The Wisdom Development Scale: Translating the Conceptual to the Concrete

Scott C. Brown Jeffrey A. Greene

In a previous study, a conceptual model of wisdom was created (Brown, 2004a) to better understand integrated learning outcomes. The purpose of this study is to develop a scale to measure this wisdom construct. This article discusses salient aspects of the extant professional literature regarding the measurement of wisdom and details the efforts to develop a valid and reliable Wisdom Development Scale (WDS) through exploratory and confirmatory factor analyses. Six of the seven factors were validated and the scale had acceptable confirmatory factor analysis fit. The article concludes with limitations of the study, implications for future research, and potential applications in higher education.

As stakeholders are raising questions about what students are getting out of their educational experiences, higher education has responded by increasing its attention toward the assessment of student learning. However, while a college education is more accurately conceptualized as a sum that is greater than its individual parts, many studies related to learning outcomes isolate discrete aspects of the college experience. Since discrete measures provide easily understood measures of accountability, they miss more complex and meaningful college-related student growth. Increasingly there is a greater interest in understanding the more ineffable outcomes of students' aggregate college experiences that account for what they reflect, integrate, and apply what they learn in and out of class, on and off campus (Brown, 2002a). Wisdom is a construct that subsumes many of the integrated and complex learning outcomes collegiate faculty and staff associate with a college education (Brown, 2004a). A means is necessary to measure wisdom in a collegiate environment.

The literature on wisdom has many rich parallels to the literature on learning within colleges and universities, and encompasses many outcomes normally associated with higher education (Brown, 2004a). While it is unlikely that college students would have achieved an extremely high level of wisdom, this article's conception of wisdom is as a continuum on which all people, including college students, can be placed. In a previous study, a conceptual model of wisdom was created (Brown, 2004a). Although Brown's Model of Wisdom Development provides a framework to consider the multi-dimensional changes a student might go through during college, it did not have an empirical way to measure them. The purpose of this study is to develop a scale to measure this wisdom construct. This article discusses salient aspects of the extant professional literature regarding the measurement of wisdom and details the efforts to develop a valid and reliable wisdom scale through exploratory and confirmatory factor analyses. The article concludes with limitations of the study and implications for future research.

LITERATURE REVIEW

Wisdom has long been a phenomenon of interest, particularly in terms of religion and

Scott C. Brown is the Director of the Daniel L. Jones Career Development Center at Mount Holyoke College. Jeffrey A. Greene is a graduate student in Educational Psychology at the University of Maryland, College Park.

philosophy (Kekes, 1983; Robinson, 1990). More recently, there have been various conceptual treatments of wisdom as a multidimensional construct, examining its: problem-finding attributes (Arlin,1990); distinction from intelligence (Clayton, 1982); combination of cognition and affect (Blanchard-Fields, Brannan & Camp, 1987; Csikszentmihalyi & Rathunde, 1990, Labouvie-Vief, 1990); integration of the multiple dimensions of self (Kramer, 1990); inclusion of response and recognition of human limitation and knowledge (Meacham, 1983, Taranto, 1989); differences based on gender (Orwoll & Archenbaum, 1993), differences between western and eastern frames of thought (Takahasi, 2000); and balance of intrapersonal, interpersonal, and extrapersonal interests in the environmental context to achieve a common good (Sternberg, 1998).

Additionally, there have been several efforts to study wisdom empirically (Ardelt 2003; Baltes & Smith, 1990; Sternberg, 1985; Webster, 2003; Wink & Helson, 1997). This literature review focuses on quantitative measures of wisdom and is divided into three categories: implicit theories, wisdom-related performance theories, and latent factor theories.

Empirical Research on Wisdom

Empirical research on wisdom can be grouped into three categories. Many studies (Holliday & Chandler, 1986; Sternberg, 1985) have examined what Sternberg (1998) would call implicit theories of wisdom: how the general public defines wisdom. Another category of wisdom research involves the analysis of wisdom-related performance (WRP) and is primarily done by researchers at the Max Planck Institute (Baltes & Staudinger, 1993, 2000). Finally, there are studies, such as Webster (2003) and Ardelt (1997, 2003), that try to measure wisdom as a construct through latent analysis. This study is informed by the first two categories, but is most similar to those in the last.

Implicit Theories of Wisdom. Any study of a socially constructed, complex idea such as wisdom must not only examine the philosophical bases of the term, but also the implicit or folk-psychology understandings of the term by people outside of academia. Sternberg (1985) asked both academicians and laypersons to describe the behaviors of wise, intelligent, and creative people. Multidimensional scaling analysis using these terms, along with research into how well individuals characterized themselves and others using these terms, revealed that people's implicit ideas of wisdom are, for the most part, separate from one another. Implicit ideas about wisdom included aspects of reasoning ability, sagacity, the ability to learn from ideas and the environment, judgment, the expeditious use of information, and perspicacity.

Holliday and Chandler (1986) also conducted a study to determine (a) if a prototype construct could be constructed from individual's implicit theories about wisdom, (b) if this prototype was consistent among age groups, and (c) if wisdom was distinct from intelligence. The first phase entailed generating descriptions of wise people. One hundred fifty research participants representing young, middle-aged, and older adults were sampled. The research participants tried to delineate differences between the following attributes: wise, intelligent, shrewd, spiritual, perceptive, and foolish. A trained rater transcribed the data, combined them thematically, and removed tangential information. The second phase involved 150 new participants, categorized in the same three cohort groups. Participants were given a series of words that people used to describe wisdom and were

asked to sort them into groups. Principal components analysis yielded five overarching factors explaining 41% of the variance: exceptional understanding (of essences, contexts and the self), judgment and communication skills, general competencies, interpersonal skills, and social unobtrusiveness. The authors concluded that though there is an overlap of the implicit theory of wisdom with intelligence, perceptiveness, spirituality and shrewdness, it is evident that wisdom is a distinct term and not a composite of other terms.

Thus, the research on implicit theories of wisdom demonstrates that the construct, while complicated, does seem to be distinct in laypersons' minds. Several studies (Holliday & Chandler, 1986; Sternberg, 1985) found some overlap of the construct with intelligence, but overall the terms shared little variance in people's definitions. With the knowledge that wisdom has a distinct, if not common, definition among researchers and laypersons, the next step is to determine how it can best be measured. Two approaches to the measurement of wisdom include wisdom-related performance and latent measurement models.

Wisdom-Related Performance. The Max Planck Institute in Berlin, Germany has a long history of studying wisdom. Rather than relying on implicit theories of wisdom, Baltes and Smith (1990) developed their own explicit, or researcher-derived, theory (Sternberg, 1985). The authors defined wisdom as "expert knowledge involving good judgment and advice about important but uncertain matters of life" (p. 95). According to them, the wise person embodies several essential criteria: rich factual knowledge, rich procedural knowledge, life span contextualism, relativism, and uncertainty. Smith and Baltes (1990) researched this construct by examining wisdom-related performance (WRP) and age/cohort differences. Participants were asked to think aloud while responding to life-planning and morally challenging problems. Trained raters then graded this verbal protocol. These researchers did not attempt to measure the latent construct of wisdom directly, but rather through participants' performance in scenarios that seemingly called for wise responses. Their analyses of 60 participants found that younger research participants displayed more wisdom in scenarios closer to their own experience, whereas older research participants were more likely to display wisdom in unfamiliar or non-normative scenarios. Overall, the researchers concluded there was evidence for a weak developmental trend in wisdom, but that the important finding was that WRP could be consistently measured using their methods.

In a subsequent study (Staudinger, Maciel, Smith, & Baltes, 1998), WRP was studied using participants' responses to morally difficult or life-planning scenarios. In this analysis, clinical psychologists were tested as a subgroup expected to score more highly, due to the researcher's belief that these professionals would have more cause to engage in the types of reflection and self-understanding necessary for wisdom. This subgroup was compared to other laypersons, and significant differences were found. Specifically, clinical psychologists were more likely to exhibit high WRP, and WRP across all groups showed only limited overlap with personality and intelligence measures.

Later studies referred to the researchers' work as the "Berlin Wisdom Paradigm" and proposed that wisdom was a:

> Metaheuristic . . . that organizes, at a high level of aggregation, the pool (ensemble) of bodies of knowledge and commensurate, more specific heuristics that are available to individuals in planning,

> managing, and evaluating issues surrounding the fundamental pragmatics of life. (Baltes & Staudinger, 2000, p. 132)

Their further work in this area covered descriptive examinations of people nominated as wise, the continued performance of theoretically wise groups such as clinical psychologists, and the examination of wisdom in common proverbs (Baltes & Staudinger, 2000). Finally, Staudinger and Pasupathi (2003) have examined wisdom in adolescence and found that measures of intelligence and personality are stronger predictors of WRP in young people, whereas more integrative measures of these constructs are stronger predictors of WRP in adults.

However, as Webster (2003) notes, it is unclear whether WRP is truly an accurate measure of wisdom in the individual. It is possible those participants are providing socially desirable answers or that groups found to be wiser, such as clinical psychologists, are guessing the purpose of the research and providing acquiescent answers. In any case, it is not at all clear that those who display WRP are truly wise, and would act in such a way in a naturalistic setting or in their real lives. While the complex scenarios used by the researchers have the advantage of soliciting rich responses, they also provide the participants an opportunity to provide an ideal response, rather than a more genuine one.

Latent Factor Analyses of Wisdom. Finally, wisdom as a construct can be measured through latent factor techniques. In these studies, rather than examining WRP as a proxy for wisdom, the researchers have attempted to assess wisdom using survey methods. These surveys, however, must be shown to be both reliable and valid with target populations before they can be used as indicators of wisdom. Some studies put forth instruments but fail to adequately demonstrate reliability or validity. For example, Wink and Helson's (1997) measures of practical and transcendent wisdom from a longitudinal study of women (N = 94, at age 22, 27, 43, 52) and their partners (N = 44, at age 27, 52) show adequate reliability with those items and those samples, but the actual factor structure of the instrument was never tested for validity. Their results revealed weak to moderate correlations of their scale with other measures hypothesized to be associated with wisdom. Correlations can demonstrate convergent validity, but the issue of construct validity requires other kinds of evidence (DeVellis, 2003). A better approach to establishing construct validity would be to perform factor analysis on the instruments themselves, to determine whether they are valid indicators of unitary constructs, such as wisdom.

Other researchers have used factor analysis to establish construct validity. Ardelt (1997) performed secondary analysis on data from the Berkeley Guidance Study (82 women and 39 men interviewed in 1968/9 as part of 40-year follow-up study), to examine whether a proposed wisdom composite of cognitive, reflective, and affective measures could predict life satisfaction. Using structural equation modeling to establish construct validity, Ardelt (1997) demonstrated that this composite improved the fit of the model above and beyond other factors such as gender, physical health, socioeconomic status, and financial situation. However, many of the measures Ardelt (1997) utilized lacked internal reliability (with many levels below .5) and that the data were first collected over 60 years prior to her analysis. It is not clear that the experiences of people three generations ago are comparable to those today, nor is it clear that measures that lacked reliability at that time could be combined to produce a valid wisdom composite, particularly given that the original

study was not investigating this construct.

More recently, Ardelt (2003) has developed a three-dimensional wisdom scale (3D-WS) with 180 older adult participants from close-knit social groups (age 52+). Despite acknowledging the difficulties in measuring wisdom even through latent factor methods, Ardelt (2003) believes that the construct should manifest through self-report indicators. She posited that three dimensions make up wisdom: cognitive, reflective, and affective, with this last dimension fostering the other two. Ardelt (2003) found acceptable fit in her confirmatory factor analyses, and the scale showed good predictive and discriminant validity with other measures. While Ardelt's (2003) findings are encouraging, her scale was analyzed using less than 200 research participants, all of whom were from an older population. Additional analyses using larger data sets and a broader range of research participants would lend further credence to her model and scale. In addition, it is not clear how Ardelt (2003) went about the process of first pruning down her scale items and then confirming the overall reduced scale. Ideally, scale items are first analyzed using exploratory factor analysis or some other technique to determine which items are working effectively and which are not. Then, the final scale of acceptable items is given to another set of participants and confirmatory factor analysis performed to cross-validate the scale's item and factor structure. It is not clear that Ardelt (2003) used different research participants for the exploratory and confirmatory analyses in her study, and if not, then it is not surprising that exploratory and confirmatory analyses would agree on good fit. In essence, the confirmatory study was almost assured to find that the exploratory results were a good fit, since it was the same set of data (DeVellis, 2003).

Webster (2003) did a 3-part exploratory analysis of a self-assessed wisdom scale (SAWS) to examine its initial consistency and validity (39 men and 46 women, 22-78 years of age). The SAWS is a 30-item, 6-point Likert-type scale addressing five dimensions: experience, emotions, reminiscence, openness, and humor. Results showed that the scores from the scale with that sample showed an overall reliability of .78. A principal components analysis was performed to determine if the five dimensions were present in the item responses, and the data supported this claim. Webster attempted to demonstrate the construct validity of the SAWS by linking it to the Erikson's theory of generativity and the concept of ego integrity. Webster found mild positive correlations between the SAWS and measures of generativity and ego integrity. Interestingly, the correlations with neither age nor educational level were found to be significant.

Despite the findings above, there are several cautions regarding Webster's (2003) scale. The development of the scale items themselves was not based on any research or expert analysis. Conceptually, the inclusion of humor as an aspect of wisdom is not well supported, as the concept has not been seen in other implicit wisdom studies, and humor's factor loadings were somewhat low in the data analysis. In addition, Webster performed only exploratory factor analyses, whereas a confirmatory factor analysis would lend support to the scale's replicability. Finally, the somewhat weak correlations between ego integrity, generativity, and SAWS are surprising given that the author used them as evidence of the construct validity of his measure.

The research on wisdom can be seen as covering three aspects of construct analysis. First, numerous researchers (Holliday & Chandler, 1986; Sternberg, 1985) have established that wisdom stands as a rather

distinct concept in people's minds. With a somewhat common idea of wisdom, the next step would be to determine how best to measure this construct. Baltes, Staudinger, and associates at the Max Planck institute have chosen to examine WRP as opposed to attempting to measure the construct through latent means (Baltes, Staudinger, Maerker, & Smith, 1995, Smith & Baltes, 1990; Staudinger, Maciel, Smith, & Baltes, 1998; Staudinger & Pasupathi, 2003). Others (Ardelt, 1997, 2003; Webster, 2003) have chosen to attempt to measure wisdom as a latent factor influencing participants' responses to survey research. In both cases, there are methodological concerns, including respondent acquiescence, a failure to fully validate the reliability and replicability of the instruments used, and unclear or unacceptable methods of confirming the fit of the model to the data.

Current Model: An Implicit and Explicit Analysis of Wisdom

Brown's (2004a) study defined wisdom, examining how it develops and speculating as to what conditions facilitate its development. Brown (2004a) asked 10 recent graduates to reflect on the salient aspects of their undergraduate college experiences in three semi-structured interviews. Participants were nominated by campus administrators on the basis of their strong academic achievement, involvement in co-curricular offerings, and observed ability to integrate their college experiences. The participants represented diversity on a number of dimensions, including major, ethnicity, gender, religion, and types of involvements in college. Grounded theory techniques were used to analyze the data compiled from three interviews with each of the 10 participants (Brown, Stevens, Troiano, & Schneider, 2002). Open coding of the raw transcripts of the data yielded more than 1000 individual concepts, combined in axial coding to create 54 more comprehensive, abstract categories, and finally aggregating into five main "key" categories and one crucial "core" category in the final selective coding. One key category, "wisdom," was defined as self-knowledge, understanding of others, judgment, life knowledge, life skills, and willingness to learn. Wisdom develops when students go through the core "learning-from-life" process, comprised of reflection, integration, and application. The conditions that facilitate the development of wisdom by directly or indirectly stimulating the "learning-from-life" process are the student's (a) orientation to learning, (b) experiences, (c) interactions with others, and (d) environment. Depending on how deeply and how often students were stimulated to go through the learning from life process, they experienced growth on one or more of the six dimensions of wisdom listed above. Brown's Model of Wisdom Development is presented in Figure 1. Trustworthiness was achieved because two peer debriefers reviewed each transcript and helped test emerging designs and hypotheses, assisting in clarifying and deepening aspects of data analysis that may have been missed by the researcher. Additionally, a separate inquiry auditor ensured that grounded theory procedures were followed properly (Lincoln & Guba, 1986). Thus, this model is both implicit (by exploring folk ideas of wisdom) and explicit (through the use of theoretical texts and expert review). The next step in this research is to take a quantitative look at the model's viability.

This study will use a latent factor approach to refine Brown's Model of Wisdom Development. Through the use of survey methodology, both Brown's Model of Wisdom Development and the instrument designed to

measure it will be tested. This study differs from those above in that it has taken aspects of implicit theories of wisdom and integrated them with a scale development and validation plan that, if successful, provides greater confidence in both the constructs measured and the means of doing so.

Hypotheses

This study sought to develop a measure to test and explore Brown's (2004a) Model of Wisdom Development using latent factor techniques. Specifically, both exploratory and confirmatory factor analyses were performed on the model using different sets of participants as a means of cross-validating the latent factor structure of the model. The process of scale development and validation are outlined. The hypotheses were that a six-factor structure, made explicit in Brown's (2004a) model, would be found using both exploratory and confirmatory analyses, and that the overall scale would have an acceptable level of reliability of at least .70 (DeVellis, 2003), predicted factor pattern coefficients in the exploratory analyses of .3 or higher, and acceptable confirmatory factor analysis fit as specified by the criteria of Hu and Bentler (1999). In addition, the six factors were hypothesized to intercorrelate, given that the model predicts that all six contribute to the overall concept of wisdom.

METHOD

This section details the methods employed in the development of the Wisdom Development Scale (WDS), including survey item creation, participant selection, and survey administration procedures. The goal of this study was to create a theoretically and psychometrically valid instrument. Only when this process has been successfully completed can new evaluation initiatives, such as using scale scores

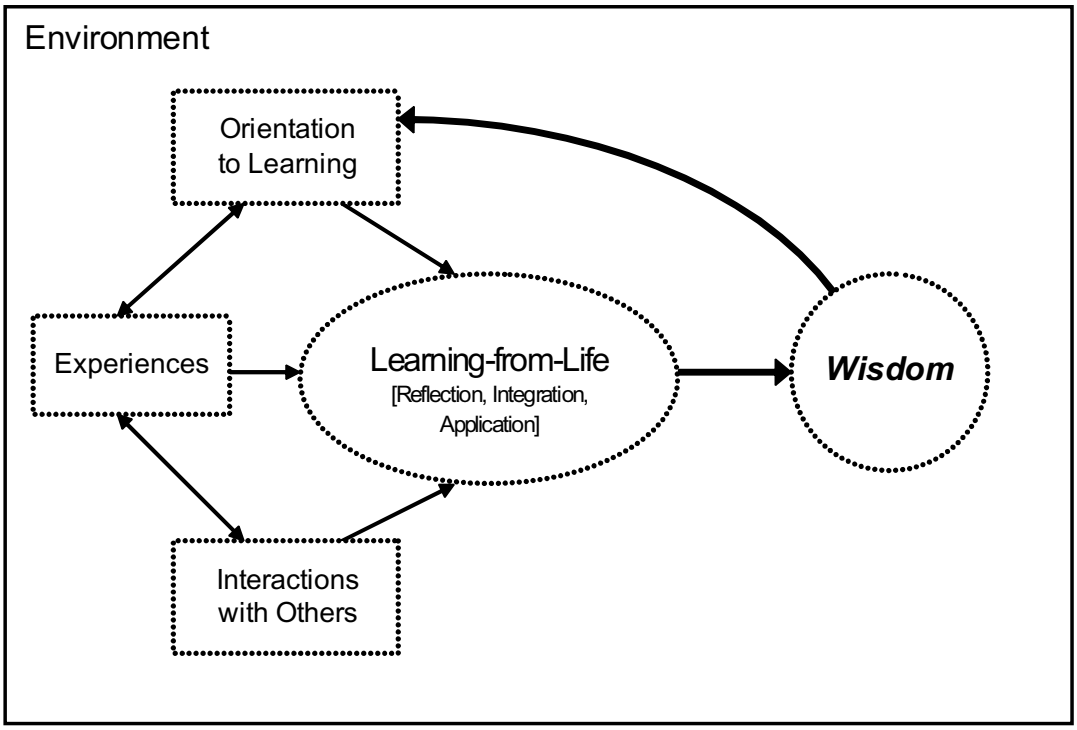

FIGURE 1. Brown's Model of Wisdom Development

to measure intervention effectiveness, be investigated.

Content Validity of the Wisdom Measure

Several steps were taken to establish the content validity of scores from the wisdom measure. Focus groups with students and educators were conducted to refine an instrument with items derived from the original study related to each dimension of the wisdom construct (i.e., self-knowledge, understanding of others, judgment, life knowledge, life skills, and willingness to learn; Brown, 2004a). Focus groups were employed early in the research process to provide a basis for the larger quantitative research involving many more participants (Stewart & Shamdasani, 1990). In essence, an analysis of the model's implicit nature was performed.

Three student focus groups were selected by variation sampling to ensure diversity of major and ethnicity. Prior to the instrument being administered, all focus group participants were asked to describe someone who they believed was wise and then took the pilot questionnaire individually. After they completed the questionnaire, they were asked several questions regarding the survey's comprehensiveness and clarity. In addition, 24 career counselors and student affairs administrators from a convenience sample based on age and gender were also interviewed, individually and in groups, to refine the wisdom construct and instrument. Discussions followed the same protocol as the student focus groups.

Revisions were made to the survey from all focus group interviews. If there were multiple interpretations to an item, they were turned into separate items (Ouimet, Bunnage, Carini, Kuh, & Kennedy, 2004). Items were also added that articulated aspects of Brown's (2004a) original wisdom construct.

WDS Questionnaire

A 141-item, seven-point Likert-type scale web-based questionnaire, revised after the group and individual interviews, was administered to a random sampling of 7050 undergraduate students from a large state university in the Northeast. The Likert-type scales ranged from *Strongly Disagree* (1) to *Strongly Agree* (7). No items were negatively worded for fear that it would confuse respondents given the large number of items (DeVellis, 2003). The web survey was formatted so several items were seen per screen, and the survey scrolled down. Radio buttons were used instead of pull-down menus because of their greater ease of use (Heerwegh & Loosveldt, 2002).

The Marlowe-Crowne Social Desirability Scale-Short Form (SDS; Reynolds, 1982) was added to the survey as a brief check for the confounding influence of socially-desirable responding. The SDS is a short 13-item form of the validated and frequently used Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960), a measure of an individual's tendency to respond in a socially-desirable fashion. Participants are asked to mark whether each statement is true or false as it pertains to them personally. A sample item is, "No matter who I'm talking to, I'm always a good listener." Scores range from 0 to 1, and high scores indicate more socially-desirable responding. The SDS has evidenced internal consistency and construct validity (Fischer & Fick, 1993; Fraboni & Cooper, 1989).

Participants

This study aimed for 1000 valid responses. Since Internet administrations often have low response rates, the survey invitation was sent to 7050 students. Although the survey had a relatively high number of total items, survey

length has not been shown to decrease response rate (Cook, Heath, & Thompson, 2000). Removing duplicate submissions (due to students pressing the "submit" button multiple times) and problematic response sets (three students responded with the same value for all items and were removed), the response rate was 1188 out of 7050 (17%). Of this sample, 32.6% were male and 60.7% female (6.7% did not report). The mean age of the sample was 21.1 years (*SD* = 4.1). The ethnic breakdown was 80.2% White, 2.7% Black, 6.4% Asian, 3.2% Hispanic, .5% Native American, and .3 % Other. This sample was roughly comparable to the undergraduate population. For scale validation purposes, Gorsuch (1983) recommends a sample to item ratio of 5 to 1, but more recent research suggests that even factors with low factor pattern coefficients can be reasonably substantiated with sample sizes above 500 (DeVellis, 2003). There were no missing data because the on-line administration did not allow students to submit their survey unless there was a valid response for each item.

Procedure

Electronic mail messages were sent to the sample, giving them an overview of the study and confidentiality information, and also inviting them to take the web survey. The following principles of constructing web surveys were followed: The web survey was formatted like a paper survey: an easy, comprehensible question appeared at the top of the screen as did periodic reminders of where participants were in the survey; and participants were not required to provide an answer to each question before being allowed to answer subsequent ones (Dillman, Tortora, & Bowker, 1999). A chance to win one of five gift certificates to an on-line merchant was given as an incentive. The initial invitation and two reminders were sent out at the end of the fall semester, and the data collection lasted one week. Students were contacted three times in an attempt to increase web response rates (Cook et al., 2000). E-mail invitations included the date the survey would be taken down which has been shown to have a positive effect on response rates (Porter & Whitcomb, 2003).

RESULTS

Scale validation consisted of both exploratory and confirmatory methods. The sample was split in half through random assignment, with exploratory factor analysis performed on the first half and confirmatory factor analysis performed on the second. Results of the exploratory analysis were used to produce the final scales used for confirmatory analyses of the model. The confirmatory factor analysis on the second half of the data allowed for another means of validation regarding the scale scores' reliability and factor structure (DeVellis, 2003). Details of both procedures are presented, along with analysis of the Social Desirability Scale information.

Exploratory Factor Analysis

Exploratory factor analysis (EFA) is often used in scale design to determine the covariance of items within hypothesized latent factor structures (Loehlin, 1998). The goal of EFA in scale development is to determine whether items hypothesized to load on common factors in fact do so. Items that do not load as hypothesized are examined and most often dropped from the final scale, with the intent that the final version of each scale have from five to ten items strongly loaded on it.

To determine if EFA was appropriate for this set of a data, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was

TABLE 1.
Dimensions of the Wisdom Development Scale (WDS)

Original		Wisdom Development Scale
Self-Knowledge	→	Self-Knowledge
Interpersonal Understanding	→	Altruism
	→	Inspirational Engagement
Judgment	→	Judgment
Life Knowledge	→	Life Knowledge
Life Skills	→	Life Skills
	→	Emotional Management
Willingness to Learn		Construct not found

examined. A KMO analysis determines whether factor analysis of any kind is warranted given the data. Using SPSS 11.5, the KMO was .944, indicating excellent intercorrelations between each pair of items after partialling out the linear effects of all other items, thus supporting a factor analysis (Loehlin, 1998). Given that the scales were hypothesized to be measuring latent factors, rather than composites, principal axis factoring (PAF) was used (Gorsuch, 1983). Given that there were over 100 items, a more sophisticated rule for factor extraction was needed than the commonly used Kaiser-Guttman rule. Thus, the scree plot was visually examined and the eigenvalues analyzed using the Cattell-Nelson-Gorsuch objective scree test (Gorsuch). This test showed a clear jump in eigenvalues and provided support for an eight-factor solution. In addition, direct oblimin rotation (correlated factors) was employed given the hypothesized correlation between latent factors. This eight-factor PAF with oblique rotation yielded a significant value for Bartlett 's Test of Sphericity, but research suggests that large sample sizes often yield significance due to the test's conservative nature (Loehlin). Thus, any concerns regarding sphericity were dismissed and the factor analysis was performed. The extracted sums of squared factor pattern coefficients for the eight factors prior to rotation was 56.365, explaining 40% of the variance in the sample.

The rotated pattern matrix was examined for hypothesized factor and item relations. Items grouping on a common factor were examined for interpretability given the hypothesized relations. Items were excluded from the factor interpretation if they loaded less than .3 on the factor, if they had cross-loadings on another factor greater than .3, or if their hypothesized factor relation was contrary to the majority of other EFA-derived items loading on the common factor (DeVellis, 2003; Gorsuch, 1983). Results revealed strong support for the hypothesized Self-Knowledge

and Life-Knowledge factors. The original "interpersonal understanding" construct was split into two factors, which upon review were determined to be appropriately entitled Altruism and Inspirational Engagement. Likewise, the hypothesized construct of "life skills" was also split between two factors, which we entitled Emotional Management and Life Skills (Table 1).

A seventh factor consisted of items from multiple hypothesized factors, but upon review was determined to be an accurate measure of the Judgment factor. The eighth factor consisted of five items but was not interpretable based upon a review of the items. The items' factor pattern coefficients were .321, .342, .470, .500, and .524. These items will be set aside for future research into why they loaded together, but given that their grouping did not fit with the theoretical framework of the study; this factor was not investigated further. The remaining factor intercorrelations ranged from .119 to .345.

Items with corrected item-total correlations below .5 and alpha-if-item-deleted values below total scale alpha were examined for removal as well. This resulted in the removal of 3 items overall. Reliability analyses on the final scale scores revealed Cronbach alpha values above .8 for each of the remaining seven factors. A list of final factors with total number of items and Cronbach alphas for each factor is provided in Table 2, and a list of factors, communalities, items and individual factor pattern coefficients can be found in Table 3.

Social Desirability Scale. Research participants' Social Desirability Scale (SDS) scores were determined by summing their responses to those items. This total score was then correlated with each item of the WDS (Ardelt, 2003). No item correlated higher than .195 with the SDS total score. Nonetheless, due to the large sample size, even relatively small correlations were found to be significant, and with so many items, the threat of Type I errors increases. Given that both of these factors would make the test more liberal, we took a more conservative approach to item removal. In the final scales derived from the EFA, the Altruism scale had 7 of 14 items with significant correlations with SDS, the Inspirational Engagement scale had 8 of 11 items, the Judgment scale had 4 of 11 items, the Life Skills scale had 3 of 9 items, and the Emotional Management scale had 9 of 11 items. Nonetheless, no item had a correlation higher than .2 with the SDS, and Ardelt (2003) advocated removing only those items that correlated at .3 or higher. Therefore, it was determined that the correlations were not strong enough to warrant removal of the items. Future administrations of the WDS should continue to examine the potential impact of socially-desirable responding and perhaps utilize principal factor deletion when creating wisdom total scores (Paulhus, 2002).

TABLE 2.

Factor Scales with Number of Items and Loading Range

Factor	Number of Items	Cronbach Alpha
Self-Knowledge	6	.8440
Altruism	14	.8737
Like Knowledge	9	.8359
Emotional Management	9	.8425
Inspirational Engagement	11	.8772
Judgment	11	.8784
Life Skills	11	.8750

TABLE 3.

Factor Communalities, Individual Items, and Factor Pattern Coefficients in Exploratory Sample

Scale Item	Item Communality	Factor Pattern Coefficient
Self-Knowledge		
SK1	.460	.511
SK2	.404	.538
SK3	.687	.603
SK4	.707	.574
SK5	.700	.656
SK6	.629	.618
Emotional Management		
EM1	.450	.513
EM2	.527	.581
EM3	.585	.614
EM4	.442	.541
EM5	.420	.633
EM6	.399	.482
EM7	.426	.491
EM8	.304	.405
EM9	.443	.465
Altruism		
A1	.306	.374
A2	.377	.512
A3	.419	.532
A4	.384	.518
A5	.398	.453
A6	.379	.474
A7	.306	.429
A8	.327	.330
A9	.480	.443
A10	.449	.573
A11	.387	.301
A12	.466	.521
A13	.419	.316
A14	.526	.465
Inspirational Engagement		
IE1	.384	.429
IE2	.387	.433
IE3	.441	.359
IE4	.495	.491
IE5	.547	.480
IE6	.532	.392
IE7	.568	.390
Inspirational Engagement (continued)		
IE8	.431	.450
IE9	.498	.419
IE10	.394	.337
IE11	.460	.331
Judgment		
J1	.431	.464
J2	.443	.324
J3	.420	.356
J4	.416	.373
J5	.416	.426
J6	.385	.319
J7	.344	.368
J8	.474	.368
J9	.541	.372
J10	.524	.372
J11	.466	.509
Life Knowledge		
LK1	.404	.603
LK2	.484	.631
LK3	.308	.495
LK4	.299	.436
LK5	.463	.657
LK6	.534	.734
LK7	.413	.402
LK8	.407	.457
LK9	.347	.301
Life Skills		
LS1	.380	.391
LS2	.479	.654
LS3	.319	.385
LS4	.539	.665
LS5	.484	.398
LS6	.404	.600
LS7	.456	.407
LS8	.545	.544
LS9	.491	.544
LS10	.412	.345
LS11	.528	.567

TABLE 4.

Item Means, Standard Deviations, Skewness and Kurtosis for Confirmatory Sample

Scale Item	Mean	Standard Deviation	Skewness	Kurtosis
Self-Knowledge				
SK1	5.60	1.44	-1.04	1.80
SK2	5.40	1.46	-0.99	0.51
SK3	5.60	1.34	-1.07	0.76
SK4	5.42	1.35	-0.85	0.37
SK5	5.72	1.26	-1.06	0.76
SK6	5.56	1.39	-1.02	0.73
Emotional Management				
EM1	4.61	1.50	-0.51	-0.24
EM2	4.49	1.61	-0.30	-0.58
EM3	4.71	1.55	-0.57	-0.18
EM4	5.10	1.36	-0.64	0.15
EM5	4.45	1.80	-0.30	-0.92
EM6	4.36	1.77	-0.20	-0.91
EM7	5.66	1.26	-1.03	0.85
EM8	4.83	1.51	-0.54	-1.5
EM9	4.85	1.09	-0.82	0.39
Altruism				
A1	5.47	1.13	-0.75	0.95
A2	6.11	1.00	-1.54	3.16
A3	5.92	1.00	-1.26	2.52
A4	5.80	1.11	-0.99	0.80
A5	6.00	1.05	-1.37	2.46
A6	4.48	1.53	-0.32	-0.61
A7	5.39	1.29	-0.87	0.83
A8	5.92	0.96	-0.78	0.41
A9	5.85	1.00	-1.04	1.80
A10	6.02	1.07	-1.28	2.03
A11	5.88	1.01	-1.20	2.58
A12	5.75	1.15	-1.05	1.34
A13	5.82	1.12	-1.27	2.09
A14	5.98	1.03	-1.08	1.09
Inspirational Engagement				
IE1	4.96	1.31	-0.71	0.39
IE2	5.61	1.06	-0.82	1.38
IE3	5.26	1.26	-0.58	0.15
IE4	4.25	1.33	-0.81	0.53
IE5	5.06	1.28	-0.44	-0.31
IE6	5.69	1.17	-1.12	1.49
EI7	5.57	1.30	-1.04	0.89
IE8	5.65	1.18	-0.87	0.66
Inspirational Engagement (continued)				
IE9	5.75	1.10	-0.90	0.69
IE10	5.61	1.23	-1.00	0.89
IE11	5.52	1.09	-0.57	0.27
Judgment				
J1	6.23	0.98	-1.83	5.09
J2	5.96	1.01	-0.96	0.85
J3	6.00	1.11	-1.33	2.32
J4	5.85	1.16	-1.11	1.40
J5	6.02	1.06	-1.19	1.44
J6	5.81	1.09	-0.82	0.39
J7	6.19	1.09	-1.70	3.30
J8	5.92	1.09	-1.17	1.81
J9	5.91	0.97	-1.04	1.50
J10	5.95	1.05	-1.31	2.67
J11	6.20	0.99	-1.44	2.32
Life Knowledge				
LK1	5.18	1.44	-0.59	-0.15
LK2	5.63	1.18	-0.86	0.52
LK3	5.50	1.40	-0.79	0.05
LK4	5.50	1.37	-0.88	0.26
LK5	5.19	1.56	-0.68	-0.26
LK6	5.17	1.55	-0.60	-0.44
LK7	5.43	1.28	-0.74	0.22
LK8	5.11	1.23	-0.52	0.27
LK9	5.74	1.25	-1.10	1.23
Life Skills				
LS1	5.40	1.24	-0.69	0.22
LS2	4.71	1.60	-0.35	-0.74
LS3	5.32	1.48	-1.00	0.66
LS4	4.54	1.72	-0.38	-0.71
LS5	5.31	1.16	-0.75	0.78
LS6	5.66	1.51	-1.19	0.80
LS7	5.36	1.58	-0.99	0.27
LS8	5.42	1.30	-0.90	0.69
LS9	5.81	1.08	-1.12	1.68
LS10	5.80	1.16	-1.04	1.13
LS11	5.52	1.19	-0.90	0.96

TABLE 5.

Intercorrelations Between Factors in Confirmatory Sample
(all values significant at .05 level)

	Self-Knowledge	Altruism	Life Knowledge	Life Skills	Inspirational Engagement	Judgment	Emotional Management
Self-Knowledge	1						
Altruism	.524	1					
Life Knowledge	.450	.656	1				
Life Skills	.426	.484	.394	1			
Inspirational Engagement	.594	.700	.608	.720	1		
Judgment	.540	.783	.751	.519	.718	1	
Emotional Management	.552	.637	.468	.689	.859	.595	1

Confirmatory Factor Analysis

With empirically-derived sets of items reflecting the latent factor structure of the WDS, the second half of the sample was used for confirmatory purposes. Specifically, the proposed factor and item structures were analyzed for fit using EQS 6.1 (Bentler, 2004), allowing for the factors to correlate. Item means, standard deviations, skewness, and kurtosis can be found in Table 4 (covariance matrix available from authors). Robust maximum likelihood estimation was used due to the non-normality in the data, as evidenced by the item skewness and kurtosis values as well as the extremely high multivariate kurtosis normalized estimate of 168.1203 (Bentler, 2004; Byrne, 1994). No start values were given for any model parameter, and the model converged in five iterations. The Satorra-Bentler scaled chi square value (Bentler, 2004) was 5116.3436 with 2393 degrees of freedom, resulting in a p value of less than .00001. Given the chi-square's sensitivity to sample size, we paid less attention to this statistic compared to other indices of fit (Bentler, 1990). For this model, the chi-square value divided by its degrees of freedom is 2.949, a value considered a sign of good fit by some researchers (Kline, 2004). The standardized root mean residual (SRMR) was .068 and the root mean-square error of approximation (RMSEA) was .058 with a 90% confidence interval of .056 to .059. Therefore, using Hu and Bentler's (1999) joint criteria for model fit, the WDS was found to have acceptable fit in the confirmatory sample. However, it should be noted that other indices of fit, primarily incremental fit indices based upon a comparison of the theoretical and independence models, did not meet common standards (Hu & Bentler). The robust-adjusted comparative fit index (CFI) was .811 and the Bentler-Bonett Non-Normed Fit Index (NNFI) was .804. These values are below common standards, but recent research

into these types of fit indices suggests that they may degrade when models contain large numbers of variables (Kenny & McCoach, 2003). Therefore, following the advice of Kenny and McCoach, the finding of acceptable RMSEA and SRMR fit indices suggests that the scale has met the conditions for confirmatory factor analysis in this sample.

Intercorrelations between factors are listed in Table 5. The overall Cronbach alpha for the entire WDS (comprising all scale scores) was .963. Intercorrelations between the scales are high as hypothesized, but not so high as to suggest that the scales all measure a single construct (Kline, 2004).

Model Respecification

A Wald test (Byrne, 1994) was requested to determine whether any parameters could be dropped without a significant loss in data-model fit. The Wald test reported no parameters that could be dropped. In support of this, all unstandardized factor pattern coefficients were significant using robust standard errors as reported in EQS 6.1 (Bentler, 2004; Byrne). In terms of parameters to be added, the LaGrange Multiplier Test results listed numerous error covariance paths to be added, however, these were not added given that the model had already obtained sufficient fit given Hu and Bentler's (1999) joint criteria.

Summary

Exploratory factor analysis revealed strong support for two of the six hypothesized factors, with another two factors measured through two correlated sub-scales each. It is not at all unusual for hypothesized factors to end up being split into multiple subscales, and in these cases the factor integrity of the scales were maintained between each subscale as the item factor pattern coefficients were as hypothesized. The Judgment scale, although containing only a modest number of items originally hypothesized, proved to be theoretically coherent and was kept in the analysis. One benefit of the confirmatory factor analysis techniques used is that any factors not sufficiently described or hypothesized in the original model can be tested on a validation sample. Sample-specific factors that are not valid indicators of latent constructs usually result in poor fit indices in a CFA (DeVellis, 2003). Thus, the fact that all seven scales (including the two factor subscales and the new Judgment scale) had good indices of fit in the CFA demonstrates strong evidence that these factors are more likely to be reflections of latent constructs and not merely the result of sample-specific error. The overall good fit of the model to both the a priori theory and validation techniques provide strong supporting evidence for both the model itself and the scale measuring it. Thus, the hypotheses of model fit, strong scale score reliability, and intercorrelations among the scales were retained.

DISCUSSION

Brown's (2004a) model posited six factors of wisdom, and this study provides empirical support for five of those factors. The sixth factor, Willingness to Learn was not found, possibly due to population concerns, discussed in the limitations below. This study provides several useful contributions to the extant literature on wisdom. One of the most important advances concerns the advancement of a measure with strong content validity. Based on our findings, wisdom is defined as a multidimensional construct with seven dimensions: Self-Knowledge, Emotional Management, Altruism, Inspirational Engagement, Judgment, Life Knowledge and Life Skills. Limitations, areas for future research

and implications for educators are also outlined below.

Content Validity

Brown's (2004a) Model of Wisdom Development has a strong implicit and explicit nature. It has an implicit nature because the model was developed from participants' conceptions of wisdom. The explicit nature was demonstrated by the grounded theory techniques used to define wisdom, how it develops, and what conditions facilitate its development. The items crafted for the wisdom measure have strong content validity due to their being vetted by several focus groups of individuals and experts. Although other measures of wisdom may be termed in various ways and overlap in different configurations, the WDS has much content validity with other scales and treatments of wisdom. Holliday and Chandler's (1986) five variables related to the WDS: "exceptional understanding of essences and contexts and the self" (Life Knowledge/ Judgment and Self Knowledge), "judgment" (Judgment), "general competencies" (Life Skills), and "interpersonal skills" and "social unobtrusiveness" (Altruism/Inspirational Engagement). Sternberg's (1985) conception of wisdom has similarities to the WDS: "a deep understanding of self and others" (Self-Knowledge, Judgment), "sagacity" (Judgment, Altruism), and "expeditious use of information," "ability to learn from ideas and environment," "perspicacity," "discernability," and "judgment" (Judgment). Baltes and Smith (1990) defined wisdom as "expert knowledge involving good judgment and advice about important but uncertain matters of life" (p. 95). The wise person must have broad and specific knowledge about life (Life Knowledge), multiple ways of dealing with multiple life circumstances (Emotional Management, Life Skills), an understanding of different contexts in life over time (Life Knowledge, Judgment, Life Skills), knowledge about different ways of approaching life's major components to reason from multiple perspectives (Judgment), and recognition and management of uncertainty (Life Skills). Ardelt's (2003) 3D-WS also connects with the WDS: affective (Self Knowledge, Emotional Management), cognitive (Judgment), and reflective (Life Knowledge). Webster's (2003) experience, reminisce (Life Knowledge), and emotions (Emotional Management) also relate to the WDS.

Limitations of Study

There are several limitations of this study that lend themselves to future research. First, we would not expect college students to display the highest levels of wisdom, given that wisdom accrues over time. Thus, it is possible that the responses to the Willingness to Learn factor were too homogenous within this population. We hypothesize that sampling from a broader range of participants of a more varied age and experience would provide the variance necessary to elicit this factor. Therefore, future studies should include a broader sample in terms of age and life experience. Second, while there was a large sample size, the return rate was low. While Internet sampling is helpful, future studies should attempt to capture a larger percentage of any population sampled. Third, the predictive, convergent, and divergent validity of the scale has not been tested. These are all issues for future research.

Implications for Future Research

There are a number of promising ways to make the WDS more robust. First, in terms of the scale to model fit, during the exploratory factor analysis of the WDS, there were a number of items related to humility and uncertainty that

were surprisingly not included in the original model and present in much of the wisdom literature. A lack of variance in these items may reflect the youth of respondents. Thus, future studies should include a broader range of participants, both in terms of age and experience, who might display more variability, thus providing evidence for the construct. In addition, future research should include other measures that should theoretically either correlate (convergent validity) or not correlate (divergent) with wisdom to further establish the scalet's validity. Other wisdom-related constructs, such as WRP, should also be examined for their relation to this scale. Further demonstrations of the scale scores' reliability and validity with multiple samples will legitimize more applied work to be done with the scale, such as charting a student's movement over time through pre- and post-testing.

Applications

The potential applications of this scale are many, but are dependent upon future research. At this time, there is more work to be done regarding cross-validation and scaling. In particular, the question of scale scores will be addressed to determine where different populations fall in terms of their scores on this measure. Thus, both aggregate score across scales and individual scale scores will be examined by group, to determine the best means by which to measure student development using this instrument.

The scale could be used to examine the impact of the aggregate college experience, the effect of interventions within the college experience, such as those based on Brown's Model of Wisdom Development (Brown, 2002b, 2004b), and the differences in wisdom development across different types of institutions that strive to develop wisdom-related outcomes. Additionally, this scale can help educational leaders better understand the more ineffable value of the college experience and communicate these outcomes to external and internal stakeholders. Further, use of the scale may shed light on whether there are levels or phases of wisdom development. Lastly, the relationship between the WDS and existing psychosocial (Chickering & Reisser, 1993) and cognitive (Baxter Magolda, 2002; King & Kitchener, 1994; Perry, 1970) theories of human development should also be examined.

CONCLUSION

Brown's (2004a) Model of Wisdom Development provides a framework to consider the multi-dimensional changes a student might go through during college. This study has established a preliminary scale to measure this wisdom construct, and scores were found to be reliable and valid through both exploratory and confirmatory factor analyses with this sample. This study is of particular interest to constituencies who have a vested interest in undergraduate education, and with further work, the scale can inform educational leaders' efforts to increase the holistic, integrative learning experience for their students.

Correspondence concerning this article should be addressed to Scott C. Brown, scbrown@mtholyoke.edu

REFERENCES

Ardelt, M. (1997). Wisdom and life satisfaction in old age. *Journal of Gerontology, 52,* 15-27.

Ardelt, M. (2003). Empirical assessment of a three-dimensional scale. *Research on Aging, 25,* 275-324.

Arlin, P. K. (1990). Wisdom: The art of problem finding. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 230-243). Cambridge: Cambridge University Press.

Baltes, P. B., & Smith, J. (1990). Toward a psychology of wisdom and its ontogenesis. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 87-120). Cambridge: Cambridge University Press.

Baltes, P. B., & Staudinger, U. M. (1993). The search for a psychology of wisdom. *Current Directions in Psychological Science, 2*(3), 75-80.

Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist, 55,* 122-136.

Baltes, P. B., Staudinger, U. M., Maerker, A., & Smith, J. (1995). People nominated as wise: A comparative study of wisdom-related knowledge. *Psychology and Aging, 10,* 155-166.

Baxter Magolda, M. B. (2002). *Making their own way: Narratives for transforming higher education to promote self-development.* Sterling, VA: Stylus.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M. (2004). *EQS: A structural equations program.* Encino, CA: Multivariate Software, Inc.

Blanchard-Fields, F., Brannan, J. R., & Camp, C. J. (1987). Alternative conceptions of wisdom: An onion peeling exercise. *Educational Gerontology, 13,* 497-503.

Brown, S.C. (2002a, June). *Understanding the Ineffable: College and the Development of Wisdom*. Paper presented at the American Association for Higher Education Assessment Forum, Boston, MA.

Brown, S. C. (2002b). A model for wisdom development-and its place in career services. *Journal of College and Employers, 62*(4), 29-36.

Brown, S. C. (2004a). Learning across campus: How college facilitates the development of wisdom. *Journal of College Student Development, 45,* 134-148.

Brown, S. C. (2004b). Where this path may lead: Understanding post-college decision-making. *Journal of College Student Development, 45,* 375-390.

Brown, S. C., Stevens, R. A., Troiano, P. F., & Schneider, M. K. (2002). Exploring complex phenomena: Grounded theory in student affairs research. *Journal of College Student Development, 43,* 173-183.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming.* Thousand Oaks, CA: Sage.

Chickering, A. W., & Reisser, L. (1993). *Education and identity* (rev. ed.). San Francisco: Jossey-Bass.

Clayton, V. (1982). Wisdom and intelligence: The nature and function of knowledge in the later years. *International Journal of Aging and Human Development, 15,* 315-321.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web or internet-based surveys. *Educational and Psychological Measurement, 60,* 821-836.

Csikszentmihalyi, M., & Rathunde, K. (1990). The psychology of wisdom: An evolutionary interpretation. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 25-51). Cambridge: Cambridge University Press.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed., Vol. 26). Thousand Oaks: Sage.

Dillman, D. A., Tortora, R. D., & Bowker, D. K. (1999). Principles for constructing web surveys. Retrieved September 23, 2005, from http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf

Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement, 53*(2), 417-424.

Fraboni, M., & Cooper, D. (1989). Further validation of three short forms of the Marlowe-Crowne Scale of Social Desirability. *Psychological Reports, 65,* 595-600.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Heerwegh, D., & Loosveldt, G. (2002). An evaluation of the effect of response formats on data quality on web surveys. *Social Science Computer Review, 20,* 471-484.

Holliday, S. G., & Chandler, M. J. (1986). *Wisdom: Explorations in adult competence.* New York: Karger.

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

Kekes, J. (1983). Wisdom. *American Philosophical Quarterly, 20,* 277-286.

Kenny, D. A., & McCoach, B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*(3), 333-351.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco: Jossey-Bass.

Kline, R. B. (2004). *Principles and practice of structural equation modeling* (2nd Ed.). New York: Guilford Press.

Kramer, D. A. (1990). Conceptualizing wisdom: The primacy of affect-cognition relations. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 279-316). Cambridge: Cambridge University Press.

Labouvie-Vief, G. (1990). Wisdom as integrated thought: Historical and developmental perspectives. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 52-86). Cambridge: Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In D. D. William's (Ed.), *Naturalistic evaluation, No 30.* (pp. 73-84). San Francisco: Jossey-Bass.

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Meacham, J. A. (1983). Wisdom and the context of knowledge: knowing that one doesn't know. In D. Kuhn & J. A. Meacham (Eds.), *On the development of developmental psychology* (pp. 113-134). Basel, Switzerland: Karger.

Orwoll, L., & Achenbaum, W. A. (1993). Gender and the development of wisdom. *Human Development, 36,* 274-296.

Ouimet, J. A., Bunnage, J. C., Carini, R. M., Kuh, G. D., & Kennedy, J. (2004). Using focus groups, expert advice, and cognitive interviews to establish the validity of a college student survey. *Review of Higher Education, 45*, 233-250.

Paulhus, D. L. (2002). Social desirable responding: The evolution of a construct. In H. I. Braun & D. N. Jackson (eds.)., *The role of constructs in psychological and educational measurement.* Mahwah, NJ: Lawrence Erlbaum.

Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme.* New York: Holt, Rinehart and Winston.

Porter, S. R., & Whitcomb, M. E. (2003). The impact of contact type on web survey response rates. *Public Opinion Quarterly, 67,* 579-588.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the

Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38,* 1, 119-125.

Robinson, D. N. (1990). Wisdom through the ages. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 12-24). Cambridge: Cambridge University Press.

Smith, J., & Baltes, P. B. (1990). Wisdom-related knowledge: Age/cohort differences in response to life-planning problems. *Developmental Psychology, 26,* 494-505.

Staudinger, U. M., Maciel, A. G., Smith, J., & Baltes, P. B. (1998). What predicts wisdom-related performance? A first look at personality, intelligence and facilitative experiential contexts. *European Journal of Personality, 12,* 1-17.

Staudinger, U. M., & Pasupathi, M. (2003). Correlates of wisdom-related performance in adolescence and adulthood: Age graded differences in "paths" toward desirable development. *Journal of Research of Research and Adolescence, 13,* 239-268.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49,* 607-62.

Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology, 2,* 347-365.

Stewart, D. W., & Shamdasani, P. N. (1990). *Focus groups: Theory and practice.* Newbury Park, CA: Sage.

Takahashi, M. (2000). Toward a culturally inclusive understanding of wisdom: Historical roots in the east and west. *International Journal of Aging and Human Development, 51,* 217-230.

Taranto, M. A. (1989). Facets of wisdom: A theoretical synthesis. *International Journal of Aging and Human Development, 29,* 1-21.

Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development, 10,* 13-22.

Wink, P., & Helson, R. (1997). Practical and transcendent wisdom: Their nature and some longitudinal findings. *Journal of Adult Development, 4,* 1-15.