is a prerequisite to conducting IRT analyses, given that IRT traditionally assumes the item set is essentially unidimensional (Toland, 2014).

IRT. IRT analysis was used on the exploratory subsample to systematically select an optimal array of items that could properly discriminate among individuals at various levels of the construct (i.e., those with unfavorable vs. neutral vs. favorable attitudes). See Supplemental Material for a description of the advantages of using IRT to select items. A prescriptive and Rasch approach to IRT using Winsteps version 3.92 (Linacre, 2016) was used because the construct and number of items desired was known. Winteps uses joint maximum likelihood estimation (JMLE) to estimate model parameters. JMLE has numerous advantages as an estimation method, including robustness against missing data (Andersen, 1973). Using Winsteps, we examined graded response polytomous models, which included as output item "difficulty" values. While in dichotomous response models this value represents the distribution of correct and incorrect responses, in polytomous models it indicates the ability of an item to differentiate between individuals high or low on the underlying variable. We began with the initial pool of 49 evaluative items, and selected items based on (a) having an infit/outfit score between 0.6 and 1.4,

and (b) dispersing items across difficulty values. We considered six, nine, 12, 15, and 18-item versions of the potential instrument. We also assessed item threshold, which represent the points of intersections among the responses options; disordered thresholds suggest that response options that are not consistent with the intended order of the response scale options.

Confirmatory Factor Analysis (CFA). After the final set of items for the MHSAS were identified through IRT, we sought to determine the stability of the MHSAS' unidimensional factor structure using the confirmatory subsample. Ad hoc EFA (both traditional and bifactor) of the nine items using the confirmatory subsample indicated that no more than one factor could be extracted, obviating the need to test alternative confirmatory measurement models. We conducted a CFA using M*plus*' MLR estimator to protect against deviations from multivariate normality. Model fit was evaluated using the scaled chi-square test statistic (scaled χ^2), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standard Root Mean Square Residual (SRMR). The following fit criteria were used: RMSEA ≤ .06, CFI ≥ .95, TLI ≥ .95, SRMR < .08 for good fit and RMSEA ≤ .10, CFI ≥ .90, TLI ≥ .90, SRMR < .10 for acceptable fit (Weston & Gore, 2006).

Factor determinacy (FD) and construct reliability/replicability (H index). The FD and H index for the MHSAS score were examined in the confirmatory subsample. An FD > .90 would indicate that any observed differences in the help seeking attitudes factor score is indicative of true individual differences on the factor, while a H index > .80 would indicate that the help seeking attitudes latent variable is likely to be replicable across studies and useful in a SEM measurement model (Rodriguez, Reise, & Haviland, 2016a).

Internal consistency. Internal consistency of reliability coefficient (α) estimates for the MHSAS score were calculated separately for the exploratory and confirmatory subsamples.

ME/I. Configural, metric, and scalar ME/I (see Supplemental Material for description) was examined for gender, previous help seeking, and psychological distress using the total Study 1 sample. We used multiple-group CFA to compare a series of nested models, following the Sequential Constraint Imposition approach (Dimitrov, 2010). Given the limitations of using

significant $\Delta\chi 2$ as an indicator of invariance (e.g., sensitivity to sample size), we used the ΔCFI < -.01 and ΔRMSEA < .015 rules for determining model equivalence (Chen, 2007; Cheung & Rensvold, 2002; Shigemoto, Thoen, Robitschek, & Ashton, 2015).

First, to test for configural invariance, a baseline model was specified and tested for adequacy of data fit in each of the two groups separately. Second, the fit of the multiple-group model comparing the two samples was investigated. Third, to test metric invariance, we compared a fully invariant model where each item loads on its corresponding factor to the same degree in both groups to the previous nested configural model (i.e., each item is allowed to freely load on its corresponding factor without the constraint of forced equality across groups). Fourth, to test scalar invariance, we compared a fully invariant model where each item intercept is set to be equal across groups to the previous nested metric model. When scalar invariance is supported, it can be concluded that the MHSAS demonstrates strong ME/I (i.e., invariance of all factor loadings and intercepts) across the two groups.

Evidence regarding relationships with conceptually related constructs. *Convergent evidence of validity***.** Bivariate correlations between the MHSAS score and the scores of the other instruments were examined.

*Incremental evidence of validity***.** To determine whether the MHSAS score accounted for incremental variance in intention beyond other instruments, a hierarchical linear regression was conducted in which the ATSPPH-SF and IASMHS-PO were entered at Step 1, the MHSAS score was entered at Step 2, and intention was entered as the criterion variable.

Known-group evidence of validity. Independent sample *t*-tests determined whether the MHSAS score is sensitive to known differences in attitudes among key groups (i.e., men vs. women, those who have previously sought mental health services vs. those who have not).

Results and Discussion

Internal structure evidence of validity. *EFA*. The results pertaining to the EFA's are described in the Supplemental Material.

IRT. The six-item version had good fit but less than desirable test information function of less than 5. The nine-item version of the MHSAS had good fit and an approximate test information function of 8.5. The 12-item version had an approximate test information function of 12.5, but fit began to weaken; fit continued to drop for the 15- and 18-item versions. The nine-item MHSAS had infit/outfit statistics that remained between 0.6 and 1.4. Difficulties were appropriately dispersed across the nine items (from -1.00 to 0.97), suggesting that the nine items encompass a wide range of the help seeking attitudes construct. In addition, none of the nine items displayed disordered thresholds. Therefore, the nine-item version of the MHSAS was chosen for subsequent use.

CFA. The unidimensional solution demonstrated good fit (scaled χ^2 [27] = 54.02, p < .001, RMSEA = .052 [90% CI of .032, .072], CFI = .976, TLI = .969, SRMR = .025). All items loaded significantly on the help seeking attitudes factor (β's > .74, .86 < B's < 1.15, p's < .001), supporting H16.

FD and H index. The MHSAS score's FD (.97) and H index (.94) scores were above the recommended thresholds.

Internal consistency. The MHSAS score demonstrated internal consistency in both the exploratory (α = .93) and confirmatory (α = .94) subsamples, supporting H14.

ME/I. *Gender*. Each gender's model fit the data (men: scaled χ2 [27, n = 257] = 40.24, p < .001; RMSEA = .044 [90% CI of .004, .070]; CFI = .988; TLI = .984; SRMR = .022; women = scaled χ2 (27, n = 592) = 66.22, p < .001; RMSEA = .050 [90% CI of .035, .065]; CFI = .977; TLI = .969; SRMR = .026), with all item factor loadings > .68 (p < .001) for each sample. Furthermore, the multiple-group model comparing the male and female samples showed an overall acceptable fit to the data (see Table 2 for all ME/I comparative model fit results). Thus, configural invariance was supported. Metric invariance results indicated that the MHSAS was fully metric invariant across men and women per the ΔCFI < -.01 and ΔRMSEA < .015 rules. Likewise, results indicated that the MHSAS was fully scalar invariant across men and women

per the ΔCFI < -.01 and ΔRMSEA < .015 rules. In summary, we can conclude that the MHSAS demonstrated strong ME/I across men and women, supporting H11.

Previous Help Seeking. Each group's model fit the data (previous help seeking: scaled $\chi 2$ [27, n = 623] = 87.60, p < .001; RMSEA = .060 [90% CI of .046, .074]; CFI = .971; TLI = .961; SRMR = .024; no previous help seeking = scaled $\chi 2$ (27, n = 234) = 38.82, p < .001; RMSEA = .043 [90% CI of .000, .072]; CFI = .985; TLI = .980; SRMR = .027), with all item factor loadings > .70 (p < .001) for each sample. The multiple-group model comparing the samples showed an overall acceptable fit to the data (see Table 2). Metric and scalar invariance testing results indicated that the MHSAS was fully metric and scalar invariant across the two groups. Thus, the MHSAS demonstrated strong ME/I across those who have (not) sought previous help, supporting H12.

Psychological Distress. Each group's model fit the data (below cutoff: scaled $\chi 2$ [27, n = 387] = 48.79, p < .001; RMSEA = .046 [90% CI of .024, .066]; CFI = .982; TLI = .975; SRMR = .024; above cutoff = scaled $\chi 2$ (27, n = 470) = 75.82, p < .001; RMSEA = .062 [90% CI of .046, .079]; CFI = .972; TLI = .962; SRMR = .025), with all item factor loadings > .70 (p < .001) for each sample. The multiple-group model comparing the samples showed an overall acceptable fit to the data (see Table 2). Metric and scalar invariance testing results indicated that the MHSAS was fully metric and scalar invariant across the two groups. Thus, the MHSAS demonstrated strong ME/I across those with low versus moderate to high psychological distress, supporting H13.

Evidence Regarding Relationships with Conceptually Related Constructs.

Convergent evidence of validity. Examination of the bivariate correlations in Table 1 reveals that H1, H2, H3, H5, H6, H8, H9, and H10 were supported by the data. Specifically, the MHSAS score correlated, in both anticipated direction and strength, with all criterion variables.

Incremental evidence of validity. In Step 1 of the hierarchical linear regression, ATSPPH-SF (β = .71, p < .001) and IASMHS-PO (β = -.02, p = .69) explained 48% of the variance in intention. In Step 2, the MHSAS (β = .39, p < .001) explained an additional 9% of

the variance in intention ($\Delta R^2 = .085$, $p < .001$). Therefore, H17 was supported by the data. When this regression was re-run without including ATSPPH-SF (MHSAS $\Delta R^2 = .085$) or IASMHS-PO (MHSAS $\Delta R^2 = .24$), the hypothesized effect remained.

Known-group evidence of validity. Participants who have previously sought mental health services ($M = 5.76$, $SD = 1.00$) reported more positive attitudes than those who had never sought mental health services ($M = 5.48$, $SD = 1.03$), $t(855) = 3.60$, $p < .001$, $d = .28$. Likewise, women ($M = 5.77$, $SD = .95$) reported more positive attitudes than men ($M = 5.51$, $SD = 1.12$), $t(847) = 3.23$, $p < .001$, $d = .25$. Thus, H4 and H7 were supported by the data.

Study 2

Study 2 examined the stability of the unidimensional factor structure and the three-week test-retest reliability (i.e., temporal stability) of the MHSAS score.

Method

Procedure. Time 1 data from the Study 2 sample was recruited in the same manner as the Study 1 sample, with a few exceptions. First, the Time 1 survey contained only the final nine-item version of the MHSAS, previous help seeking, and demographic items. Second, inclusion criteria restricted participation to adults who self-identified as currently experiencing a mental health concern (e.g., anxiety, depression). Third, participants were contacted via email three weeks after completing the Time 1 survey and invited to complete the Time 2 survey, which contained the final nine-item version of the MHSAS. Median survey completion time was 3.5 minutes for Time 1 and 1.1 minutes for Time 2.

Participants. Participants were 285 (57 men, 225 women, 3 other gender identity) community adults. The participants ranged in age from 18 to 76 (M = 39.97, *SD* = 14.78). Approximately 80% of the sample identified as White, 7% as Multiracial, 4% as African-American/Black, 4% as Latino/a, 2% as Asian American/Pacific Islander, 2% Other, 1% Native American/Alaskan Native, and 1% preferred not to answer. Approximately 1% reported having less than a high school diploma, 5% earned a high school diploma or GED, 9% earned an associate's degree or attended vocational school, 20% had some college experience, 32% earned

a bachelor's degree, and 33% earned a graduate or professional degree. Approximately 6% of participants reported they had not ever sought help from a mental health professional, while 93% reported they had sought help from a mental health professional in the past. Seventy three percent of Time 1 participants (n = 207) provided data at Time 2.

Measures. The Study 1 Measures section (see above) describes the nature of the MHSAS, previous help seeking, and demographic items administered in Study 2.

Analysis Plan. After data cleaning (see Supplemental Material), we sought to further examine the stability of the MHSAS' unidimensional factor structure using Study 2 data from Time 1. We also used this data to calculate FD, H index, and internal consistency of reliability coefficient (α) estimates. For all analyses, we followed the same procedures as used in Study 1.

Test-retest reliability was examined in three ways. First, the bivariate correlation between respondents' Time 1 and Time 2 MHSAS scores was calculated. Second, the average measures intraclass correlation coefficient (2,k) between the Time 1 and Time 2 MHSAS scores was computed. Third, a paired-samples t-test was conducted to determine if respondents' MHSAS scores demonstrated a mean difference between Time 1 and Time 2.

Results and Discussion

The unidimensional solution demonstrated adequate fit (scaled χ^2 [27] = 59.31, $p < .001$, RMSEA = .065 [90% CI of .042, .087], CFI = .959, TLI = .945, SRMR = .032). All items loaded significantly on the help seeking attitudes factor (β's > .61, .79 < B's < 1.26, p's < .001), further supporting H16. The MHSAS score's FD (.96), H index (.93), and internal consistency (α = .92) estimates were above the recommended thresholds, further supporting H14.

Test-Retest Reliability. The bivariate correlation between the Time 1 and Time 2 scores was .76 and the intraclass correlation coefficient (2,k) for the two scores was .86. The paired-samples t-test suggested the absence of a significant mean difference, t(207) = .147, p = .88, between participants' Time 1 (M = 5.68, SD = 1.12) and Time 2 (M = 5.66, SD = 1.07) scores. Thus, all three test-retest reliability analyses supported H15.

General Discussion

The present investigation sought to develop and evaluate the nine-item Mental Help Seeking Attitudes Scale (MHSAS), designed to measure a respondents' overall evaluation (unfavorable vs. favorable) of their seeking help from a mental health professional.

Initial Evidence of Reliability and Validity for the MHSAS Score

Regarding internal structure evidence of validity, IRT analysis was used to select nine items that discriminate among individuals at different levels (i.e., unfavorable, neutral, or favorable attitudes) of the construct. A unidimensional measurement model accurately reproduced the observed covariation among the nine MHSAS items in both Study 1 and Study 2, providing initial support for the structural generalizability of the MHSAS. Furthermore, results from the ME/I analyses indicated that the MHSAS demonstrated strong ME/I (i.e., invariance of all factor loadings and intercepts) across gender, past help seeking experience, and psychological distress. This suggests that, for these key help seeking groups, the MHSAS had a similar theoretical structure and meaning, relations between the MHSAS and external variables could validly be compared, and mean differences in the MHSAS score could validly be compared.

Regarding reliability, the MHSAS items demonstrated internal consistency across both studies, and temporal stability over three weeks. Test content evidence of validity was also provided via feedback from experts and community adults, who rated the MHSAS instructions and items as sufficiently clear, relevant, and representative of the intended construct.

Validity evidence regarding relationships with conceptually related constructs was also presented. Convergent evidence of validity was demonstrated when the MHSAS score demonstrated the hypothesized relationships with the following variables: subjective norms, perceived behavioral control, intention, public stigma, self-stigma, anticipated risks and benefits, and the ATSPPH-SF and IASMHS-PO scores. Incremental evidence of validity was demonstrated when the MHSAS demonstrated the ability to account for unique variance, beyond that accounted for by the ATSPPH-SF and IASMHS-PO, in intention to seek help. Finally, known-group evidence of validity was provided when women and those who had previously

sought mental health services were shown to report more favorable attitudes than did their demographic counterparts.

Addressing Limitations through Future Research

Instrument validation is an ongoing process (AERA et al., 2014). There are several issues that future research must address to advance our understanding of the utility and limitations of the MHSAS. First, our findings are tied to the nature of our sample of community adults living in the USA. While one-third of the participants in our sample identified as people of color, our cell sizes for each racial/ethnic group were not large enough to allow properly-powered ME/I analyses. These analyses are an important next step to verifying the conceptual and measurement equivalence of the attitudes construct, as operationalized by the MHSAS. Until these analyses are conducted, the appropriateness of using the MHSAS to make racial/ethnic comparisons remains an open question. For example, it is possible that the terminology of disempowerment versus empowerment used in item 7 may be unfamiliar to adults from certain cultural groups, or that people living below the poverty line with lack of access to mental healthcare may consider such items irrelevant to their socioeconomic situation. Similarly, our sample was generally well-educated, which makes testing the MHSAS with people with less education another priority. The average participant also had favorable attitudes toward seeking help and past experience with seeking mental health services, which can exert an influence on the psychometric performance of the MHSAS and its items (e.g., negative skew can reduce internal consistency). The fact that 35% of the sample was over the age of 50 may partially account for these more positive attitudes (see Mackenzie, Scott, Mather, and Sareen, 2008, as could selection bias (i.e., because the study was advertised as a therapy questionnaire, participants interested in the topic may have been more likely to complete the survey). Examining the performance of the MHSAS among those with unfavorable attitudes toward seeking mental health services and those with little exposure to mental healthcare is another important next step. Those who wish to use the MHSAS with international populations or adapt it into another language should seek to establish the ME/I and cross-cultural validity of the

instrument prior to making firm conclusions based on MHSAS data. Second, test-criterion evidence of validity would provide additional support for the utility of the MHSAS. Specifically, comparing the ability of the MHSAS, ATSPPH-SF, and IASMHS-PO to predict future help seeking behavior would provide important information about the degree to which these instruments accurately measure help-seeking attitudes. Our research team intends to begin addressing these current limitations through upcoming research and invites other researchers to collaboratively and independently do the same. Verification of the stability of the present findings in similar and different samples is also of interest, as this would facilitate further confidence in the utility of the MHSAS.

Implications for Prevention, Practice, and Research

Help seeking attitudes is a central construct in the help seeking literature, given its influence on treatment seeking behavior via intention to seek help (Freyer et al., 2007). Previous research has heavily utilized instruments such as the ATSSPH-SF and IASMHS to understand how help seeking attitudes differ across groups and relate to other key constructs such as gender, acculturation/enculturation, stigma, and social support (Nam et al., 2010; Nam et al., 2013; Sun et al., 2016). However, we argued in the present paper that the psychometric limitations of these popular instruments necessitated the development the MHSAS. The MHSAS differs from the ATSPPH-SF and IASMHS-PO in that it (a) uses items less likely to measure construct-irrelevant variance or underrepresent the construct of interest, (b) demonstrated a cleaner and more stable unidimensional structure in initial testing, (c) achieved stronger internal consistency in initial testing, and (d) demonstrated initial evidence of ME/I across key groups (see Supplemental Material for results of a psychometric comparison of the three instruments). Therefore, we believe the MHSAS has the potential to increase the accuracy of help seeking attitudes measurement in the context of prevention, practice, and research.

The MHSAS could potentially be used in a variety of applications such as assessing the (a) population-level impact of prevention efforts like the National Institute of Mental Health's *Real Men, Real Depression* campaign (Rochlen, Whilde, & Hoyer, 2005), (b) efficacy of a self-

affirmation intervention (Lannin, Guyll, Vogel, & Madon, 2013), or (c) openness of new or potential clients to behavioral health treatment in the context of an integrated healthcare clinic. Counseling professionals may find it helpful to administer the MHSAS to new clients to assess clients' attitudes toward counseling, which can help professionals better anticipate client reservations regarding talk therapy. Furthermore, the MHSAS may help researchers study the relationship between help seeking attitudes and other constructs of interest with more precision and less risk of tautology.

References

Addis, M. E., & Mahalik, J. R. (2003). Men, masculinity, and the contexts of help seeking. *American Psychologist*, *58*(1), 5-14.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognitions to behavior* (11-29). New York, NY: Springer.

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*(1), 27-58.

Ajzen, I. (2006). Constructing a Theory of Planned Behavior questionnaire. Available online: http://people.umass.edu/aizen/pdf/tpb.measurement.pdf

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Andersen, E. B. (1973). Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, *26*(1), 31–44.

Andrykowski, M. A., & Burris, J. L. (2010). Use of formal and informal mental health resources by cancer survivors: Differences between rural and nonrural survivors and a preliminary test of the theory of planned behavior. *Psycho-Oncology*, *19*(11), 1148-1155.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471-499.

Bayer, J. K., & Peay, M. Y. (1997). Predicting intentions to seek help from professional mental health services. *Australian and New Zealand Journal of Psychiatry*, *31*(4), 504-513.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (Vol. 17). Quantitative applications in the social sciences. Beverly Hills: Sage Publications.

Chang, H. (2007). Depressive symptom manifestation and help-seeking among Chinese college students in Taiwan. *International Journal of Psychology, 42*(3), 200-206.

Cheung, G. W., & Lau, R. S. (2012). A direct comparison approach for testing measurement invariance. *Organizational Research Methods, 15*(2), 167-198.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255.

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

Cho, S. J., Wilmer, J., Herzmann, G., McGugin, R. W., Fiset, D., Van Gulick, A. E., ... & Gauthier, I. (2015). Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychological Assessment, 27*(2), 552-566.

Chong, S. S. (2016). *Help-seeking models for Asian international and American students* (Unpublished doctoral dissertation). Arizona State University, Tempe, Arizona.

Codd, R. T., & Cohen, B. N. (2003). Predicting college student intention to seek help for alcohol abuse. *Journal of Social and Clinical Psychology, 22*(2), 168-191.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage publications.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43*(2), 121-149.

Dueber, D. M. (2016). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. Retrieved from http://sites.education.uky.edu/apslab/resources

Dunford, E., & Granger, C. (2017). Maternal guilt and shame: Relationship to postnatal depression and attitudes towards help-seeking. *Journal of Child and Family Studies, 26*(6), 1692-1701.

Drapeau, C. W., Cerel, J., & Moore, M. (2016). How personality, coping styles, and perceived closeness influence help-seeking attitudes in suicide-bereaved adults. *Death Studies, 40*(3), 165-171.

Elhai, J. D., Schweinle, W., & Anderson, S. M. (2008). Reliability and validity of the Attitudes Toward Seeking Professional Psychological Help Scale-Short Form. *Psychiatry Research, 159*(3), 320-329.

Fischer, E. H., & Farina, A. (1995). Attitudes toward seeking professional psychological help: A shortened form and considerations for research. *Journal of College Student Development, 36*(4), 368-373.

Fischer, E. H., & Turner, J. I. (1970). Orientations to seeking professional help: Development and research utility of an attitude scale. *Journal of Consulting and Clinical Psychology, 35*, 79-90.

Freyer, J., Coder, B., Bischof, G., Baumeister, S. E., Rumpf, H. J., John, U., & Hapke, U. (2007). Intention to utilize formal help in a sample with alcohol problems: A prospective study. *Drug and Alcohol Dependence, 87*(2), 210-216.

Gulliver, A., Griffiths, K. M., Christensen, H., & Brewer, J. L. (2012). A systematic review of help-seeking interventions for depression, anxiety and general psychological distress. *BMC Psychiatry, 12*, 81.

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry, 10*(1), 113.

Hammer, J. H., & Vogel, D. L. (2010). Men's help seeking for depression: The efficacy of a male-sensitive brochure about counseling. *The Counseling Psychologist, 38*(2), 296-313.

Hammer, J. H., & Vogel, D. L. (2013). Assessing the utility of the willingness/prototype model in predicting help-seeking decisions. *Journal of Counseling Psychology, 60*, 83-97.

Hammer, J. H., & Vogel, D. L. (2017). Development of the Help-Seeker Stereotype Scale. *Stigma and Health, 2*, 121-136.

Heise, D. (1970). The semantic differential and attitude research. In G. F. Summer. Attitude Measurement (235-253). Chicago, IL: Rand McNally.

Hess, T. R., & Tracey, T. J. (2013). Psychological help-seeking intention among college students across three problem areas. *Journal of Counseling & Development, 91*, 321-330.

Hyland, P., Boduszek, D., Dhingra, K., Shevlin, M., Maguire, R., & Morley, K. (2015). A test of the inventory of attitudes towards seeking mental health services. *British Journal of Guidance & Counselling*, *43*(4), 397-412.

Kaiser, F. G., & Gutscher, H. (2003). The proposition of a general version of the theory of planned behavior: Predicting ecological behavior. *Journal of Applied Social Psychology*, *33*(3), 586-603.

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., ... & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalence and trends in non-specific psychological distress. *Psychological Medicine*, *32*(06), 959-976.

Komiya, N., Good, G. E., & Sherrod, N. B. (2000). Emotional openness as a predictor of college students' attitudes toward seeking psychological help. *Journal of Counseling Psychology, 47,* 138-143.

Lannin, D. G., Guyll, M., Vogel, D. L., & Madon, S. (2013). Reducing the stigma associated with seeking psychotherapy through self-affirmation. *Journal of Counseling Psychology, 60*(4), 508-519.

Lannin, D. G., Vogel, D. L., Brenner, R. E., & Tucker, J. R. (2015). Predicting self-esteem and intentions to seek counseling: The internalized stigma model. *The Counseling Psychologist*, *43*(1), 64-93.

Leach, M. M., Jana-Masri, A., & Priester, P. E. (2009). Attitudes Toward Seeking Professional Psychological Help Scale in Egypt: Its factor structure and relationship to Islamic beliefs and behaviors. *Research in the Social Scientific Study of Religion, 20,* 281-296.

Li, W., Dorstyn, D. S., & Denson, L. A. (2014). Psychosocial correlates of college students' help-seeking intention: A meta-analysis. *Professional Psychology: Research and Practice*, *45*(3), 163-170.

Linacre, J. M. (2016). *Winsteps® (Version 3.92.0).* Beaverton, OR: Winsteps.com. Retrieved from http://www.winsteps.com/

Mackenzie, C. S., Gekoski, W. L., & Knox, V. J. (2006). Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. *Aging and Mental Health*, *10*(6), 574-582.

Mackenzie, C. S., Knox, V. J., Gekoski, W. L., & Macaulay, H. L. (2004). An adaptation and extension of the attitudes toward seeking professional psychological help scale. *Journal of Applied Social Psychology, 34*(11), 2410-2435.

Mackenzie, C. S., Scott, T., Mather, A., Sareen, J. (2009). Older adults' help-seeking attitudes and treatment beliefs concerning mental health problems. *American Journal of Geriatric Psychiatry, 16*, 1010-1019.

Mahalik, J. R., Lagan, H. D., & Morrison, J. A. (2006). Health behaviors and masculinity in Kenyan and U.S. male college students. *Psychology of Men & Masculinity*, *7*(4), 191-202.

Mak, H., & Davis, J. (2014). The application of the theory of planned behavior to help-seeking intention in a Chinese society. *Social Psychiatry & Psychiatric Epidemiology*, *49*(9), 1501-1515.

Masuda, A., Anderson, P. L., & Edmonds, J. (2012). Help-seeking attitudes, mental health stigma, and self-concealment among African American college students. *Journal of Black Studies, 43*(7), 773-786.

Miller, M. J., & Sheu, H. (2008). Conceptual and measurement issues in multicultural psychology research. In S. D. Brown & R. W Lent (Eds.), *Handbook of counseling psychology* (103-120). Hoboken, NJ: John Wiley & Sons.

Mojaverian, T., Hashimoto, T., & Kim, H. S. (2013). Cultural differences in professional help seeking: A comparison of Japan and the US. *Frontiers in Psychology*, *3*, 615.

Mo, P. H., & Mak, W. S. (2009). Help-seeking for mental health problems among Chinese: The application and extension of the theory of planned behavior. *Social Psychiatry and Psychiatric Epidemiology*, *44*(8), 675-684.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide (7th ed.).* Los Angeles, CA: Authors.

Nam, S. K., Choi, S. I., Lee, J. H., Lee, M. K., Kim, A. R., & Lee, S. M. (2013). Psychological factors in college students' attitudes toward seeking professional psychological help: A meta-analysis. *Professional Psychology: Research and Practice*, *44*(1), 37-45.

Nam, S. K., Chu, H. J., Lee, M. K., Lee, J. H., Kim, N., & Lee, S. M. (2010). A meta-analysis of gender differences in attitudes toward seeking professional psychological help. *Journal of American College Health*, *59*(2), 110-116.

Nam, S. K., & Lee, S. M. (2015). The role of attachment and stigma in the relationship between stress and attitudes toward counseling in South Korea. *Journal of Counseling & Development*, *93*(2), 212-224.

Nigg, C. R., Lippke, S., & Maddock, J. E. (2009). Factorial invariance of the theory of planned behavior applied to physical activity across gender, age, and ethnic groups. *Psychology of Sport and Exercise*, *10*(2), 219-225.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

Parent, M. C. (2013). Handling item-level missing data: Simpler is just as good. *The Counseling Psychologist*, *41*, 568–600.

Pederson, E. L., & Vogel, D. L. (2007). Male gender role conflict and willingness to seek counseling: Testing a mediation model on college-aged men. *Journal of Counseling Psychology*, *54*(4), 373-384.

Picco, L., Abdin, E., Chong, S. A., Pang, S., Shafie, S., Chua, B. Y., & ... Subramaniam, M. (2016). Attitudes toward seeking professional psychological help: Factor structure and socio-demographic predictors. *Frontiers in Psychology*, *7*.

Prochaska, J. J., Sung, H. Y., Max, W., Shi, Y., & Ong, M. (2012). Validity study of the K6 scale as a measure of moderate mental distress based on mental health treatment need and utilization. *International Journal of Methods in Psychiatric Research*, *21*(2), 88-97.

Rochlen, A. B., Whilde, M. R., & Hoyer, W. D. (2005). The Real Men. Real Depression Campaign: Overview, theoretical implications, and research considerations. *Psychology of Men & Masculinity*, *6*(3), 186-194.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137-150.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98(*3*),* 223-237.

Schomerus, G., Matschinger, H., & Angermeyer, M. C. (2009). The stigma of psychiatric treatment and help-seeking intentions for depression. *European Archives of Psychiatry and Clinical Neuroscience*, *259*(5), 298-306.

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*(1), 80-110.

Shaffer, P. A., Vogel, D. L., & Wei, M. (2006). The mediating roles of anticipated risks, anticipated benefits, and attitudes on the decision to seek professional help: An attachment perspective. *Journal of Counseling Psychology*, *53*(4), 442-452.

Shigemoto, Y., Thoen, M. A., Robitschek, C., & Ashton, M. W. (2015). Assessing measurement invariance of the Personal Growth Initiative Scale-II among Hispanics, African Americans, and European Americans. *Journal of Counseling Psychology*, *62*(3), 537-544.

Sun, S., Hoyt, W. T., Brockberg, D., Lam, J., & Tiwari, D. (2016). Acculturation and enculturation as predictors of psychological help-seeking attitudes (HSAs) among racial and ethnic minorities: A meta-analytic investigation. *Journal of Counseling Psychology*. Advance online publication.

Toland, M. D. (2014). Practical guide to conducting an item response theory analysis. *The Journal of Early Adolescence, 34*(1), 120-151.

Vogel, D. L., & Heath, P. (2015). Men, masculinities, and help-seeking patterns. in S. R. Wester & J. Wong (Eds.), *APA handbook for the psychology of men and masculinities*. Washington, DC: American Psychological Association.

Vogel, D. L., Wade, N. G., & Haake, S. (2006). Measuring the self-stigma associated with seeking psychological help. *Journal of Counseling Psychology, 53*(3), 325-337.

Vogel, D. L., & Wester, S. R. (2003). To seek help or not to seek help: The risks of self-disclosure. *Journal of Counseling Psychology, 50*, 351-361.

Vogel, D. L., Wester, S. R., Wei, M., & Boysen, G. A. (2005). The role of outcome expectations and attitudes on decisions to seek professional help. *Journal of Counseling Psychology, 52*(4), 459-470.

Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E. J., & ... Wells, J. E. (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet, 370*(9590), 841-850.

Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*, 719-751.

Wong, Y. J., Ho, M. R., Wang, S., & Miller, I. K. (2017). Meta-analyses of the relationship between conformity to masculine norms and mental health-related outcomes. *Journal of Counseling Psychology, 64*(1), 80-93.

Table 1

Bivariate Correlations, Means, Standard Deviations, and Internal Consistency Estimates for all Instruments used in Study 1

	MHSAS	SN	PBC	PS	SS	AR	AU	ATSPPH-SF	IASMHS-PO	INT	M	SD	α (95% CI)
MHSAS											5.69	1.02	.93 (.925, .938)
SN	**.41**										5.52	1.30	.81 (.791, .835)
PBC	**.44**	**.38**									6.04	0.99	.78 (.756, .804)
PS	**-.47**	**-.39**	**-.34**								2.26	0.61	.80 (.782, .824)
SS	**-.60**	**-.30**	**-.38**	**.61**							2.25	0.78	.90 (.885, .906)
AR	**-.40**	**-.25**	**-.32**	**.48**	**.58**						3.00	1.12	.86 (.844, .875)
AU	**.62**	**.27**	**.34**	**-.34**	**-.55**	**-.31**					3.76	0.94	.84 (.821, .856)
ATSPPH-SF	**.67**	**.37**	**.42**	**-.47**	**-.67**	**-.41**	**.66**				2.56	0.59	.86 (.848, .876)
IASMHS-PO	**.50**	**.31**	**.35**	**-.50**	**-.62**	**-.49**	**.52**	**.74**			2.73	0.86	.83 (.815, .849)
INT	**.68**	**.51**	**.51**	**-.45**	**-.59**	**-.42**	**.54**	**.69**	**.51**		5.50	1.49	.95 (.949, .959)
K6	**-.12**	**-.11**	**-.19**	**.23**	**.17**	**.22**	-.03	.00	**-.13**	-.09	6.49	5.24	.89 (.879, .902)

Note. Study 1 *N* = 857. MHSAS = Mental Help Seeking Attitudes Scale; SN = Subjective Norms, PBC = Perceived Behavioral Control; PS = Public Stigma of Seeking Help, SS = Self-Stigma of Seeking Help; AR = Anticipated Risk; AU = Anticipated Utility; ATSPPH-SF = Attitudes Toward Seeking Professional Psychological Help - Short Form Scale; IASMHS-PO = Psychological Openness subscale of the Inventory of Attitudes Toward Seeking Mental Health Services; INT = Intention; K6 = Kessler Psychological Distress Scale. Bold indicates significance at $p < .05$.

Table 2

Measurement Equivalence/Invariance of the Mental Help Seeking Attitudes Scale

	Scaled $\chi 2$	*df*	RMSEA	ΔRMSEA	CFI	ΔCFI	Δdf	$\Delta\chi 2$	Model comparison
Men vs. Women									
Configural	107.26	54	.048		.981				
Metric	111.98	62	.044	-.004	.982	.001	8	3.74	Configural
Scalar	127.03	70	.044	.000	.980	-.002	8	15.28	Metric
Those Who Have Previously Sought Help from a Mental Health Professional vs. Those Who Have Not									
Configural	128.33	54	.057		.974				
Metric	140.15	62	.054	-.003	.973	-.001	8	12.22	Configural
Scalar	154.84	70	.053	-.001	.970	-.003	8	13.19	Metric
Those Reporting Low vs. Moderate-to-High Psychological Distress									
Configural	122.41	54	.054		.976				
Metric	142.53	62	.055	.001	.972	-.004	8	**20.11**	Configural
Scalar	160.44	70	.055	.000	.968	-.004	8	**17.68**	Metric

Note. RMSEA = root-mean-square error of approximation; CFI = comparative fit index. Bold indicates significant chi-square difference.