Measuring a Dimension of Spirituality for Health Research

Validity of the Spirituality Index of Well-Being

BRUCE B. FREY
University of Kansas
TIMOTHY P. DAALEMAN
University of North Carolina at Chapel Hill
VICKI PEYTON
University of Kansas

Health-related studies of *spirituality* are threatened by the lack of conceptual distinctions between religion and spirituality, the use of small, nongeneralizable samples, and by measurement error in many instruments that unreliably and invalidly capture this domain. The authors review the construct and validity evidence for the Spirituality Index of Well-Being (SIWB), an instrument designed to measure a dimension of spirituality linked to subjective well-being in patient populations. The SIWB was developed using qualitative research methods and subsequently conceptualized with two dimensions; *self-efficacy* and *life scheme*. Primary psychometric data from three sample populations are reviewed and summarized. A secondary, confirmatory factor analysis, using pooled data from all samples, supports the theoretical two-factor structure. In addition, SIWB scores correlate more strongly with established measures of well-being than the Spiritual Well-Being Scale (SWB) or other recognized religiosity instruments. The SIWB is a valid and reliable instrument that can be used in health-related studies.

Keywords: *spirituality; health; well-being; self-efficacy; life scheme*

There is ongoing interest in examining the association of religion and spirituality with health-related outcomes (Koenig, McCullough, and Larson 2001). A series of articles in the late 1980s suggested that

AUTHORS' NOTE: The authors wish to thank Bonnie J. Johnson for her contributions to this article. Address correspondence regarding this article to Bruce Frey, Department of Psychology and Research in Education, JR Pearson Hall, 1122 West Campus Road, Room 643, Lawrence, KS 66045; phone: (785) 864-9706; fax: (785) 864-3820; e-mail: bfrey@ku.edu.

RESEARCH ON AGING, Vol. 27 No. 5, September 2005 556-577
DOI: 10.1177/0164027505277847

morbidity and mortality rates varied by religious denomination and that high levels of involvement with religious activities were associated with better overall health status (Jarvis and Northcott 1987; Levin and Vanderpool 1987; Troyer 1988). Although health status has become a more recent, specific outcome of interest, one literature review found an inconsistent relationship between religiosity and perceived health, limited by a failure to control for known covariates of self-reported health (Musick 1996). In addition to confounding, two additional methodological shortcomings continue to threaten the validity of studies of religiosity, spirituality, and health outcomes: the use of small, nongeneralizable samples and the lack of valid and reliable instruments that reliably measure spirituality and/or religiosity (Sloan, Bagiella, and Powell 1999).

In considering measurement, conceptual distinctions between religion and spirituality, and the operationalization of constructs that reliably capture these domains, are critical in studies that examine health outcomes (George et al. 2000). Religion or religiosity has been viewed as the various organized, individual, and attitudinal manifestations of different faith traditions (Wulff 1997). There are many measures that gauge multiple dimensions of religiosity (Hill and Hood 1999); however, it remains unclear if these instruments are useful in studies of individual or population health. To address the lack of valid and reliable measures of religiosity and spirituality in health care settings, a group sponsored by the National Institute on Aging (NIA) and the Fetzer Institute was convened to identify appropriate measures for specific research objectives and to conceptualize the relationships between the measured concepts and health (Ory and Lipman 1998). The output of this work has led to a multidimensional measure of religion and spirituality for use in health research that was embedded in the 1998 General Social Survey (Idler et al. 2003). This instrument has demonstrated satisfactory psychometric properties in a general population; however, the majority of its items measure religiosity domains, such as private religious practices, religious support, and religious history rather than spirituality (Fetzer Institute and National Institute on Aging 1999). In addition, despite the contribution from prominent social scientists and health service researchers, the measure is not contextually sensitive to a patient and health care perspective.

The importance of context cannot be overstated as spirituality has become an accepted part of contemporary American culture (Wuthnow 1998) and by extension, part of the patient experience of health and illness as well (Koenig 2002). The health-related quality-of-life (HRQOL) field offers a framework to view spirituality within a health care context and from the patient perspective. HRQOL studies highlight the centrality of assessing patient experiences, beliefs, expectations, and perceptions in understanding health (Testa and Simonson 1996). One conceptual model of HRQOL locates the general perception of health as a subjective variable that directly affects overall quality of life (Wilson and Cleary 1995). In clinical practice, the measurement of HRQOL has been advocated to promote patient-centered care and to assist with clinical decision making (Rumsfeld 2002). From a population health standpoint, increased life expectancy in the United States highlights the importance of HRQOL, particularly in the areas of chronic illness, aging, and end-of-life care (National Center for Health Statistics 2002).

This article examines the Spirituality Index of Well-Being (SIWB), a 12-item instrument designed to measure one dimension of spirituality linked to subjective well-being in patient populations. First, we perform a limited concept analysis of spirituality to examine the characteristics of this concept across multiple disciplines (Walker and Avant 1995). We then outline SIWB scale development, which was derived from a qualitative study of patients to generate a conceptual framework and stimulus material for the instrument. The results of preliminary validity and reliability testing of the measure, conducted in a geriatric outpatient population, are presented. We also review two additional validity studies; one in a sample of primary care outpatients and the other in high school students. Finally, we pool data from these three studies and perform a confirmatory factor analysis to evaluate the dimensions of the scale.

Concept Development

CONCEPT ANALYSIS OF SPIRITUALITY

Concept analysis is a linguistic process, used in the social and nursing sciences, that determines the defining and irrelevant attributes of a

concept (Walker and Avant 1995). Although there are multiple steps in concept analysis, we will limit our analysis to determining the defining attributes (Wilson 1963). At the onset, a concept analysis is purposeful in its aims and inclusive in the sources and disciplines that use the concept. Our own work focused on distinguishing between religion and spirituality in the context of health-related research.

We have previously reviewed some uses of religion, which shares multiple connotations and interpretations as does its sibling, spirituality. For example, disciplines as varied as sociology (Coleman 1997), psychology (Wulff 1997), and theology (McGinn 1993) have conceptualized spirituality in various ways. In an exhaustive review of spirituality from the social work literature, one concept has been proposed that assumes an underlying human drive toward spirituality, which operationalizes the process of making sense of self and world (Canda and Furman 1999). The term *nonreligious spiritual propensity*, found in this literature, describes one who does not use religion as a foundational belief system but believes that all features of spiritual propensity can take on nonreligious forms. Spirituality is depicted as a condition of asking and answering major existential questions: Who am I? Why do I exist? What is my purpose? How do I fit in the world? Sociological perspectives are often congruent with this existential orientation because spirituality is represented as that which Roof (1993) described as giving expression to our inner being, a power that comes from knowing our deepest selves and what it is that is sacred to us.

A psychological perspective may be useful when considering spirituality's relationship with health or well-being, or when theories are presented that postulate a causal, mediating, or moderating relationship between the variables of interest. In psychology, the characteristics of spirituality are presented as a web of theoretical relationships within the domain of well-being (Pargament and Mahoney 2002). For example, positive psychology, which studies positive subjective experience, depicts spirituality as providing a framework for adjustment, growth, and reaching one's human potential (Seligman 2002; Snyder and Lopez 2002). Another psychological use includes coping mechanisms, although this largely arises from religiosity constructs (Pargament 1997).

Theological definitions of spirituality have roots within a largely contemporary, Catholic tradition (McGinn 1993). Spirituality often centers on beliefs in a divine being, as well as the sociological, philo-

sophical, and psychological manifestations of these beliefs (Cox 1996). For example, a recent white paper on professional chaplaincy depicts spirituality as "an awareness of relationships with all creation, an appreciation of presence and purpose that includes a sense of meaning" (Association of Professional Chaplains 2001). An additional theological use places spirituality within the stories, practices, and beliefs that are developed within largely religious traditions and communities but are carried into ordinary and everyday life events (Shea 2000). As a result, people can characterize themselves as religious and/or spiritual, understanding these concepts as independent, although perhaps related, qualities of what it is to be human (Zinnbauer et al. 1997).

These uses of spirituality from varied fields suggest that the concept of spirituality may be tied to attributes of personal meaning that may or may not be tied to religious traditions. A second attribute appears to be a largely positivistic characteristic; spirituality, in some ways, empowers and enables individuals to cope and to grow more fully.

CONCEPTUAL DEVELOPMENT OF THE SIWB

Our conceptualization of the SIWB was grounded from a qualitative study, which described the patient perspective of spirituality and its relationship with subjective health and well-being (Daaleman, Cobb, and Frey 2001). Rather than using expert panels or criteria, we used focus groups to capture the patient voice, an important approach in defining well-being constructs (Frey and Daaleman 1999). Figure 1 depicts the conceptual framework and the relationship of spirituality and well-being derived from this work. According to the framework, perceived threats or actual changes to functioning or health status trigger two patient-initiated tasks: (1) the gathering and processing of health-related information and (2) the interpretation and incorporation of these data within the context of the patient's life experience. These tasks rely on both lay illness explanations and professional information to construct or maintain an individual meaning system (Kleinman, Eisenberg, and Good 1978). Core beliefs are sources that contribute to these meaning systems and originate from social struc-

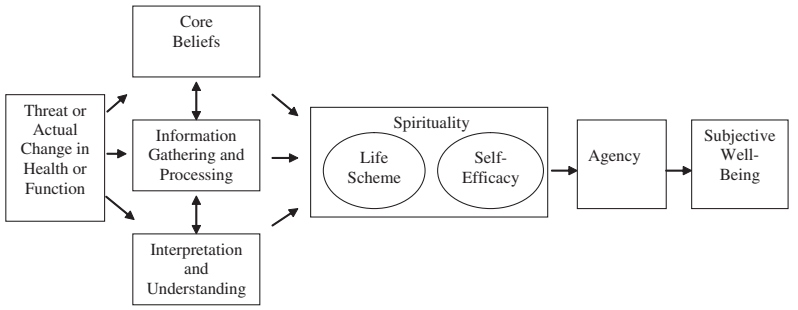

Figure 1: Conceptual Framework for Spirituality

tures, such as networks of friends or family, faith communities, and secular institutions.

The coping literature refers to the activity of "meaning making" as a cognitive representation of one's life that provides a sense of order and purpose (Thompson and Janigian 1988). This representation, or life scheme, is one primary domain of spirituality within the conceptual framework. It is similar to the construct of sense of coherence, which is described as a positive, pervasive way of viewing the world and one's life in it, lending elements of comprehensibility, manageability, and meaningfulness (Antonovsky 1987). Self-efficacy, the second domain, depicts an individual's belief in the capacity to organize and perform activities required for a prescribed goal (Bandura 1997). Within the framework of spirituality, self-efficacy beliefs are specific to the task of overcoming challenges to global functioning. A belief in overcoming real or apparent threats to individual problems and difficulties—regardless of perceived resources or individual capacities—is a key assumption within this domain.

Our understanding of spirituality, within the context of good health status and well-being, can be conceptualized as a congruent, meaningful life scheme and high functional self-efficacy beliefs that synergistically promote personal agency. Agency beliefs refer to the individual's self-view as an active participant who constructs his or her own life course through the choices and actions that he or she takes, given the opportunities and constraints of his or her circumstances (Bandura 1997).

Empirical Studies

VALIDITY EVIDENCE

New concepts, and the tests that are presumed to measure them, can only reasonably claim construct validity when they have been securely placed within a network of laws, rules, predictions, and expectations, according to Cronbach and Meehl (1955) in their classic work on the validity of psychological constructs. The web of observed relationships involving the construct, which make up part of a nomological net, or set of observable natural laws and testable predictions, will be limited in the early history of a construct. Indeed, a theoretical construct is more than a matter of elaborating the nomological network in which it occurs. Embretson (1983) provided a refinement to Cronbach and Meehl's definition of construct validity and referred to a *nomothetic span* as the network of relationships of a test to other measures that provide evidence as to the importance of a test as a measure of individual differences. Campbell and Fiske (1959) emphasized that because hypothetical human traits are usually invisible, inferences about the constructs must be made by observing human behaviors and individual differences and contended that convergent and discriminant evidence is required for construct validation. A test has convergent evidence of validity when it correlates highly with scores on tests that measure similar constructs. A test has discriminant evidence of validity when it does not correlate highly with tests that measure dissimilar constructs. Consequently, modern standards for psychological measurement emphasize analyses of the relationship of the scale scores to other variables as an important source of validity evidence (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education 1999).

The SIWB (see appendix) has been used to measure a dimension of health and well-being in peer-reviewed studies across a variety of populations. In every study, scale scores have shown meaningful and important associations with health and well-being constructs. In a sample of 277 community-dwelling elders (Geriatric Outpatient Study), the SIWB had significant moderate correlations with other recognized, well-validated quality-of-life instruments such as the Geriatric Depression Scale, the EuroQol, and SF-36 (Daaleman et al.

2002). In addition, no significant correlation was found with a measure of religiosity derived from the General Social Survey. A secondary analysis of the same data used univariate and multivariate regression analyses to examine determinants of self-reported health status (Daaleman, Perera, and Studenski 2004). Whereas physical functioning, race, age, geriatric depression, quality of life, and the SIWB total score were all predictive of self-reported good health, religiosity was not. In the full multivariate model, adjusting for all covariates, spirituality, as determined by the SIWB, remained a significant independent predictor of self-rated good health, whereas age and depression did not.

In a study that systematically sampled more than 500 adult outpatients at primary care clinic sites in the greater Kansas City area (Primary Care Study), SIWB scores strongly correlated with a validated measure of well-being, the General Well-Being Scale, in addition to the recognized Spiritual Well-Being Scale (SWB) (Daaleman and Frey 2004). Moderate associations in the expected directions were also found with the Zung Depression Scale and the Religious Well-Being subscale from the SWB (Ellison 1983). A sample of 577 students attending Catholic high schools in the Midwest (Catholic High School Study) was used to explore the relationship between SIWB scores and scores on the Children's Hope Scale (Snyder et al. 1997) (Frey et al. 2004). Moderate correlations between the SIWB and the full scale hope score, in addition to the two conceptually distinct hope subscales were reported. In addition, a gender difference was found, with women reporting higher scores on the SIWB Life Scheme subscale.

RELIABILITY EVIDENCE

The studies previously described include reliability estimates from three samples. Table 1 presents three coefficient alpha estimates of internal consistency for each of the SIWB indices; the full scale score and the two subscale scores. As with any statistic, reliability coefficients are an estimate of a population value and fall within a theoretical confidence interval (Thompson 2003) that is included in the table. In addition to these internal reliability estimates, the Primary Care Study found two-week test-retest correlations of .79 for the full scale, .86 for the Life Schema subscale, and .86 for the Self-Efficacy

TABLE 1
Coefficient Alpha Reliability Estimates for the Spirituality Index of Well-Being

	Coefficient Alpha Reliability Estimates		
	Full Scale Score	*Life Schema Subscale*	*Self-Efficacy Subscale*
Study Sample	*(95% CI)*	*(95% CI)*	*(95% CI)*
Catholic high school	.91 (.90-.92)	.86 (.84-.88)	.84 (.82-.86)
Primary care	.91 (.90-.92)	.89 (.87-.90)	.86 (.84-.88)
Geriatric	.87 (.85-.89)	.80 (.76-.83)	.83 (.80-.86)

subscale. These reliability indices suggest that the SIWB measures with consistency and stability across a wide age range. Most of the reliability estimates are the same across different studies, with the exception of the Geriatric Outpatient Study. For that sample, the full scale and the Life Schema subscale demonstrated statistically less internal consistency than was shown in other populations. Although statistically different, the estimates still represent very good to high reliability.

CONVERGENT AND DISCRIMINANT VALIDITY EVIDENCE

Empirically derived correlations between the SIWB scale scores and scores on measures of other constructs of interest provide information on the underlying constructs tapped by the scale (i.e., spirituality related to well-being, life scheme, self-efficacy). In early conceptual work, the construct of spirituality had been defined with the following attributes: a patient-centered aspect of subjective well-being, a cognitive process of interpreting information through a framework of core beliefs, and an independence from religious constructs (Daaleman et al. 2001; Frey and Daaleman 1999). Following this definition, SIWB scores should correlate highly with general measures of well-being and, although less so, with specific well-being measures, especially those that have a cognitive component. In addition, the SIWB should correlated weakly with functional health indi-

TABLE 2
Validity Coefficients for the Spirituality Index of Well-Being (SIWB)

	SIWB Full Scale	*Self-Efficacy Subscale*	*Life Schema Subscale*
Existential well-being[a]	.75***	.61***	.75***
General well-being[a]	.64***	.61***	.57***
Spiritual well-being[a]	.62***	.49***	.63***
Depression (1)[a]	–.42**	–.39**	–.39**
Hope[b]	.40***	.38***	.36***
Fear of death[c]	–.39**	–.33**	–.38**
Hope (agency)[b]	.36***	.34***	.32***
Hope (pathways)[b]	.36***	.34***	.32***
Depression (2)[c]	–.35**	–.31**	–.31**
Religious well-being[a]	.35***	.27***	.38***
Poor health[c]	–.35**	–.27**	–.36**
Physical functioning[c]	.28**	.28**	.23**
Quality of life[c]	.18**	.19**	.14*
Religiosity[c]	.12	.03	.18**

Variables and Measures

Construct	*Instrument*	*Source*
Depression (1)	Zung Depression Scale	Zung (1965)
Depression (2)	Geriatric Depression Scale	Yesavage et al. (1982)
Existential well-being	spiritual well-being scale	ellison (1983)
fear of death	Death Attitude Profile Scale–Revised	Wong, Reker, and Gesser (1994)
General well-being	General Well-Being Scale	McDowell and Newell (1996)
Hope	Children's Hope Scale	Snyder et al. (1997)
Hope (agency)	Children's Hope Scale	Snyder et al. (1997)
Hope (pathways)	Children's Hope Scale	Snyder et al. (1997)
Physical functioning	Short Form 36 (SF-36)	Stewart, Hays, and Ware (1988)
Poor health	Years of Healthy Life Scale	Erickson, Wilson, and Shannon (1995)
Quality of life	European Quality of Life Scale (EuroQol)	EuroQol Group (1990)
Religiosity	Religious Belief Scale	Frey and Daaleman (1999)
Religious well-being	Spiritual Well-Being Scale	Ellison (1983)
Spiritual well-being	Spiritual Well-Being Scale	Ellison (1983)

a. $N = 509$, reported in Daaleman and Frey (2004).
b. $N = 577$, reported in Frey et al. (2004).
c. $N = 277$, reported in Daaleman et al. (2002).
*$p < .05$. **$p < .01$. ***$p < .001$.

cators, and not at all with measures of religiosity. Table 2 presents a summary of reported correlations between the total SIWB scale and

subscale scores and other measures. There are large correlations of the SIWB with broad well-being measures (i.e., General Well-Being Scale), moderate correlations with more specific well-being measures with cognitive components (i.e., depression, hope, fear of death), and smaller correlations with physical functioning measures. Although there was a modest correlation of the SIWB with the Religious Well-Being subscale, correlations with a Religiosity Scale, which included *belief in God* and religious practices items, were small or nonsignificant.

DISTINGUISHING SPIRITUALITY CONSTRUCTS

Although these findings place our concept of spirituality within a nomological web, correlational evidence of construct validity evidence for the SIWB should go beyond a pattern of expected correlations with other measures. To be useful to researchers, SIWB scores should have stronger associations than other spirituality or religiosity scales with salient variables of interest. Using data from the Geriatric Outpatient Study and the Primary Care Study, we compared the correlations of the SIWB, SWB, and Religiosity Scale with other key well-being measures (Table 3). An analysis of these correlations found a pattern that further distinguishes the SIWB from other spirituality measures. In the Geriatric Outpatient sample, spirituality shared approximately six times as much variance with depression as did religious belief, and accounted for approximately four times as much variance with fear of death. The Primary Care Study included administration of the Spiritual Well-Being Scale (Ellison 1983), an instrument that provides a measure of overall perception of spiritual well-being. In a secondary analysis of this data set, we found that the SIWB total score was a stronger estimate of both depression and general well-being than the SWB, with SIWB total scores accounting for approximately twice the variance of both these constructs than did the Spiritual Well-Being Scale.

GENDER DIFFERENCES

The Catholic High School Study examined SIWB scores for gender differences. Female students scored about one point higher on the Life Scheme subscale than did male students, which was a statistically

TABLE 3
Comparisons of Selected Validity Coefficients

	Spirituality Index of Well-Being	*Religious Belief Scale*
Geriatric Outpatient Study, $N = 274$		
Depression	–.36*	–.15*
Fear of death	–.39*	–.20*
		Spiritual Well-Being Scale
Primary Care Study, $N = 494$		
Depression	–.42*	–.30*
General well-being	.64**	.48**

NOTE: All pairs of correlations in the same rows differ from each other, $p \leq .01$. Estimates of the proportion of variance shared between two variables can be obtained by squaring the correlation coefficient.
*$p < .05$. **$p < .01$.

significant difference but also resulted in a small effect size ($d = .18$). No gender differences were found on the Self-Efficacy subscale or the full SIWB scale scores. We performed secondary analyses of data from the Geriatric Outpatient and Primary Care Studies to further examine gender differences. Female geriatric outpatients scored higher on the Life Scheme subscale ($p = .004$, $d = .25$), and primary care outpatients reported gender differences in both Life Scheme and the full scale score. Female persons reported higher SIWB scores than male persons on the total score, $t(507) = 2.28$, $p = .02$, $d = .22$, a finding largely driven by gender difference on the Life Scheme subscale, $t(507) = 3.12$, $p = .002$, $d = .30$. In three study samples, female respondents report a significantly more coherent view of meaning in their lives than do male respondents across the life span. However, we found no gender differences in functional self-efficacy.

EVIDENCE FOR THE BIDIMENSIONAL NATURE OF HEALTH-RELATED SPIRITUALITY

The SIWB was designed to measure two components: self-efficacy and life scheme, and a model with two correlated factors reflecting this intent is shown as Figure 2. The Geriatric Outpatient Study, the

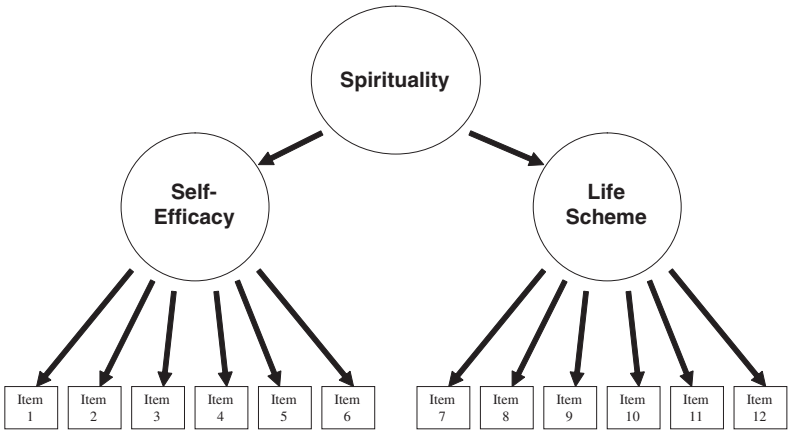

Figure 2: Two-Factor Model for Spirituality

Primary Care Study, and the Catholic High School Study all included results of exploratory factor analyses, and each concluded that two-factor solutions were appropriate for their study samples. The Primary Care Study also included a confirmatory factor analysis and reported fit indices generally supportive of a two-factor solution.

To examine the a priori theory-based model of two distinct but correlated factors (self-efficacy and life scheme), we pooled data from all three samples to conduct a confirmatory factor analysis with maximum likelihood estimation using EQS (Bentler 1995) for each of three age-groups separately (ages 14 to 18, 19 to 64, and 65 to 94). In these models, each item was specified with a single factor, the two factors were correlated, and the error variances associated with each item were uncorrelated. To begin, we examined the fit of this initial model for the 65- to 90-year-old group; the model yielded a significant chi-square, $\chi^2(53) = 82.71$, $p = .006$, a Comparative Fit Index (CFI) of .98, and a root mean square error of approximation (RMSEA) of .04. With a large number of observations, often the chi-square value is inflated producing statistically significant results, and therefore the chi-square ratio (χ^2 / *df*), which adjusts for model complexity and sample size, is reported. Although the cutoffs for interpreting this statistic vary, in general, a chi-square ratio between 1 and 3 indicates good fit (Arbuckle and Wothke 1999). The CFI and RSMEA are also used to

make judgments regarding the adequacy of the model. For the CFI, the closer the value is to 1.00, the better the model fit with cutoff values close to .95, indicating good fit. For the RMSEA index, the closer the value is to 0, the better the fit of the conceptual model to the empirical data; indices of .06 and smaller are considered good fit (Hu and Bentler 1999).

Next, the initial model was evaluated for the 19- to 64-year-olds. Consistent with the older group of respondents, the model fit the data adequately, χ^2 ratio = 3.13, CFI of .96, and the RMSEA was .07. Last, the correlated factor model was fit to the data for the youngest group (14- to 18-year-olds). In general, the fit indices were adequate to fair, χ^2 ratio = 5.68, CFI of .93, and the RMSEA was .09, giving rise to the possibility that the factor structure for this particular group is different than that of mature adults. To examine whether the factor structures were comparable for this group compared with the middle adult and older adult respondents, a multisample analytic approach was undertaken.

First, we analyzed the confirmatory factor models for the middle and older adult groups, estimating simultaneously the same pattern of relationships among indicators for these two samples, while eliminating the younger group data from this particular analysis. Using this multisample approach, equivalence among samples is evaluated by constraints that impose identical estimates for the model's parameters (Byrne 1994a, 1994b). Using EQS, the plausibility of these equality constraints is examined by the LaGrange multiplier test (Bentler 1995), which evaluates whether freeing up any constrained parameters would improve the model. For each of the constraints specified, the LaGrange multiplier test provides evidence that the constraint applies to the populations being studied.

To begin, a series of models was tested. The first model in this series allows for the factor loadings of each item, the covariance between the two factors, and each item's uniqueness (error variance) to differ between the middle and older groups. To proceed to the next set of steps in the modeling process, the multigroup fit indices need to be acceptable, and they were, χ^2 ratio = 2.35, CFI = .97, and RMSEA = .04. In a sequential succession of steps that follow, the factor loadings, the covariance between factors, and error variances were constrained to equality one at a time, respectively, and model fit was evaluated at each step. If none of the parameters differ between the groups, the fit

of the model across the series of steps will remain significantly unchanged and a case can be made for invariance across groups. This was found for our sample, χ^2 ratio = 2.86, CFI = .95, RMSEA = .05.

Next, we added the younger-group data into the analyses and evaluated the fit of the three-group multisample model, following the same procedures as with the two-group multisample approach. Allowing for the factor loadings of each item, the covariance between the two factors, and each item's error variance to differ between the middle, older, and younger group produced a good fit to the data, χ^2 ratio = 3.46, CFI = .95, RMSEA = .04. The final model, when all constraints were in place across the three groups, resulted in an adequate fit to the data, χ^2 ratio = 3.83, CFI = .93, RMSEA = .05. Examination of the LaGrange multiplier tests for the final model revealed that fit could be improved by releasing a few of the error variances across two or all of the groups, or by releasing a select few factor loadings across only two of the groups, but not the third. In the interest of parsimony and the lack of a significant improvement in fit in the three fit indices when any modifications were made, we chose to retain our final multisample model and not include any changes as suggested by the LaGrange multiplier tests.

In summary, the factor structure of the SIWB is consistent for the three groups, indicating that it measures the same components of self-efficacy and life scheme across the life span and that these two entities are correlated. Modification indices suggested further refinements to the initial model such as cross loading an item on both factors. However, none of these suggested modifications were supported across the three groups. For the sake of simplicity and because the fit of the individual group models as well as the multisample models were good, the initial two-factor correlated simple structure model was retained.

Limitations

CONSTRUCT LIMITATIONS

The face validity of the SIWB is one potential limitation in its usefulness. To some researchers, due to the absence of a religious component, the construct may appear to be misdefined or mislabeled. However, spirituality is a social construct, and a secular view of this

concept is well established in the literature across fields and has recently become more accepted in the healing professions (Van Ness 1996).

A second limitation in the validation of our conceptualization of spirituality is the absence of work that tests the conceptual framework (Figure 1). A variety of pathways, sequences, and relationships are proposed in the framework but lack empirical testing. The framework was developed from qualitative data, but its generalizability to a broader and more diverse patient population has not been examined. Theoretically, the variable has meaning even beyond a health context, as the life scheme and self-efficacy components have importance in education and other fields. The Catholic High School Study introduced the concept *cognitive spirituality* and argued that it is a construct that represents the spiritual nature of the whole person (Frey et al. 2004). However, examination of the usefulness of this concept outside of health has yet to be made.

A third validity concern arises from linking spirituality with two well-recognized social and psychological concepts: self-efficacy and life scheme. Studies have not been designed or conducted that evaluate the correlation of the SIWB and its subscales with established measures of those two constructs. Those seem to be, at a minimum, necessary evidence to accept the labels as appropriate and independent.

PSYCHOMETRIC LIMITATIONS

The SIWB is composed entirely of items in one direction, that is, negatively worded items. There are two potential concerns that arise when a scale is composed of items in this fashion. One concern is that response bias or a "response set" may occur (Guilford 1954; Nunnally 1978; Sax 1997). Respondents may answer very quickly after the first few items and mark similar answers on all items without considering the specific wording of items. Using items in both "directions" presumably forces respondents to respond in a more reflective and less biased fashion. A response set appears unlikely with the SIWB, as the scale is made up of two differently focused sections of only six items that are intrinsically interesting to respondents. In addition, the strong convergent and discriminant validity evidence presented earlier suggests that there is not much construct-irrelevant variance in scores. A

second potential concern with a scale with unidirectional items relates to the technical interpretations made when creating different subscales from a larger pool of items using factor analysis. Artifacts theory posits that from a pool of both negatively and positively worded items, negatively worded items will occasionally load onto a separate factor all by themselves (King n.d.; Marsh 1996; Schmitt and Stults 1985). This is problematic if a researcher interprets this artificial factor as a result of the nature of the construct itself and begins to define concepts and build theory on the basis of this measurement artifact alone. The SIWB was, a priori, conceptually developed from a qualitative study, and factor analysis was employed subsequently.

A second potential limitation of the SIWB is the use of two subscale scores as indicators of two different dimensions. Although the two-factor model worked very well across the age span, with items loading as expected, there was less good fit between the ages of 14 and 18. An alternative model with slightly better fit for that age group would allow for cross loadings for one or two items. This would not result in substantial improvement, however, and the parsimonious nature of a two-factor model, which offers two meaningful subscale scores and a total scale score, is advantageous. Gender differences found for one subscale in this age group, but not for the other subscale, as well as the theoretical differences in meaning between the two components, suggest that there is utility in using both subscale and full scale scores in research.

Conclusion

The SIWB appears to be a valid and reliable measure of spirituality. The instrument's conceptualization of spirituality is useful, simple, and grounded in a qualitative study of the patient perspective of spirituality and well-being. The scales perform with high reliability, and scores from the SIWB show associations with health and well-being constructs across different populations. In univariate and multivariate analyses, the SIWB was an independent predictor of self-reported good health status, even after adjusting for known covariates such as race and depression.

A confirmatory factor analysis with three large samples found good fit between empirically gathered data and the theoretical bidimensional definition of spirituality as consisting of two independent, but related, components—life scheme and self-efficacy. Gender differences have been found in subscale performance across the age span, with women displaying a higher sense of meaning in their lives. Limitations on the use of the scale include an absence of studies testing the conceptual framework in which this conceptualization of spirituality resides. In summary, the SIWB may be a useful research tool for investigators seeking to measure a dimension of spirituality in patient populations. It may be best situated in studies of chronic illness, aging, and end-of-life care that are inclusive of HRQOL and spirituality. Future validation studies with multiple populations and a longitudinal design are needed to refine, modify, or verify the SIWB as an additional, complementary instrument of well-being.

APPENDIX
Spirituality Index of Well-Being

Which response best describes how you feel about each statement?

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*
1	2	3	4	5

There is not much I can do to help myself.

1	2	3	4	5

Often, there is no way I can complete what I have started.

1	2	3	4	5

I can't begin to understand my problems.

1	2	3	4	5

I am overwhelmed when I have personal difficulties and problems.

1	2	3	4	5

I don't know how to begin to solve my problems.

1	2	3	4	5

There is not much I can do to make a difference in my life.

1	2	3	4	5

I haven't found my life's purpose yet.

1	2	3	4	5

I don't know who I am, where I came from, or where I am going.

1	2	3	4	5

I have a lack of purpose in my life.

1	2	3	4	5

In this world, I don't know where I fit in.

1	2	3	4	5

I am far from understanding the meaning of life.

1	2	3	4	5

There is a great void in my life at this time.

1	2	3	4	5

SOURCE: Copyright 2000, University of Kansas Medical Center.

REFERENCES

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 1999. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Antonovsky, Aaron. 1987. *Unraveling the Mystery of Health: How People Manage Stress and Stay Well*. San Francisco: Jossey-Bass.

Arbuckle, James L. and W. Wothke. 1999. *AMOS Users' Guide, Version 4.0*. Chicago: Small Waters.

Association of Professional Chaplains, Association for Clinical Pastoral Education, Canadian Association for Pastoral Practice and Education, National Association of Catholic Chaplains, and National Association of Jewish Chaplains. 2001. "A White Paper. Professional Chaplaincy: Its Role and Importance in Healthcare." *Journal of Pastoral Care* 55 (1): 81-97.

Bandura, Albert 1997. *Self-Efficacy, the Exercise of Control*. New York: Freeman.

Bentler, Peter M. 1995. *EQS Structural Equations Program Manual*. Encino, CA: Multivariate Software.

Byrne, Barbara M. 1994a. *Structural Equation Modeling with EQS and EQS Windows: Basic Concepts, Applications, and Programming*. Thousand Oaks, CA: Sage.

———. 1994b. "Testing the Factorial Validity, Replication, and Invariance of a Measuring Instrument: A Paradigmatic Application Based on the Maslach Burnout Inventory." *Multivariate Behavioral Research* 29:289-311.

Campbell, Donald T. and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81-105.

Canda, Edward R. and Leola Dyrud Furman. 1999. *Spiritual Diversity in Social Work Practice: The Heart of Helping*. New York: Free Press.

Coleman, J. A. 1997. "Exploding Spiritualities: Their Social Causes, Social Location, and Social Divide." *Christian Spirituality Bulletin* 5 (1): 9-15.

Cox, James L. 1996. *Expressing the Sacred: An Introduction to the Phenomenology of Religion*. 2d ed. Harare: University of Zimbabwe Publications.

Cronbach, Lee J. and Paul E. Meehl. 1955. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52:281-302.

Daaleman, Timothy P., A. K. Cobb, and Bruce B. Frey. 2001. "Spirituality and Well-Being: An Exploratory Study of the Patient Perspective." *Social Science and Medicine* 53:119-27.

Daaleman, Timothy P. and Bruce B. Frey. 2004. "The Spirituality Index of Well-Being: A New Instrument for Health-Related Quality of Life Research." *Annals of Family Medicine* 2 (5): 499-503.

Daaleman, Timothy P., Bruce B. Frey, Dennis Wallace, and Stephanie A. Studenski. 2002. "The Spirituality Index of Well-Being: Development and Testing of a New Measure." *Journal of Family Practice* 51:952.

Daaleman, Timothy P., S. Perera, and Stephanie A. Studenski. 2004. "Religion, Spirituality, and Health Status in Geriatric Outpatients." *Annals of Family Medicine* 2:49-53.

Ellison, Christopher W. 1983. "Spiritual Well-Being: Conceptualization and Measurement." *Journal of Psychology and Theology* 11:330-40.

Embretson, Susan. 1983. "Construct Validity: Construct Representation Versus Nomothetic Span." *Psychological Bulletin* 93 (1): 179-97.

Erickson, P., R. Wilson, and I. Shannon I. 1995. *Years of Healthy Life*. Statistical notes, No. 7. Hyattsville, MD: U.S. Department of Health and Human Services, CDC, National Center for Health Statistics.

EuroQol Group. 1990. "EuroQol—A New Facility for the Measurement of Health-Related Quality of Life." *Health Policy* 16:199-208.

Fetzer Institute and National Institute on Aging. 1999. *Multidimensional Measurement of Religiousness/Spirituality for Use in Health Research*. Kalamazoo, MI: John A. Fetzer Institute.

Frey, Bruce B. and Timothy P. Daaleman. 1999. "Toward a Patient-Centered Measure of Spirituality." *The Behavioral Measurement Letter* 6 (2): 2-4.

Frey, Bruce B., Jennifer T. Pedrotti, Lisa M. Edwards, and Diane McDermott. 2004. "Cognitive Spirituality and Hope in Catholic High School Students." *Catholic Education: A Journal of Inquiry and Practice* 7 (4): 479-91.

George, Linda K., David B. Larson, Harold G. Koenig, and Michael E. McCullough. 2000. "Spirituality and Health: What We Know, What We Need to Know." *Journal of Social and Clinical Psychology* 19:102-16.

Guilford, J. P. 1954. *Psychometric Methods*. New York: McGraw-Hill.

Hill, Peter C. and Ralph W. Hood, eds. 1999. *Measures of Religiosity*. Birmingham, AL: Religious Education Press.

Hu, Li-tze and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6:1-55.

Idler, Ellen L., Marc A. Musick, Christopher G. Ellison, Linda K. George, Neal Krause, Marcia G. Ory, Kenneth I. Pargament, Lynda H. Powell, Lynn G. Underwood, and David R. Williams. 2003. "Measuring Multiple Dimensions of Religion and Spirituality for Health Research." *Research on Aging* 25 (4): 327-65.

Jarvis, George K. and Herbert C. Northcott. 1987. "Religion and Differences in Morbidity and Mortality." *Social Science and Medicine* 25:813-24.

King, Craig V. N.d.. *Factor Analysis and Negatively Worded Items*. Retrieved August 14, 2003, from http://www.populus.com/techpapers/download/faand_neg_worded.pdf

Kleinman, A., L. Eisenberg, and B. Good. 1978. "Culture, Illness, and Care, Clinical Lessons from Anthropologic and Cross-Cultural Research." *Annals of Internal Medicine* 88:251-58.

Koenig, Harold G. 2002. "An 83-Year Old Woman with Chronic Illness and Strong Religious Beliefs." *Journal of the American Medical Association* 288 (4): 487-93.

Koenig, Harold G., Michael E. McCullough, and David B. Larson. 2001. *Handbook of Religion and Health.* New York: Oxford University Press.

Levin, J. S. and H. Y. Vanderpool. 1987. "Is Frequent Religious Attendance Really Conducive to Better Health? Toward an Epidemiology of Religion." *Social Science & Medicine* 24 (7): 589-600.

Marsh, Herbert W. 1996. "Positive and Negative Global Self-Esteem: A Substantively Meaningful Distinction or Artifacts?" *Journal of Personality and Social Psychology* 70 (4): 810-19.

McDowell, I. and C. Newell. 1996. *Measuring Health: A Guide to Rating Scales and Questionnaires*. New York: Oxford University Press.

McGinn, Bernard. 1993. "The Letter and the Spirit: Spirituality as an Academic Discipline." *Journal of the Society for the Study of Christian Spirituality* 1:1-10.

Musick, Marc A. 1996. "Religion and Subjective Health Among Black and White Elders." *Journal of Health and Social Behavior* 37:221-37.

National Center for Health Statistics. 2002. *Health, United States with Chartbook on Trends in the Health of Americans.* Hyattsville, MD: Department of Health and Human Services.

Nunnally, Jum C. 1978. *Psychometric Theory.* 2d ed. New York: McGraw-Hill.

Ory, Marcia G. and Paula Darby Lipman. 1998. *Religion and Spirituality in Aging and Health*. Rockville, MD: National Institute on Aging, Behavioral and Social Research Programs.

Pargament, Kenneth I. 1997. *The Psychology of Religion and Coping: Theory, Research, Practice*. New York: Guilford.

Pargament, Kenneth I. and A. Mahoney. 2002. "Spirituality: Discovering and Conserving the Sacred." Pp. 646-59 in *Handbook of Positive Psychology*, edited by C. Richard Snyder and Shane J. Lopez. New York: Oxford University Press.

Roof, Wade C. 1993. *A Generation of Seekers: The Spiritual Journeys of the Baby Boom Generation*. San Francisco: Harper.

Rumsfeld, John S. 2002. "Health Status and Clinical Practice, When Will They Meet?" *Circulation* 106: 5-7.

Sax, Gilbert. 1997. *Principles of Educational and Psychological Measurement and Evaluation*. London: Wadsworth.

Schmitt, N. and D. M. Stults. 1985. "Factors Defined by Negatively Keyed Items: The Result of Careless Respondents?" *Applied Psychological Measurement* 9 (4): 367-73.

Seligman, Martin E. P. 2002. "Positive Psychology, Positive Prevention and Positive Therapy." Pp. 3-9 in *Handbook of Positive Psychology*, edited by C. Richard Snyder and Shane J. Lopez. New York: Oxford University Press.

Shea, John. 2000. *Spirituality and Health Care, Reaching Toward a Holistic Future*. Chicago: The Park Ridge Center.

Sloan, Richard P., E. Bagiella, and T. Powell. 1999. "Religion, Spirituality, and Medicine." *The Lancet* 353:664-667.

Snyder, C. Richard, B. Hoza, W. E. Pelham, M. Rapoff, L. Ware, M. Danovsky, L. Highberger, H. Rubinstein, and K. J. Stahl. 1997. "The Development and Validation of the Children's Hope Scale." *Journal of Pediatric Psychology* 22:399-421.

Snyder, C. Richard and Shane J. Lopez, eds. 2002. *Handbook of Positive Psychology.* New York: Oxford University Press.

Stewart, A. L., R. D. Hays, and J. E. Ware Jr. 1988. "The MOS Short-Form General Health Survey: Reliability and Validity in a Patient Population." *Medical Care* 26 (7): 724-35.

Testa, Marcia A. and Donald C. Simonson. 1996. "Assessment of Quality-of-Life Outcomes." *New England Journal of Medicine* 334:835-40.

Thompson, Bruce, ed. 2003. *Score Reliability.* Thousand Oaks, CA: Sage.

Thompson, Suzanne C. and A. S. Janigian. 1988. "Life Schemes: A Framework for Understanding the Search for Meaning." *Journal of Social and Clinical Psychology* 7:260-80.

Troyer, H. 1988. "Review of Cancer Among Four Religious Sects: Evidence That Lifestyles Are Distinctive Sets of Risk Factors." *Social Science and Medicine* 26:1007-17.

Van Ness, Peter H., ed. 1996. *Spirituality and the Secular Quest.* New York: Crossroad.

Walker, L. O. and K. C. Avant. 1995. *Strategies for Theory Construction in Nursing*. 3d ed. Englewood Cliffs, NJ: Prentice Hall.

Wilson, J. 1963. *Thinking With Concepts*. New York: Cambridge University Press.

Wilson, I. B. and P. D. Cleary. 1995. "Linking Clinical Variables With Health-Related Quality of Life." *Journal of the American Medical Association* 273 (1): 59-64.

Wong, Paul T. P., Gary T. Reker, and Gina Gesser. 1994. "Death Attitude Profile–Revised: A Multidimensional Measure of Attitudes Toward Death." Pp. 121-48 in *Death Anxiety Handbook: Research, Instrumentation, and Application*, edited by R. A. Neimeyer. Washington, DC: Taylor & Francis.

Wulff, David M. 1997. *Psychology of Religion, Classic and Contemporary.* 2d ed. New York: John Wiley.

Wuthnow, Robert. 1998. *After Heaven: Spirituality in America Since 1950*. Berkeley: University of California Press.

Yesavage, J. A., T. L. Brink, T. L. Rose, O. Lum, V. Huang, M. Adey, and V. O. Leirer. 1982. "Development and Validation of Geriatric Depression Screening Scale: A Preliminary Report." *Journal of Psychiatric Research* 17:37-49.

Zinnbauer, Brian J., Kenneth I. Pargament, Brenda C. Cole, Mark S. Rye, Eric M. Butter, Timothy G. Belavich, Kathleen M. Hipp, Allie B. Scott, and Jill L. Kadar. 1997. "Religion and Spirituality: Unfuzzying the Fuzzy." *Journal for the Scientific Study of Religion* 36 (4): 549-64.

Zung, W. W. K. 1965. "A Self-Rating Depression Scale." *Archives of General Psychiatry* 12:63-70.

Bruce B. Frey, Ph.D., is an assistant professor in psychology and research in education in the School of Education at the University of Kansas. His research interests include cognitive spirituality, classroom assessment, and program evaluation methods.

Timothy P. Daaleman, D.O., M.P.H., is an associate professor in the Department of Family Medicine and Research Fellow in the Program on Aging, Disability, and Long-Term Care in the Cecil G. Sheps Center for Health Services Research at the University of North Carolina at Chapel Hill. His primary research interests include religion, spirituality, and medicine, and the social and cultural influences on end-of-life care and chronic illness.

Vicki Peyton, Ph.D., is an assistant professor in the Psychology and Research in Education Department in the School of Education at the University of Kansas. Her research interests include multi-level statistical modeling methods and differential item functioning in academic and attitude measures.