

Dimensionality, Reliability, and Predictive Evidence of Validity for Three Help Seeking Intention

Instruments: ISCI, GHSQ, and MHSIS

[Brief Report]

Note: This article may not exactly replicate the final version published in the APA journal. It is not the copy of record. Please use the DOI link on my website to access the PDF through your institution, allowing full access to the published type-set article.

To obtain a copy of the Mental Help-Seeking Intention Scale (MHSIS), please visit:
<http://DrJosephHammer.com>

Current APA-Style Citation:

Hammer, J. H., & Spiker, D. A. (in press). Dimensionality, Reliability, and Predictive Evidence of Validity for Three Help Seeking Intention Instruments: ISCI, GHSQ, and MHSIS. *Journal of Counseling Psychology*.

Joseph H. Hammer, Ph.D.
243 Dickey Hall
Department of Educational, School, and Counseling Psychology
University of Kentucky
Lexington, KY 40506
(847) 809-8785
jhammer@gmail.com

Douglas A. Spiker, M.A.
251 Dickey Hall
Department of Educational, School, and Counseling Psychology
University of Kentucky
Lexington, KY 40506
(859) 257-4158

Abstract

People's intention to seek help from a mental health professional is thought to be the proximal cause of help seeking behavior and thus is a dependent variable frequently measured by help seeking researchers. Using a research design that accounted for actual future help seeking behavior, the present study documented the dimensionality, internal consistency, and predictive evidence of validity of three intention instruments: the Intentions to Seek Counseling Inventory (ISCI), General Help Seeking Questionnaire (GHSQ), and Mental Help-Seeking Intention Scale (MHSIS). The sample was composed of 405 community-dwelling adults who self-identified as currently experiencing a mental health concern. Results provided support for the ISCI's three-factor structure and the internal consistency of its three subscale scores. In contrast, the GHSQ did not demonstrate clear evidence of adequate measurement model fit or internal consistency in the present sample. Results also tentatively suggested that the three-item MHSIS is a unidimensional instrument that produces an internally consistent total score with appropriate construct replicability. The ability of these instruments to predict who would seek help from a mental health professional in the next three months was also examined. The MHSIS demonstrated the strongest evidence of predictive validity (about 70% of participants were correctly classified), followed by the GHSQ and ISCI.

Keywords: help seeking; intention; validity; reliability; factor analysis

Public Significance Statement: This study compared the reliability and validity of different self-report instruments designed to measure people's intention to seek mental health services. Measuring help seeking intention accurately is important because invalid instruments can lead to invalid research conclusions that, in turn, can lead to misguided clinical and policy decisions.

Dimensionality, Reliability, and Predictive Evidence of Validity for Three Help Seeking Intention Instruments: ISCI, GHSQ, and MHSIS

Most people who need psychotherapy do not receive it (Wang et al., 2007). Historically, counseling psychologists have drawn upon psychological theory to better understand the factors that create and maintain this treatment gap (e.g., Gourash, 1978). The theory of planned behavior (TPB; Ajzen, 1985) and its predecessor (theory of reasoned action; Ajzen & Fishbein, 1980) are used by counseling psychologists to understand factors that influence help seeking for mental health concerns. The TPB posits that when people have a sufficient degree of actual control over a behavior, their intention (i.e., motivation to exert effort to perform a behavior) drives them to seek treatment (Ajzen, 2006). Measuring actual help-seeking behavior requires substantial research resources to operationalize via longitudinal and experimental design. Thus, researchers often choose to measure help seeking intention, the closest measurable proxy of actual help-seeking behavior, via a self-report intention instrument. However, there is a lack of clear evidence regarding the dimensionality (i.e., factor structure), reliability, and validity of three popular intention instruments, the Intentions to Seek Counseling Inventory (ISCI; Cash, Begley, McCown, & Weise, 1975; Cepeda-Benito & Short, 1998), General Help-Seeking Questionnaire (GHSQ; Wilson, Deane, Ciarrochi, & Rickwood, 2005), and Mental Help-Seeking Intention Scale (MHSIS; Hammer & Vogel, 2013; Hess & Tracey, 2013, Mo & Mak, 2009). Therefore, using a research design that accounted for actual future mental health help seeking behavior, the present study compared the psychometric properties of these three instruments.

Current Psychometric Evidence for Existing Intention Instruments

Historically, the ISCI is the most commonly used intention instrument. The ISCI items were originally developed by Cash et al. (1975) to measure respondents' expectancies of a counselor's helpfulness for 15 personal problems (e.g., depression, insomnia). Later publications in the *Journal of Counseling Psychology* adapted the ISCI into a 17-item instrument that asked respondents to rate how likely they would be to seek counseling if they were experiencing each of the 17 problems (e.g., Kelly & Achter, 1995). Cepeda-Benito and Short (1998) first

investigated the ISCI's dimensionality, concluding that the ISCI was composed of three correlated factors: Psychological and Interpersonal Concerns (10 items), Academic Concerns (4 items), and Drug Use Concerns (2 items). This version of the ISCI has been used frequently in subsequent help seeking research (e.g., Demyan & Anderson, 2012; Vogel, Wester, Wei, & Boysen, 2005), though using only one or two of the ISCI subscales is common (e.g., Nam & Lee, 2015; Pederson & Vogel, 2007). However, Pheko, Chilisa, Balogun, & Kgathi (2013) documented that the ISCI did not conform to the three correlated factors model.

While many researchers have created three subscales scores based on the Cepeda-Benito and Short (1998) dimensional results, others have calculated a single total score for the ISCI (e.g., Kelly & Achter, 1995; Leech, 2007; Pheko et al., 2013), despite a lack of published evidence that the ISCI is unidimensional. Furthermore, ISCI scores have not always demonstrated sufficient reliability (e.g., α 's < .70; Demyan & Anderson, 2012; Vogel et al., 2005). Most important, predictive evidence of validity has never been published for the ISCI. If the ISCI truly measures intention to seek help, then a sensible test of the ISCI scores' validity would be how accurately it predicts future help seeking behavior. In summary, the dimensionality, reliability, and predictive evidence of validity for the ISCI scores requires verification.

The GHSQ was designed to improve upon existing intention instruments such as the ISCI. An 18-item version (Deane, Wilson, & Ciarrochi, 2001) was first developed to assess intention to seek help from six sources (e.g., friend, family, mental health professional, telephone help line, doctor/GP, no one) for three problem types (e.g., personal-emotional, anxiety-depression, and suicidal thoughts). A principle-component analysis suggested retention of the six aforementioned help sources factors, and the collapsing of the personal-emotional and anxiety-depression problem types into a single personal-emotional problem type. Despite concluding that six factors defined the 18 items rather than two factors or one factor, the authors asserted that the GHSQ items for personal-emotional and suicidal problem types could either be scored as two subscales or combined to create a single total score. The GHSQ developers later published a

22-item version (Wilson et al., 2005), which asked respondents to rate the likelihood that they would seek help for two different problem types (i.e., suicidal thoughts, personal-emotional problems) from 10 different help sources (e.g., friend, family). Like the 18-item version, the developers stated that the GHSQ could be scored as one total score or two subscales scores. However, no evidence for a one-factor or two-factor structure was presented (Wilson et al., 2005).

To increase the GHSQ's utility, the developers have encouraged researchers to modify the GHSQ to include only those problem types and help sources relevant for a given study. This has led to several versions of the GHSQ with factor structures ranging from a one-factor structure for a 6-item GHSQ (Wilson & Deane, 2012) to a five-factor structure for a 21-item GHSQ (Wilson, Rickwood, Bushnell, Caputi, & Thomas, 2011). The inconsistency in the construction of the GHSQ across studies has made it difficult to firmly establish the dimensionality of the instrument and the reliability and validity of its scores. The practice of calculating scores without first establishing that the items used to create those scores successfully load on the same factor (e.g., Seward & Harris, 2016; Wilson et al., 2011) has led to insufficient score reliability (e.g., α 's < .66; Hasking, Reese, Martin, & Quigley, 2015; McDermott et al., 2017; Straiton, Hjelmeland, Grimholt, & Dieserud, 2013).

To examine predictive evidence of validity for the GHSQ, Wilson et al. (2005) measured the correlation between intention to seek help from a mental health professional for personal-emotional problems at Time 1 and self-reported help seeking behavior at Time 2, three weeks later. The GHSQ accounted for only 2.9% of the variance in prospective help seeking behavior for personal-emotional problems, which may have been due to its reliance on a single item (per problem type) or due to the brief window between Time 1 and Time 2. In summary, the dimensionality, reliability, and predictive evidence of validity for the GHSQ scores would benefit from further verification.

Like the GHSQ, different versions of the MHSIS have been used by help-seeking researchers (e.g., Hammer & Vogel, 2013; Hess & Tracey, 2013, Mo & Mak, 2009). Each

version was adapted from the three-item intention instrument presented in Ajzen's (2006) "Constructing a TPB Questionnaire" guide. These versions have demonstrated internal consistency (α 's $> .87$) and convergent evidence of validity (e.g., significant positive associations between intention and both attitudes and subjective norms around seeking professional psychological help; Hammer & Vogel, 2013; Hess & Tracey, 2013; Mo & Mak, 2009), but the dimensionality and predictive evidence of validity for the MHSIS score requires investigation. As with the GHSQ, the variation in the construction of the MHSIS across studies has created an opportunity for a formal psychometric evaluation of a standardized version of the MHSIS that could be used in future research. The present study is the first to provide such an evaluation.

The Present Study

A mental help seeking intention instrument with clear dimensionality, reliability, and strong predictive evidence of validity could improve the quality and consistency of future help seeking scholarship. Our review of the literature suggests that no intention instrument has yet demonstrated sufficient evidence of these three psychometric qualities. Therefore, the present study sought to answer three questions about the ISCI, GHSQ, and MHSIS using a community-dwelling sample of adults who self-identified as currently dealing with a mental health concern.

First, what dimensionality (i.e., factor structure) does each instrument display? Dimensionality determines how an instrument should be conceptualized and scored, which has implications for testing reliability and validity (DeVellis, 2012). Second, how internally consistent are the score(s) of each instrument? Lack of internal consistency prevents researchers from being able to detect the true degree of association between help seeking intention and other constructs of interest (Osborne, 2003). Third, how accurately does each instrument's score(s) predict subsequent, actual help seeking behavior? Accurate prediction of prospective behavior constitutes predictive evidence of validity (*Standards for Educational and Psychological Testing*; American Education Research Association [AERA], American Psychological Association, & National Council on Measurement in Education, 2014). The present findings may help

counseling psychologists and other stakeholders make informed decisions about the utility of these instruments for use in future research and practice.

Method

Participants and Procedures

Participants were recruited via ResearchMatch (RM), 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 (CTSA) program. RM 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's Office of Research Integrity. RM participants were contacted via the registry system regarding the study, advertised as a survey about factors influencing mental health help seeking. Interested participants were directed to an online Time 1 survey that began with an informed consent page, continued with the survey items, and ended with a conclusion page. The survey items page offered this clarification: "For the purposes of this survey, the term "mental health concern" refers to a reason one might visit a mental health professional, ranging from personal difficulties (e.g., loss of a loved one) to mental illness (e.g., anxiety, depression)." Time 1 participants were contacted via email three months later and invited to complete the brief Time 2 survey, which asked about their actual help seeking behavior since Time 1. Participants had the option of entering a drawing for one of several \$25 Amazon.com gift cards.

Time 1 participants were 405 (60 men, 340 women, 5 other gender identity) community-dwelling adults who answered "yes" when asked if they self-identified as currently experiencing a mental health concern (e.g., depression, anxiety). The participants ranged in age from 19 to 78 ($M = 40.74$, $SD = 14.41$). Approximately 86% of the sample identified as White, 5% as African-American/Black, 3% as Multiracial, 2% as Latino/a, 2% as Asian American/Pacific Islander, and 1% as Other. Approximately 57% reported being in a civil union or in a committed relationship, 26% single, 14% separated or divorced, 2% widowed, and 1% preferred not to answer. 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, 18% had some college experience, 34% earned a bachelor's degree, 34% earned a graduate or professional degree, and 1% preferred not to answer. Approximately 14% of participants reported they had never sought help from a mental health professional, while 86% reported they had sought help from a mental health professional in the past. Regarding U.S state residence, approximately 2% reported living in New England, 10% in Middle Atlantic, 22% in East North Central, 5% in West North Central, 22% in South Atlantic, 14% in East South Central, 4% in West South Central, 5% in Mountain, 15% in Pacific, and 1% reported not residing in U.S. Seventy three percent of Time 1 participants ($n = 294$) provided data at Time 2. Independent sample t-tests and Pearson Chi-Square tests were used to determine whether those who did and those who did not complete the Time 2 survey systematically varied on race, marital status, education, past help seeking, region of residence, ISCI total score, GHSQ total score, or MHSIS total score. Apart from education (i.e., Time 2 completers reported a higher level of education), the two groups did not significantly differ on these variables ($ps > .26$).

Measures

Copies of each instrument used in this study are provided in the Supplemental Material.

Intentions to Seek Counseling Inventory (ISCI; Cash et al., 1975; Cepeda-Benito & Short, 1998). We used the 17-item version of the ISCI to measure respondents' intention to seek counseling if they experienced a variety of specific problems (i.e., weight control, relationship difficulties, concerns about sexuality, depression, conflicts with parents, difficulties dating, difficulty in sleeping, inferiority feelings, difficulties with friends, self-understanding, loneliness, excessive alcohol use, drug problems, speech anxiety, choosing a major, test anxiety, academic work procrastination). Participants rated the likelihood they would seek counseling for each problem, on a five-point Likert scale from 1 (*very unlikely*) to 6 (*very likely*), with higher scores indicating greater intention. Post-hoc convergent evidence of the validity for the ISCI subscale scores (see Introduction) has been presented in the form of significant positive associations with attitudes toward seeking psychotherapy and past-help seeking behavior (Lannin, Vogel, Brenner,

& Tucker, 2014; Vogel, Wade, & Haake, 2006). Cepeda-Benito and Short (1998) provided evidence of internal consistency for the three subscales (α 's > .71).

General Help-Seeking Questionnaire (GHSQ; Wilson et al., 2005; Rickwood, Deane, Wilson, & Ciarrochi, 2005). We used the publicly-available 10-item version of the GHSQ presented in the appendix of Rickwood et al. (2005) to measure participants' intention to seek help for personal-emotional problems. This version includes four informal source items (i.e., intimate partner, friend, parent, other relative/family member), four formal source items (i.e., mental health professional, phone helpline, doctor/GP, minister or religious leader), an item for no one (i.e., "I would not seek help from anyone"; reverse-scored), and an item for other source (i.e., "I would seek help from another not listed above [please list in the space provided]"). Participants rated the likelihood they would seek help from each source on a 6-point Likert scale from 1 (*extremely unlikely*) to 7 (*extremely likely*), with higher scores indicating greater intention. Convergent evidence of the validity for the single GHSQ item measuring intention to seek help from a mental health professional was presented in the form of a positive association with perceived quality of previous mental health care and an inverse association with perceived barriers to seeking counseling (Wilson et al., 2005). Wilson and colleagues provided evidence of internal consistency and three-week test-retest reliability for the personal-emotional problems subscale score ($\alpha = .70$; $r = .86$).

Mental Help-Seeking Intention Scale (MHSIS). The three-item MHSIS was designed to measure respondents' intention to seek help from a mental health professional if they had a mental health concern. Participants rated their degree of intention using a 6-point Likert scale from 1 (e.g., *definitely false*) to 7 (e.g., *definitely true*), with higher scores indicating greater intention. The TPB's principle of compatibility guided the adaptation of these items to the help seeking context, as articulated in Ajzen's (2006) guide. This principal specifies that the behavior of interest and the intention instrument must be defined in terms of the same elements of target (e.g., mental health professional), action (e.g., seeking help), context (e.g., for assistance with a mental health concern), and time (e.g., upon the development of a mental health concern).

Internal consistency (α 's > .87) and convergent evidence of validity for the MHSIS score has been documented in the form of significant positive associations between intention and both attitudes and subjective norms around seeking professional psychological help (Hammer & Vogel, 2013; Hess & Tracey, 2013; Mo & Mak, 2009). Given that some researchers may be interested in using a single-item help seeking intention instrument (e.g., reduced participant burden; use in national surveys) that can produce a valid score, we also tested an alternative single-item version of the MHSIS consisting of the first MHSIS item (i.e., "If I had a mental health concern, I would intend to seek help from a mental health professional").

Prospective Help Seeking Behavior. To measure actual future help seeking behavior, participants were asked at Time 2 whether (yes/no) they had sought help from a mental health professional (i.e., psychologist, psychiatrist, social worker, counselor) in the last three months.

Analytic Approach

Twelve different instrument models (M1-M12) were considered in the present investigation. Models 1-4 represented different versions of the ISCI, M5-M10 represented different versions of the GHSQ, and M11 and M12 represented different versions of the MHSIS (see Table 1). Some versions were eligible for dimensionality testing, reliability testing, and/or predictive evidence of validity testing, while other versions were not (see below for explanations of eligibility at each kind of test). See the Supplemental Material for data cleaning procedures.

Dimensionality. Models 1-8 were eligible to be tested for the degree of global measurement model fit using a series of confirmatory factor analysis (CFA) measurement models with *Mplus* version 6.11 (Muthén & Muthén, 1998-2012). *Mplus*' MLR option for maximum likelihood estimation was used, which calculates the scaled chi-square test statistic (scaled χ^2). Model fit was evaluated using the scaled χ^2 statistic, 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 \leq .06, CFI \geq .95, TLI \geq .95, SRMR < .08 for good fit and RMSEA \leq .10, CFI \geq .90, TLI \geq .90, SRMR < .10 for acceptable fit (Weston & Gore, 2006). Regarding statistical power, Preacher and Coffman's

(2006) web utility indicated that a minimum sample size of 340 participants (cf. present study's $N = 405$) was needed to achieve a power of .80 when alpha is .05, Null RMSEA is .05, Alternative RMSEA is .08, and df is 27. Lower df requires a larger sample size, and 27 was the lowest df of all models calculated in the present study (see Table 1).

Mathematically, the GHSQ-3 (M9) and MHSIS-3 (M11) cannot be tested for degree of model fit. Therefore, standardized factor loadings and standardized residual variances from these models were examined to determine the degree to which a single factor accounted for the preponderance of the item-level variance. Such a test is less conclusive than testing for degree of global measurement model fit but still offers useful information. Finally, given that the GHSQ-1 (M10) and MHSIS-1 (M12) are single-item instruments, they could not be tested for dimensionality.

Reliability. All versions that demonstrated appropriate dimensionality (see Dimensionality Results) and had more than one item were eligible for reliability testing (i.e., M3, M4, M8, M9, M11). The internal consistency of the instruments' scores was tested using Cronbach alpha estimates. In addition, Factor Determinacy (FD) and construct reliability/replicability (H index) were calculated for all instrument factors. An $FD > .90$ would indicate that any observed differences in the factor score is indicative of true individual differences on the factor (Gorsuch, 1983). A H index $> .80$ would indicate that the latent factor variable is likely to be replicable across studies and useful in a SEM measurement model (Rodriguez, Reise, & Haviland, 2016).

Predictive Evidence of Validity. All versions that demonstrated appropriate dimensionality (see Dimensionality Results) and that contained items focused solely on seeking help from a mental health professional were eligible for predictive evidence of validity testing (i.e., M3, M4, M10, M11, M12). (It should be noted that combining two, three, or all four of the GHSQ's formal help seeking items into a score, and using this score to try and predict seeking help from mental health professionals specifically, is not a fair test of the GHSQ, as only one GHSQ item is relevant to the criterion of seeking help from a mental health professional.)

Predictive evidence of validity for the instruments' scores was examined using a series of logistic regressions in SPSS (Version 23). Specifically, the betas, Wald's χ^2 , Cox and Snell R^2 , Nagelkerke R^2 , percent correctly classified, and Odds Ratios for each instrument were calculated, with higher values interpreted as stronger predictive evidence of validity. All instrument scores were standardized through conversion to the standard normal distribution (i.e., z scores) prior to being entered in the logistic regressions.

In the first logistic regression (M3), the three subscale scores of the 17-item ISCI (i.e., with the "weight control" item) were used to predict prospective help seeking behavior. The second logistic regression (M4) used the three subscale scores of the 16-item ISCI (i.e., without the "weight control" item). The third logistic regression (M10) used the item score for the GHSQ item measuring intention to seek help from a mental health professional. The fourth logistic regression (M11) used the total score of the three-item MHSIS. The fifth logistic regression (M12) used the item score for the single-item MHSIS (i.e., "If I had a mental health concern, I would intend to seek help from a mental health professional").

Results

Dimensionality

Global fit statistics for the eligible models (M1-M8) are provided in Table 1. The ISCI has been treated as both a unidimensional instrument that produces a single total score and a multidimensional instrument with three correlated factors that produce three subscale scores. Unidimensional models for versions of the ISCI with (M1) and without (M2) the "weight control" item provided a poor fit to the data, whereas corresponding correlated factors models (M3 and M4, respectively) provided an adequate fit to the data.

The GHSQ has been treated as a unidimensional instrument that produces a single total score and a multidimensional instrument with between two and six correlated factors that produces between two to six subscale scores, depending on which items are used. Unidimensional models advocated by Wilson et al., (2005) with (M5) and without (M6) the "other source" item provided a poor fit to the data. A correlated factors model (M7), in which

the four informal help source items (i.e., intimate partner, friend, parent, other relative/family member) and the four possible formal help source items (i.e., mental health professional, phone helpline, doctor/GP, minister or religious leader) were set to load on informal and formal factors, respectively, provided a poor fit to the data. A correlated factors model (M8), in which the four informal help source items and three formal help source items (i.e., mental health professional, phone helpline, doctor/GP) used by Wilson (2010) were set to load on informal and formal factors, respectively, also provided a poor fit for the data.

A unidimensional model (M9) incorporating three formal help source items for “mental health professional”, “doctor/GP”, and “phone helpline”, which mirrors the GHSQ version used by Wilson & Deane (2012), produced standardized factor loadings (and standardized residual variances) of .55 (.70), .50 (.75), and .50 (.75). Thus, most of the variance (R^2) for the three items (70%, 75%, and 75%) was not explained by a single factor, which suggests that a unidimensional model may not provide an adequate fit to this three-item GHSQ. In summary, all tested versions of the GHSQ failed to demonstrate clear evidence of adequate fit in the present dataset.

The MHSIS is designed and conceptualized as a unidimensional instrument that produces a single total score (Ajzen, 2006). The unidimensional model (M11) for the three-item MHSIS produced standardized factor loadings (and standardized residual variances) of .92 (.15), .91 (.15), and .92 (.16). Thus, the vast majority of the variance (R^2) for each of the three items (85%, 83%, and 84%) was explained by a single factor, tentatively suggesting that a unidimensional model may provide an adequate fit in the present dataset. This provided initial support for modeling the three-item MHSIS as a unidimensional instrument.

Reliability

Dimensionality evidence supported modeling the ISCI as a correlated factors instrument (M3 and M4) that produces three scores corresponding to the three subscales articulated by Cepeda-Benito and Short (1988). Therefore, the internal consistency of these three subscale scores was examined. All scores demonstrated evidence of internal consistency: Psychological

and Interpersonal Concerns with the “weight control” item included ($\alpha = .87$, 95% CI [.848, .892]), Psychological and Interpersonal Concerns with the “weight control” item excluded ($\alpha = .87$, 95% CI [.849, .893]), Academic Concerns ($\alpha = .78$, 95% CI [.741, .821]), and Drug Use Concerns ($\alpha = .92$, 95% CI [.899, .935]). In addition, the FD (.94, .90, .98) and H index (.89, .81, .96) scores for the Psychological and Interpersonal Concerns, Academic Concerns, and Drug Abuse Concerns subscale scores, respectively, exceeded the recommended minimum cutoffs.

Dimensionality evidence did not provide clear support for modeling the three formal help source items of the GHSQ as a unidimensional instrument (M9) that produces a single total score. Consistent with this finding, the GHSQ-3 formal total score did not demonstrate evidence of internal consistency: $\alpha = .52$, 95% CI (.429, .595). (It should be noted that every other possible combination of the 10 GHSQ items demonstrated internal consistency estimates below .66.) In addition, the FD (.72) and H index (.52) scores for the GHSQ-3 formal total score fell below the recommended minimum cutoffs.

Dimensionality evidence provided initial support for modeling the three-item MHSIS as a unidimensional instrument that produces a single total score. The MHSIS-3 total score demonstrated evidence of internal consistency ($\alpha = .94$; 95% CI [.929, .949]). In addition, the FD (.97) and H index (.94) scores for the MHSIS-3 total score exceeded the recommended minimum cutoffs.

Predictive Evidence of Validity

Table 2 summarizes the predictive evidence of validity for the eligible instrument versions (i.e., M3, M4, M10, M11, M12). Results indicated that the MHSIS-3 demonstrated the strongest evidence of predictive validity, with a correct classification rate near 70%. From strongest to weakest degree of predictive evidence, the instruments were as follows: MHSIS-3, MHSIS-1, GHSQ-1, ISCI-16, and ISCI-17.

Discussion

The present study examined the dimensionality, internal consistency, and predictive evidence of validity for three help seeking intention instruments. Results supported

conceptualizing the ISCI as a multidimensional instrument that can be modeled with a correlated factors solution, regardless of the “weight control” item’s inclusion. The ISCI findings align with extant evidence supporting the correlated factor structure and internal consistency of the ISCI (Cepeda-Benito & Short, 1998). These findings do not provide support for the practice, seen occasionally in the literature (e.g., Pheko et al., 2013), of treating the ISCI as a unidimensional instrument that produces an interpretable total score.

The GHSQ has been treated as a unidimensional instrument that produces a single total score and a multidimensional instrument with between two and six correlated factors that produce between two to six subscale scores, depending on the GHSQ items used. Our examination of several plausible or previously used versions of the GSHQ indicated that no tested version of the instrument demonstrated clear evidence of adequate fit and internal consistency in the present sample. These results stand in contrast to studies reporting the GSHQ evidenced appropriate dimensionality and reliability (e.g., Tullio & Velasquez, 2014; Wilson et al., 2005).

Results provided initial support for conceptualizing the three-item MHSIS as a unidimensional instrument that produces an internally consistent total score with appropriate construct replicability. The three-item MHSIS demonstrated predictive evidence of validity by correctly predicting, with almost 70% accuracy, the future help seeking behavior of community adults with a current mental health concern. The single-item MHSIS (i.e., “If I had a mental health concern, I would intend to seek help from a mental health professional”) was the second strongest predictor.

Counseling psychologists and allied social scientists interested in improving mental health help seeking behavior can consider the present findings in making decisions about which instrument may offer the best utility for their research, clinical, or program evaluation needs. For example, users interested in a brief intention instrument that has demonstrated initial evidence of reliability and predictive evidence of validity may find the MHSIS suitable for their purposes. Users wishing to directly compare their findings to past studies may find the ISCI or GHSQ

preferable, given the more frequent use of these instruments. Users desiring an instrument that asks about intention to seek help for a variety of presenting concerns may find the ISCI advantageous. Those wanting to compare intention to seek help from formal versus informal sources may find the GHSQ appealing.

However, each instrument demonstrated limitations that temper these potential advantages. First, the MHSIS is not yet widely used, so the stability of its psychometric properties is not yet established. Second, the ISCI received the weakest predictive evidence of validity of the instruments tested and is longer than the other two instruments. Given that many researchers use the ISCI on the assumption that it measures the closest proxy to actual help-seeking behavior (i.e., intention), this weaker predictive evidence of validity is worth careful consideration. Third, the GHSQ demonstrated a lack of simple factor structure and reliability, and less predictive evidence of validity than the MHSIS. Thus, researchers should carefully consider when and if the GHSQ's ability to attend to both formal and informal help seeking intention outweighs the psychometric advantages of the other help seeking intention instruments.

Addressing Limitations Through Future Research

The three intention instruments demonstrated psychometric strengths and limitations in the present study. However, instrument validation is an ongoing process (AERA et al., 2014) and future research is a prerequisite to confident conclusions about the advantages of certain intention instruments over others. We consider these findings to be contextually dependent on the nature of the sample. Most of our sample was composed of White, educated women who had sought help in the past and were more willing than not to seek future help from a mental health professional. Though anyone can join ResearchMatch (RM), this research registry is likely overrepresented by people with personal interest in health research. In fact, this interest was a key reason for recruiting from RM: to help reduce attrition at T2. In comparison to the general US population, the present sample and RM population have greater access to the internet and are more likely to be female, middle-aged, White, and from the Southeast United States. In sum, the present sample is not a nationally-represented sample. Thus, future research investigating the

performance of these instruments using non-research registry samples primarily composed of people of color, men, people with less formal education, and those with less willingness to seek help is a prerequisite to claims about the generalizability of these results to these populations. In addition, because 27% of Time 1 respondents did not complete the Time 2 survey, it is possible that a degree of self-selection bias was introduced. Perhaps those who said they would seek help, but did not, would anticipate feeling shame when asked at Time 2, resulting in inflated classification rate for the three instruments.

The issue of gender and help-seeking requires additional comment. Men, compared to women, are less likely to seek psychological help for a wide range of psychosocial concerns (e.g., depression, substance abuse) (Addis & Mahalik, 2003; Vogel & Heath, 2016). Researchers have often explained this discrepancy as a product of internalized masculine social norms that promote both self-reliance and emotional control (Addis & Mahalik, 2003). As noted above, the present sample was primarily composed of female respondents. Interestingly, the factor analysis results for the three instruments remained consistent when analyzing only the female respondents' data (see Supplemental Material for detailed results). Given women's greater willingness to seek mental health treatment (Addis & Mahalik, 2003), the present findings may or may not extend to certain all-male samples. Because the smaller number of men in the present sample provided insufficient power for measurement equivalence/invariance testing across gender, we strongly recommend this testing in future research on these intention instruments.

In addition, formal psychometric evaluation of other forms of validity evidence (e.g., concurrent, test content, response processes, causal) in favor of all three instruments would be beneficial, given the limitations identified in the present study. Because multiple versions of these instruments exist, the generalizability of these results to those versions untested in the present study is uncertain, and we encourage future users of these instruments to verify our results.

References

- Addis, M. E., & Mahalik, J. R. (2003). Men, masculinity, and the contexts of help seeking. *American Psychologist, 58*, 5-14. <http://dx.doi.org/10.1037/0003-066X.58.1.5>
- 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. (2006). Constructing a theory of planned behavior questionnaire. Available from <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 (AERA), 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.
- Cash, T. E., Begley, P. J., McCown, D. A., & Weise, B. C. (1975). When counselors are heard but not seen: Initial impact of physical attractiveness. *Journal of Counseling Psychology, 22*, 273-279. <http://dx.doi.org/10.1037/h0076730>
- Cepeda-Benito, A., & Short, P. (1998). Self-concealment, avoidance of psychological services, and perceived likelihood of seeking professional help. *Journal of Counseling Psychology, 45*, 58-64. <http://dx.doi.org/10.1037/0022-0167.45.1.58>
- Deane, F. P., Wilson, C. J., & Ciarrochi, J. (2001). Suicidal ideation and help-negation: Not just hopelessness or prior help. *Journal of Clinical Psychology, 57*, 901-914. <http://dx.doi.org/10.1002/jclp.1058>
- Demyan, A. L., & Anderson, T. (2012). Effects of a brief media intervention on expectations, attitudes, and intentions of mental health help seeking. *Journal of Counseling Psychology, 59*, 222-229. <http://dx.doi.org/10.1037/a0026541>

- DeVellis, R. F. (2016). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Gourash, N. (1978). Help-seeking: A review of the literature. *American Journal of Community Psychology, 6*, 413-423. <http://dx.doi.org/10.1007/BF00941418>
- Gorsuch, R. L. (1983). *Factor analysis* (2nd. ed.). Hillsdale, NJ: LEA.
- 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. <http://dx.doi.org/10.1037/a0030449>
- Hasking, P., Rees, C. S., Martin, G., & Quigley, J. (2015). What happens when you tell someone you self-injure? The effects of disclosing NSSI to adults and peers. *BMC Public Health, 15*, 1039. <http://dx.doi.org/10.1186/s12889-015-2383-0>.
- 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. <http://dx.doi.org/10.1002/j.1556-6676.2013.00100.x>
- Kelly, A. E., & Achter, J. A. (1995). Self-concealment and attitudes toward counseling in university students. *Journal of Counseling Psychology, 42*, 40-46. <http://dx.doi.org/10.1037/0022-0167.42.1.40>
- Lannin, D. G., Vogel, D. L., Brenner, R. E., & Tucker, J. R. (2014). Predicting self-esteem and intentions to seek counseling: The internalized stigma model. *The Counseling Psychologist, 43*, 64-93. <http://dx.doi.org/10.1177/0011000014541550>
- Leech, N. L. (2007). Cramer's model of willingness to seek counseling: A structural equation model for counseling students. *The Journal of Psychology, 141*, 435-448. <http://dx.doi.org/10.3200/jrlp.141.4.435-448>
- Mo, P. K., & Mak, W. W. (2009). Help-seeking for mental health problems among Chinese. *Social Psychiatry and Psychiatric Epidemiology, 44*, 675-684. <http://dx.doi.org/10.1007/s00127-008-0484-0>

- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide (7th ed.)*. Los Angeles, CA: Authors.
- McDermott, R. C., Cheng, H., Wong, J., Booth, N., Jones, Z., & Sevig, T. (2017). Hope for help-seeking: A positive psychology perspective of psychological help-seeking intentions. *The Counseling Psychologist, 45*, 237-265. <http://dx.doi.org/10.1177/0011000017693398>
- 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*, 212-224. <http://dx.doi.org/10.1002/j.1556-6676.2015.00197.x>
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: Lessons from educational psychology. *Practical Assessment, Research, & Evaluation, 8*, Retrieved from <http://PAREonline.net/getvn.asp?v=8&n=11>
- Parent, M. C. (2013). Handling item-level missing data: Simpler is just as good. *The Counseling Psychologist, 41*, 568–600. <http://dx.doi.org/10.1177/0011000012445176>
- 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*, 373-384. <http://dx.doi.org/10.1037/0022-0167.54.4.373>
- Pheko, M. M., Chilisa, R., Balogun, S. K., & Kgathi, C. (2013). Predicting intentions to seek psychological help among Botswana university students: The role of stigma and help-seeking attitudes. *Sage Open, 3*, 1-11. <http://dx.doi.org/10.1177/2158244013494655>
- Preacher, K. J., & Coffman, D. L. (2006). Computing power and minimum sample size for RMSEA. [Computer software]. Available from <http://quantpsy.org>.
- Rickwood, D., Deane, F. P., Wilson, C. J., & Ciarrochi, J. (2005). Young people's help-seeking for mental health problems. *Australian E-journal for the Advancement of Mental Health, 4*, 218-251. <http://dx.doi.org/10.5172/jamh.4.3.218>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*, 137-150. <http://dx.doi.org/10.1037/met0000045>

- Seward, A., & Harris, K. M. (2016). Offline versus online suicide-related help seeking: Changing domains, changing paradigms. *Journal of Clinical Psychology, 72*, 606-620. <http://dx.doi.org/10.1002/jclp.22282>
- Straiton, M. L., Hjelmeland, H., Grimholt, T. K., & Dieserud, G. (2013). Self-harm and conventional gender roles in women. *Suicide and Life-Threatening Behavior, 43*, 161-173. <http://dx.doi.org/10.1111/sltb.12005>
- Tuliao, A. P., & Velasquez, P. A. (2014). Revisiting the General Help Seeking Questionnaire: Adaptation, exploratory factor analysis, and further validation in a Filipino college student sample. *Philippine Journal of Psychology, 47*, 1-17.
- Vogel, D. L., & Heath, P. J. (2016). Men, masculinities, and help-seeking patterns. In Y. J. Wong & S. R. Wester, (Eds.), *APA handbook of men and masculinities* (pp. 685–707). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14594-031>
- Vogel, D. L., Wade, N. G., & Haake, S. (2006). Measuring the self-stigma associated with seeking psychological help. *Journal of Counseling Psychology, 53*, 325-337. <http://dx.doi.org/10.1037/0022-0167.53.3.325>
- 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*, 459-470. <http://dx.doi.org/10.1037/0022-0167.52.4.459>
- 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*, 841-850. [https://dx.doi.org/10.1016/S0140-6736\(07\)61414-7](https://dx.doi.org/10.1016/S0140-6736(07)61414-7)
- Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*, 719-751. <http://dx.doi.org/10.1177/0011000006286345>
- Wilson, C. J. (2010). General psychological distress symptoms and help-avoidance in young Australians. *Advances in Mental Health, 9*, 63-72. <http://dx.doi.org/10.5172/jamh.9.1.63>

- Wilson, C. J., & Deane, F. P. (2012). Brief report: Need for autonomy and other perceived barriers relating to adolescents' intentions to seek professional mental health care. *Journal of Adolescence, 35*, 233-237.
<http://dx.doi.org/10.1016/j.adolescence.2010.06.011>
- Wilson, C. J., Deane, F. P., Ciarrochi, J., & Rickwood, D. (2005). Measuring help-seeking intentions: Properties of the general help-seeking questionnaire. *Canadian Journal of Counselling and Psychotherapy, 39*, 15-28.
- Wilson, C. J., Rickwood, D. J., Bushnell, J. A., Caputi, P., & Thomas, S. J. (2011). The effects of need for autonomy and preference for seeking help from informal sources on emerging adults' intentions to access mental health services for common mental disorders and suicidal thoughts. *Advances in Mental Health, 10*, 29-38.
<http://dx.doi.org/10.5172/jamh.2011.10.1.29>

Table 1

Goodness of Fit Statistics for All Tested Measurement Models

Model	Scaled χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
(1) ISCI-17 Unidimensional	689.45	119	.109 [.101, .117]	.720	.679	.087
(2) ISCI-16 Unidimensional	667.68	104	.116 [.107, .124]	.711	.667	.090
(3) ISCI-17 Correlated Factors	290.40	116	.061 [.052, .070]	.914	.899	.059
(4) ISCI-16 Correlated Factors	267.11	101	.064 [.054, .073]	.915	.899	.059
(5) GHSQ-10 Unidimensional	273.19	35	.130 [.116, .144]	.471	.320	.093
(6) GHSQ-9 Unidimensional	256.94	27	.145 [.129, .161]	.462	.283	.092
(7) GHSQ-8 Correlated Factors	69.69	19	.081 [.061, .102]	.836	.758	.056
(8) GHSQ-7 Correlated Factors	60.65	13	.095 [.072, .120]	.820	.709	.057
(9) GHSQ-3 Unidimensional*			-			
(10) GHSQ-1 Unidimensional*			-			
(11) MHSIS-3 Unidimensional*			-			
(12) MHSIS-1 Unidimensional*			-			

Note: *These models cannot be tested for degree of global measurement model fit but are displayed here to aid comparison of models across the manuscript text and tables. Bolded models indicate that this version of the instrument was eligible for predictive evidence of validity testing (see Table 2). The scaled chi-square for models 1 through 8 were statistically significant at the $p < .001$ level. ISCI = Intention to Seek Counseling Inventory. GHSQ = General Help Seeking Questionnaire. Statistics are based on MLR estimation. Scaled χ^2 = scaled chi-square test statistic, RMSEA = Root Mean Square Error of Approximation, CI = Confidence Interval, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, SRMR = Standard Root Mean Square Residual.

Table 2

Predictive Evidence of Validity derived from Logistic Regression

Model	<i>B</i>	<i>SE B</i>	Wald's χ^2	df	<i>p</i>	Odds ratio	95% CI	% correctly classified	Cox & Snell R^2	Nagelkerke R^2
(3) ISCI-17 Correlated Factors								59.1%	.03	.04
PIC	0.37	0.18	3.93	1	.05	1.44	[1.004, 2.069]			
AC	0.04	0.18	0.05	1	.82	1.04	[.737, 1.468]			
DUC	-0.01	0.15	0.00	1	.95	0.99	[.745, 1.319]			
(4) ISCI-16 Correlated Factors								60.0%	.04	.05
PIC	0.42	0.18	5.30	1	.02	1.53	[1.065, 2.189]			
AC	0.02	0.17	0.01	1	.93	1.02	[.722, 1.429]			
DUC	-0.02	0.15	0.03	1	.87	0.98	[.732, 1.301]			
(10) GHSQ-1 Unidimensional	0.81	0.15	29.57	1	.00	2.24	[1.676, 2.999]	64.6%	.11	.15
(11) MHSIS-3 Unidimensional	1.10	0.17	44.74	1	.00	3.01	[2.179, 4.154]	69.7%	.18	.24
(12) MHSIS-1 Unidimensional	1.04	0.16	41.09	1	.00	2.83	[2.058, 3.887]	67.3%	.17	.23

Note: PIC = Psychological and Interpersonal Concerns Subscale, AC = Academic Concerns Subscale, DUC = Drug Use Concerns Subscale, ISCI = Intention to Seek Counseling Inventory, GHSQ = General Help Seeking Questionnaire, MHSIS = Mental Help-Seeking Intention Scale.