2-1989

An Evaluation of the Spiritual Well-Being Scale: Reliability and Response Measurement

Daniel Dale Brinkman

Order Number 8917593

An evaluation of the Spiritual Well-Being Scale: Reliability and response measurement

Brinkman, Daniel Dale, Psy.D.

Western Conservative Baptist Seminary, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

An Evaluation of the Spiritual Well-Being Scale: Reliability and Response Measurement

by

Daniel Dale Brinkman

Presented to the Faculty of
Western Conservative Baptist Seminary
in partial fulfillment
of the requirements for the degree of
Doctor of Psychology
in Clinical Psychology

Portland, Oregon
February 28, 1989

APPROVAL

An Evaluation of the Spiritual Well-Being Scale:

Reliability and Response Measurement

by

Daniel Dale Brinkman

Signatures:

Rodger K. Bufford, Ph.D.
Committee Chairman

Vice President for
Academic Affairs

Date: April 24, 1989

Neal F. McBride, Ed.D., Ph.D.
Committee Member

Norman L. Thiesen, Ph.D.
Committee Member

Date: 31 March 1989

An Evaluation of the Spiritual Well-Being Scale:
Reliability and Response Measurement

Western Conservative Baptist Seminary
Portland, Oregon
Daniel Dale Brinkman

ABSTRACT

The Spiritual Well-Being Scale (SWB) is a 20 item self-report attitudinal measure of one's religious and existential well-being. It is the most extensively studied of the instruments developed to measure spiritual well-being (Moberg, 1986). Despite the popularity of this scale, it is still in the process of research and development. This study evaluated research done with the SWB scale considering the Standards for Educational and Psychological Testing (1985). While there has been much progress, two needs that became clear from this examination were for additional studies in the area of reliability and to try to "raise the ceiling" of the scale because scores tend to cluster near the maximum, especially for highly religious populations.

Censored scores are undesirable because they limit interpretation and practical use.

This dissertation conducted three studies. The first study designed a new rating scale with more opportunity for score variability. This study tested the new scale with evangelical Christians looking at correlations with other religious measures and comparing it with the original rating scale. The second study investigated test-retest and internal consistency reliability coefficients for both versions of the rating scale. The second study used two samples: community college students and a Baptist church. Finally, a third study examined internal consistency coefficients and other descriptive data from previously collected data using the original scale. The three samples in this study were (a) Conservative Baptists in Oregon, (b) a church in Salem, OR, and (c) Chinese American Christians in the Northwest.

Results supported the initial reliability studies for the scale, suggesting the scale is satisfactory in this area. The experimental scale was equivalent in many respects to the original including correlations with other measures, reliability coefficients, and

measures of central tendency and variability. It did not alter the basic shape of the score distribution enough to be of practical significance. There is a need for more research in this area so the scale can discriminate among persons in highly religious populations.

ACKNOWLEDGEMENTS

I would like to express my appreciation to the members of my dissertation committee, to Rodger Bufford for his guidance and enthusiasm in this project; to Neal McBride for his commitment to excellence and creations with his laser printer; and to Norm Thiesen for his interest in measurement instruments and willingness to serve on my committee.

I would also like to thank my parents, Dale and Anita Brinkman, for always believing in me; my wife, Robin, for allowing me to pursue my dreams and goals; and my daughter, Jacqueline, for her patience and for letting me use the computer once in a while.

Finally, I would like to thank God for His provision and protection throughout my life.

DEDICATION

To my grandfather, the Rev. John C. Procter.

TABLE OF CONTENTS

Approval ii
Abstract iii
Acknowledgments vi
Dedication vii
Table of Contents viii
List of Tables xiii
List of Figures xv
Chapter 1: INTRODUCTION 1
Statement of the Problem 4
Related Literature 5
Development of the Spiritual Well-Being Scale 5
Evaluation of the Spiritual Well-Being Scale 13
Validity 16
Content Validity 19
Construct Validity 21
Criterion Validity 47
Reliability 51
Test Development and Revision 59
Scaling and Norming 64
Test Publication 70
Summary 70

Reliability 71
Sources of Variance 74
Models of Reliability 81
Test-Retest 81
Internal Consistency 83
Standard Error of Measurement 85
Using the Reliability Estimate 90
Interpreting Reliability Coefficients 92
Increasing Reliability 93
Summary 95
Response Measurement 97
Measurement Theory 97
Measuring Attitudes 103
Models for Scaling Attitudes 104
Guttman Scales 105
Thurstone Scales 107
Semantic Differential Scales 110
Likert Scales 112
Properties of Rating Scales 117
Anchor Words 117
Number of Categories 119
Midpoints 121
Response Sets 122
Instructions 123

General Appearance 124
Summary 125
Purpose of the Study 128
Chapter 2: METHODS 132
Study One 133
Participants 133
Research Instruments 134
Spiritual Well-Being Scale (Form B)..... 134
Concept of God scale 138
Spiritual Maturity Index 140
Religious Orientation Scale 143
Single Item Measures 147
Procedures 147
Study Two 151
Participants 151
Research Instruments 152
Spiritual Distress Scale 153
Demographic Questionnaire 156
Procedures 157
Study Three 162
Participants 163
Research Instruments 164
Procedures 164
Chapter 3: RESULTS 166

Study One 166
Study Two 174
Study Three 192
Descriptive Statistics 195
Reliability 204
Intratest 204
Test-Retest 206
Internal Consistency 208
Response Measurement 210
Chapter 4: DISCUSSION 215
Construct Validity 222
Reliability 226
Response Measurement 229
Contributions 231
Suggestions for Future Research 233
References 236
Appendices 257
Appendix A Spiritual Well-Being Scale (Original Version) 257
Appendix B Spiritual Well-Being Scale (Test Version) 260
Appendix C Concept of God Scale 263
Appendix D Spiritual Distress Scale 267
Appendix E Spiritual Maturity Index 270

Appendix F Religious Orientation Scale 274
Appendix G Single Item Measures (Study One) ... 279
Appendix H Demographic Questions (Study Two) .. 281
Appendix I Research Announcement (Study One) .. 284
Appendix J Instructions (Study One) 286
Appendix K Research Announcements (Study Two) . 288
Appendix L Instructions (Study Two) 292
Appendix M Handout (Study Two) 294
Appendix N Raw Data 296
Appendix O Vita 307

LIST OF TABLES

Table 1 Types of Validity 18

Table 2 Correlations Between the Spiritual Well-Being Scale and the Spiritual Maturity Index 35

Table 3 Correlations Between the Spiritual Well-Being Scale and the Religious Orientation Intrinsic Scale 36

Table 4 Correlations Between the Spiritual Well-Being Scale and the Rotter Locus of Control Scale 39

Table 5 Correlations Between the Spiritual Well-Being Scale and the Tennessee Self-Concept Scale 40

Table 6 Correlations Between the Spiritual Well-Being Scale and the Religious Orientation Extrinsic Scale..... 44

Table 7 Intratest Correlations for the Spiritual Well-Being Scale 55

Table 8 Test-Retest Correlations for the Spiritual Well-Being Scale From Upshaw (1986) .. 58

Table 9 Spiritual Well-Being Scale Descriptive Statistics 66
Table 10 Sources of Test Score Variance 78
Table 11 Descriptive Data for Study One Measures . 168
Table 12 Correlations Between SWB Scales and Data from Study One 172
Table 13 Demographic Data from Study Two 177
Table 14 Descriptive Data for Study Two Measures . 184
Table 15 Correlations Between SWB Scales and Data from Study Two 189
Table 16 Demographic Data from Study Three Samples 193
Table 17 Descriptive Statistics for the Original SWB Scale 196
Table 18 Descriptive Statistics for the Test SWB Scale 200
Table 19 Intratest Correlations for the SWB Scale 205
Table 20 Test-Retest Reliability Coefficients for SWB Scale 207
Table 21 Internal Consistency Alphas for SWB Scale 209

LIST OF FIGURES

Figure 1 Score Distribution for Original SWB Scale--Study One 211

Figure 2 Score Distribution for Test SWB Scale--Study One 212

Figure 3 Score Distribution for Original SWB Scale--Study Two 213

Figure 4 Score Distribution for Test SWB Scale--Study Two 214

CHAPTER 1

INTRODUCTION

In the past few years there has been an increasing amount of interest in religious, or "spiritual," phenomenon and their place in the evaluation and treatment of the total person (Moberg, 1979). With this attention has come a demand for defining and measuring these phenomena, in ways that meet the criteria of good research. Good research relies on observation which is objective and understood, that uses intelligible and reproducible data (Spilka, Hood, & Gorsuch, 1985). According to Moberg (1986), there is a great need for empirical research on and related specifically to spiritual well-being in both the pure and applied aspects of many disciplines.

As Spilka, Hood, and Gorsuch (1985) have noted, there are many problems encountered when researching in the religious domain. One of the most fundamental of these obstacles is in defining what "religion" is. However, Ellison (1983) maintains that if the sciences can tolerate validity problems inherent to other

unobservable phenomenon, such as personality, attitudes, emotions, and intelligence, then religion and spirituality should pose no greater threat to the scientific method.

Moberg (1984) defends the use of measurement instruments to study spiritual well-being. Moberg surmised that without reliable tools any evaluation of efforts to promote spiritual well-being would remain on the level of nonrepresentative illustrations, philosophical arguments, theological exhortations, common sense folk wisdom, and careless "trial and error" experimentation rather than systematically tested conclusions.

The Spiritual Well-Being Scale (SWB) created by Craig Ellison and Ray Paloutzian in the late 70's is one of the measurement instruments developed from this interest in religion. Since then, this scale has generated much interest and has been included in over 50 studies (Moody, 1988). Even so, it must be remembered the SWB scale is not available for general use and is still in the process of research and development.

In summarizing the studies done with the SWB scale up until then, Paloutzian (1982) identified the need

for continued research in two areas: (a) more work on the scale itself, and (b) more work on how spiritual well-being interacts with other psychological and sociological variables. Most of the studies conducted since then have focused on the SWB scale's relationship with other variables and not on the scale itself.

During the process of examining the research with the SWB scale, the need for more information in three areas became evident to this researcher. The first need is for a comprehensive evaluation of the research that has been done on and with the SWB scale. A second need is reliability. Initial reliability studies were encouraging but not sufficient to satisfy professional standards for test reliability.

A third area in which the scale is weak are the scores assigned to individuals completing the scale when compared with a normal distribution. Research studies reveal a low ceiling for the scale--especially for respondents who identify themselves as highly religious. Correlations are reduced and interpretations weakened with scores that tend to cluster towards the top end of the scale and thus have a skewed distribution (Gravetter & Wallnau, 1988). It is also difficult to know what the true score of

individuals who receive the highest score would be since more room at the top is not available with the current response scale.

Statement of the Problem

The purpose of this study is to evaluate the research conducted on and with the Spiritual Well-Being Scale (SWB) and to contribute additional research in two areas: reliability and response measurement. Specifically, there are three research questions to be explored. The first question involves research on and with the SWB scale. Can a system of evaluation be devised that presents the research conducted with the SWB scale in a manner that is understandable and allows for comparison with some standards?

The second research question focuses on the issue of reliability. Can additional reliability coefficients be generated that are consistent with the original studies and defensible by professional standards? Evidence of an instruments's reliability include test-retest coefficients, internal alphas, intratest correlations, and standard errors of measurement.

The third and final research question concerns a specific problem with the SWB scale. Can the rating scale of the SWB scale be modified to minimize ceiling effects and produce scores approximating a normal distribution? Presently, the current rating scale has limited use with highly religious populations because scores cluster near the maximum. This is a problem since most of the research with this scale has studied these populations.

Related Literature

In reviewing the literature, four areas will be addressed: (a) the background and development of the Spiritual Well-Being Scale, (b) research conducted with the scale compared to criteria for psychological tests, (c) reliability, and (d) response measurement.

Development of the Spiritual Well-Being Scale

Ellison (1983) traces attempts by researchers to measure the subjective well-being of American people back to 1960 and a national survey of happiness, worries, and experiences conducted by Gurin, Veroff, and Feld. Early studies in this area focused mainly on

economic indicators, but the emphasis gradually shifted to include subjective non-economic signs as valid and essential factors in measuring well-being.

In addition, many of the well-being measures developed during the 1970's involved objective indicators and did not assess the internal feelings or perceptions of respondents. For the most part, the early studies in this area ignored or glossed over the spiritual dimension. This occurred despite growing evidence of an upsurge in religiosity in America and in the numbers of people who said religion played an important role in how they lived and experienced life (Ellison & Paloutzian, 1979).

For instance, the sociologist Angus Campbell (1981) postulated that well-being is dependent on the satisfaction of three basic needs: the need for having, the need for relating, and the need for being. While acknowledging these needs were important and necessary, Ellison (1983) believed Campbell and similar minded colleagues had ignored a fourth need: the need for transcendence. In Ellison's thinking, transcendence refers to the sense of well-being one experiences when he or she finds a purpose or purposes to commit themselves to which involves the ultimate meaning for

life. Transcendence includes a nonphysical dimension of awareness and experience which can best be termed "spiritual." All the great religions of the world acknowledge transcendence and call humans to this as the path to the highest levels of well-being.

While some researchers were choosing to ignore the spiritual dimension of well-being, others were not as reluctant to acknowledge its presence and began to develop some theoretical and empirical parameters (Ellison, 1983). The first step in this process was to try to develop an operational definition for spiritual well-being which would "specify the terms and describe the contours of the phenomenon in empirically based, measurable language" (Ellison, 1983, p. 331). This was a difficult task because operational definitions can never sufficiently describe a phenomenon. Another complication is that the construct of spiritual well-being has many facets and possible interpretations. In selecting their operational definition of spiritual well-being, the authors attempted to capture a quantitative, denotative meaning which would allow for systematic observation and public verification (Ellison, 1982a).

One definition of spiritual well-being proposed by the National Interfaith Coalition on Aging states: "spiritual well-being is the affirmation of life in a relationship with God, self, community, and environment that nurtures and celebrates wholeness" (NICA, 1975, p. 1). The NICA definition was a starting point for Paloutzian and Ellison. It suggested to them that there is a religious and social-psychological component to spiritual well-being. Religion was not presumed to be synonymous with spiritual well-being, but as a component and indicator of it (Moberg, 1984).

This definition was also consistent with David Moberg's (1971) concept of spiritual well-being as two faceted, having both vertical and horizontal components that interrelate with each other and with other areas of well-being. The vertical dimension refers to one's sense of well-being in relation to God while the horizontal dimension refers to a sense of life purpose and life satisfaction, with no reference to anything specifically religious.

With these definitions as a framework, the authors made additional clarifications and distinctions. They postulated that spiritual well-being may not be the same as spiritual health. Spiritual well-being rises

from an underlying state of spiritual health and is an expression of it. This distinction freed them from the "burden of trying to exactly or empirically measure the inner contours of one's spirit--a task which is most likely impossible" (Ellison, 1983, p. 332). Spiritual well-being refers to a psychological-experiential dimension, while spiritual health would more likely be defined according to the accepted creeds and codes of a religious body (Paloutzian, 1982).

Spiritual well-being is different from spiritual maturity, though these two should influence each other. In their view, one does not necessarily have to be spiritually mature to experience a sense of spiritual well-being.

Ellison and Paloutzian also felt the relationship between feeling psychologically healthy and spiritually healthy should be viewed as bidirectional, as being intricately intertwined. Finally, spiritual well-being was operationally defined as a continuous, rather than a dichotomous, variable. Spiritual well-being is not a matter of whether one has it, but how much one has (Ellison, 1983).

Paloutzian (1982) summarizes his thoughts on spiritual well-being by stating SWB is not synonymous

with health nor maturity, but instead refers to that sense of well-being which is a consequence of focusing one's attention beyond oneself. As such, it can take on both religious and non-religious forms.

The SWB scale was initially introduced as a 15 item instrument. Nine items made up the religious well-being subscale and six items the existential well-being subscale, with a four point answering format. After administering the scale to 115 women a factor analysis was done. Results from this pilot study suggested a need for more testing and revision. The authors dropped poorly worded items, added new items, and adopted a six point response scale.

This revision led to the current version of the Spiritual Well-Being Scale, originally called version 2. The revised version consists of 20 items. Ten of the items are designed to measure Religious Well-Being (RWB) based on Moberg's vertical dimension. RWB refers to one's sense of well-being in relationship to God. The items are carefully worded so respondents are free to interpret the word "God" in any way they choose (Paloutzian & Ellison, 1979a). The other ten items measure Existential Well-Being (EWB), or Moberg's horizontal dimension. EWB refers to a sense of life

purpose and life satisfaction, with no reference to God or anything else specifically religious. To have a sense of existential well-being is "to know what to do and why, who we are, and where we belong in relation to ultimate concerns" (Ellison, 1982a, p. 6).

The Spiritual Well-Being Scale is an objective, self-report, attitudinal survey. Examinees respond to statements on a six point Likert scale ranging from strongly agree to strongly disagree, and each response receives a numerical value from 1 to 6. Higher numbers suggest more well-being. The response categories are: Strongly Agree (1), Moderately Agree (2), Agree, (3), Disagree (4), Moderately Disagree (5), and Strongly Disagree (6). To avoid an acquiescence response bias half of the responses are negatively worded. A midpoint in response options was omitted to discourage neutral responses. The SWB scale yields three scores: (a) a total SWB score for all twenty items, (b) a score for the ten religious well-being (RWB) items, and (c) a score for the ten existential well-being (EWB) items.

Ellison (1982b) highlights eight features of the scale. The first is all the items deal with transcendent concerns or those aspects of our

experience which involve meaning, ideals, faith, commitment, purpose in life, and relationship to the divine. This means the only type of well-being measured is spiritual. Second, responses to the items convey personal experience. It is not a measure of belief, doctrinal correctness, ideology, or values. Instead, it is a measure of the tone of one's inner, subjective life.

Third, the items refer to feelings of satisfaction, affect, purpose and meaning, and a sense of being valued. According to Ellison, these are commonly accepted indicators of well-being and intrapersonal health. Fourth, the scale is multi-dimensional because it allows for a general measure of spiritual well-being while also differentiating between religious and existential well-being. Fifth, the scale allows measurement of spiritual well-being as a continuous, quantifiable variable. For each item six responses are available. This enables comparisons with other measures and scientific study as it approximates interval level data.

Sixth, while the scale arises from the Judeo-Christian view of religious well-being, it is

non-sectarian and can be used across Catholic, Protestant, Jewish and other religions which conceive of God in personal terms. Seventh, the scale provides a general measure of spiritual well-being while not getting bogged down in specific theological issues or a priori standards of well-being which may vary from one religious belief system or denomination to another. Finally, the scale is short and easy to use. It is therefore not expensive to administer or score. It can readily be used individually in counseling, within the church, or in research.

The following section examines the initial studies with the Spiritual Well-Being Scale, along with subsequent research. This section will attempt to provide a framework and a context for evaluating the SWB scale.

Evaluation of the Spiritual Well-Being Scale

The governing bodies of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education have compiled technical standards for evaluating the effectiveness of tests (Standards for Educational and Psychological Tests and Manuals, 1966).

The introduction to the 1966 edition states:

> Psychological and educational tests are used in arriving at decisions which may have great influence on the ultimate welfare of the persons tested, on educational points of view and practices, and on development and utilization of human resources. Test users, therefore, need to apply high standards of professional judgment in selecting and interpreting tests, and test producers are under obligation to produce tests which can be of the greatest possible service. The test producer, in particular, has the task of providing sufficient information about each test so that users will know what reliance can safely be placed on it. (p. 1)

This manual asserts that these standards should cover not only tests as narrowly defined, but also most published devices for diagnosis, prognosis, and evaluation. Included are interest related clinical techniques, tests of aptitude or ability, and achievement tests.

There are many benefits to having standardized measures as opposed to personal judgments or other subjective appraisals when trying to measure something.

According to Nunnally (1978), there are five distinct advantages. The first one is objectivity; measurement takes the guesswork out of scientific observation, allowing the data to be independently verified by another person. The second advantage is quantification. Numerical data makes it possible to report results in finer detail and to use methods of mathematical analysis.

Third, standardized measures allow for efficient communication among researchers. Researchers are able to build on past learning and compare their findings with others. The fourth advantage is economy. Once developed, standardized measures are more economical of time and money than subjective evaluations, and free highly trained professionals for other work. Finally, standardized measures allow for scientific generalization, helping the process of hypothesis testing and the formulation of scientific principles and laws.

Moberg (1986) believes there is a great need for research on and related to spiritual well-being. Since research is a major component in "the language of science" this should attract the attention of even skeptics and agnostics and open the subject for further

investigation. Thus, there are many reasons the Spiritual Well-Being Scale should be subject to the same criteria and standards as other instruments.

The Standards for Educational and Psychological Tests and Manuals (1985) specify five areas of technical standards for test construction and evaluation to consider when creating or evaluating a test: (a) validity; (b) reliability and errors of measurement; (c) test development and revision; (d) scaling, norming, score comparability, and equating; and (e) test publication: technical manuals and user's guides. The following section will examine the Spiritual Well-Being Scale using the guidelines from these five areas.

Validity

According to the Standards, validity is the most important consideration in test evaluation. Validity refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Test validation is the process of accumulating evidence to support such inferences. Although evidence may be accumulated in many ways, validity always refers to the degree to which that evidence supports the inferences made from the scores. The inferences

regarding specific uses of a test are validated, not the test itself (Messick, 1975).

There are a variety of ways to accumulate evidence for validity. Validity is traditionally grouped under the categories of content, construct, and criterion related validity. However, the use of these categories does not imply there are distinct types of validity because rigorous distinctions between the categories are not possible. Table 1 presents a summary of the different types of validity used for this study.

The Standards state that evidence of validity should be presented for the major types of inferences the test is recommended for along with a rationale to support that evidence. Validity which has not been investigated should be noted when it could affect interpretation of the test. Validity relating to the subscales and the procedures, sample composition, and any factors which may influence validity should also be reported. The evidence presented below for the SWB scale is separated into different categories for organizational purposes. These categories are not rigid but overlap with the other areas.

Table 1

Types of Validity

Content
Face
Logical
Construct
Developmental Changes
Correlations with Other Tests in Same Domain
Factor Analysis
Internal Consistency
Convergent
Discriminant
Experimental Intervention
Known-Group Differences
Criterion
Concurrent
Predictive

Note. Sources: Aiken (1979), Allen & Yen (1979), Anastasi (1988), D. Mueller (1986).

Content Validity. Content validation involves the systematic examination of test content to determine whether it covers a representative sample of the behavior domain being measured (Kaplan & Saccuzzo, 1982). Content validity is the only type for which evidence is logical rather than statistical. This type of validity is more commonly sought in achievement tests and built into a test from the outset through the choice of appropriate items. In Anastasi's (1988) opinion, primary reliance on content validation is usually inappropriate for aptitude and personality tests and may be misleading. Although considerations of relevance and content enter into the initial stages of constructing any test, eventual validation of aptitude or personality tests require empirical verification by other types of validity.

Anastasi continues by saying personality tests are not based on a specified course of instruction or uniform set of prior experiences from which test content can be drawn. Because of this, individuals are more likely to vary in their psychological processes employed in responding to the same test items, thus measuring different functions. The content of aptitude and personality tests can do little more than reveal

the hypotheses that led the test constructor to choose a certain type of content for measuring a specified trait.

Allen & Yen (1979) identify two types of content validity: face validity and logical validity. A test has face validity when an examination of the items leads to the conclusion that the items are measuring what they are supposed to be measuring. Ellison (1983) maintains the SWB scale has good face validity. A closer examination of the item statements raises some questions. Some of the SWB scale items, particularly those containing the word "believe," suggest that a cognitive component, or stable belief, is being measured and not the experience of well-being (Brinkman, Capes, Kunkel, & Tackett, 1988).

Concerning logical validity, the Standards state that when content-related evidence serves as a significant demonstration of validity for a particular test use, a clear definition of the universe represented, its relevance to the proposed test use, and the procedures followed in generating test content to represent that universe should be described. When using subject matter experts in this process,

qualifications should be listed and the procedure used to obtain a consensus reported.

D. Mueller (1986) agrees with Anastasi's assertion that content validity applies less well to the measurement of affective traits. According to Mueller, this results from the difficulty in circumscribing the "universe" of a psychological construct. With spiritual well-being it would be hard to identify all the possible positive and negative affective statements an individual could have towards spiritual well-being. Therefore, this type of validity, which has no statistical index, can only be documented as a process.

Little is known about the selection process that went into choosing the items for the Spiritual Well-Being Scale. The early manuscripts by Ellison and Paloutzian do not discuss item selection. Given the above considerations, this is not a critical issue when accumulating evidence of validity for the SWB scale.

Construct Validity. Construct validity is the most general type of validity in that it incorporates evidence from studies of the content and criterion related validity of a test (Aiken, 1979). The construct measured should be embedded in a conceptual framework which specifies the meaning of the construct,

distinguishes it from other constructs, and suggests how measures of the construct should relate to other variables (Standards, 1985).

The construct validity of an instrument is the extent to which one can be sure it represents the construct whose name appears in its title (Henerson, Morris, & Fitz-Gibbon, 1978). According to Messick (1975), the concern of construct validity is not to explain a single behavior or item response but to account for consistency in behaviors or item responses which often have a small number of determinants and sometimes a major one.

When a test is proposed as a measure of a construct the proposed interpretation of the test score should be explicitly stated, and evidence presented to support such inferences. Evidence should demonstrate the test scores are more closely associated with variables of theoretical interest than with variables not included in the theoretical network. This also applies to any subscales.

Anastasi (1988) lists six specific techniques which contribute to construct validation: developmental changes, correlations with other tests, factor analysis, internal consistency, convergent and

discriminant validation, and experimental interventions. D. Mueller (1986) adds an additional category: known-group differences.

Regarding developmental changes, the technique of age differentiation does not apply to any construct not exhibiting clear-cut and consistent age changes. In the area of personality measurement age differences have found limited use (Anastasi, 1988). The authors of the SWB scale did not theorize that spiritual well-being would show any differences with age. Most of the studies conducted with the SWB scale do not show any correlation with age. Jang (1987) did find a difference with ethnic Chinese church attenders in the Pacific Northwest. In this study age significantly correlated with SWB fullscale and EWB subscale scores with subjects 26 years and older having higher scores than those 18-25 years old. Jang speculated that this may be due to the younger respondents being in school and financially dependent, but it also may have been affected by acculturation.

In studies reporting the relationship between age and SWB scale scores, Bressem, Waller, and Powers (1985), Mullins (1986), and Palmer (1985) found a positive correlation between age and SWB scale scores,

while Bufford (1984), Durham (1984), Hawkins (1986), and Hawkins & Larson (1984) found a negative relationship. Other studies have found no relationship (Bressem, 1986; Carpenter & Dean, 1985; Carr, 1986; E. Mueller, 1986; Temple, 1987). One observation that should be made is all the cited studies except Jang looked for a linear relationship. Perhaps there is a relationship between SWB scale scores and age but not in a linear way. Another explanation may be that age is not related to well-being but to life issues associated with certain age ranges. For example, mid-life changes may alter one's sense of well-being.

These findings are consistent with Diener's (1984) review of subjective well-being. Diener refers to a meta-analysis conducted prior to 1980 revealing the correlation between age and subjective well-being as near zero. Since spiritual well-being is hypothesized as an aspect of subjective well-being, the lack of significant correlations with age is theoretically expected.

Correlations between a new test and earlier tests measuring in the same general area are sometimes cited as evidence of construct validity. Correlations should be moderately high, but not too high. High

correlations would represent needless duplication (Anastasi, 1988). Correlations with similar tests and dissimilar ones are also used to demonstrate the new test is generally free from the influence of certain irrelevant factors (Anastasi, 1988). This section will examine studies using other measures of spiritual well-being. Correlations with similar and dissimilar constructs will be discussed later.

Moberg (1986) has identified several instruments designed to measure the phenomenon of spiritual well-being. Moberg describes the SWB scale as the first and most widely used of these instruments. Other scales measuring SWB, which Moberg lists, include the semantic differentiation scales by Calvin Farnham, the Spiritual Distress scale by Ruby Flesner, J. H. Kauffman's Religious Life Scale, and a few others in various stages of development.

Moberg (1984) attempted to explore subjective spiritual well-being with an 82 item questionnaire constructed from items previously used in spiritual well-being research plus new items taken from studies, interviews, and other sources. The 82 items were reduced to 45 continual variable items, factor analyzed, and correlated with other variables. Seven

indexes emerged from factor analysis clusters. They were: Christian Faith, Self-Satisfaction, Personal Piety, Subjective Spiritual Well-Being, Optimism, Religious Cynicism, and Elitism.

In later research using diverse American groups (N = 1,535) the EWB subscale correlated r = .73 with Moberg's Self-Satisfaction index and RWB correlated r = .86 with Christian Faith and r = .70 with Personal Piety.

Ruby Flesner (1981) constructed the Spiritual Distress Scale (SDS). The SDS is a 22 item self-report Likert scale based upon five major areas in which people can experience distress of the spirit (forgiveness, love, hope, trust, meaning, and purpose). The higher the score the more the individual is reporting spiritual distress. Scores on this scale correlate negatively with the SWB scale. In a sample of 88 first year nursing students, the two scales correlated r = -.45 (p <.001). One week later, 83 subjects from the original sample completed the same two scales resulting in correlations of r = -.90 (p <.001). There was no mention of subscale correlations. It should be noted the difference between correlations

is puzzling when the consistency of the scale scores over the time span is considered.

No other correlations between measures of spiritual well-being are known. The two studies which have been done suggest moderate to high correlations between scales in this domain but much more evidence needs to be accumulated.

Another area of construct validity is factor analysis. Factor analysis is a refined statistical technique for analyzing the interrelationships of data, looking for common traits that would account for obtained correlations. According to Anastasi (1988), the process of factor analysis involves reducing the number of variables or categories from several tests down to a small number of factors. After the factors have been identified they can be used to describe the factorial composition of the test. The tests can then be characterized in terms of the major factors determining their scores, together with the weight or loading of each factor and the correlation of the test with that factor.

Ellison (1983) reports a factor analysis on the 20 item SWB scale suggesting the scale clusters together as expected. Using a varimax rotation, three eigenvalues

emerged from the analysis at 7.136, 2.716, and 0.859 and two factors retained. Ellison reported all the RWB subscale items loaded on the first factor and several of the EWB subscale items loaded onto the second factor, connoting life direction. The sample for this analysis consisted of 206 students from three religiously oriented colleges. However, these findings have not been replicated and need further study (Bufford, 1984). The analysis is questionable since one of the eigenvalues identified is less than 1 and the particular factor analysis approach used assumed the items did not correlate with each other. An oblique rotation is probably the more appropriate choice with SWB scale items because these calculations assume the items intercorrelate with each other (Norusis, 1986). Another problem is the labeling of three factors when only two emerged from the analysis.

Cooper (1987) studied the construct validity of the SWB scale together with the Spiritual Maturity Index (SMI) and concluded one "general factor" was being measured rather than two separate constructs of spiritual well-being and spiritual maturity. Because the SWB scale and SMI were mislabeled in the study and the data, additional analysis of the separate scales

from his study cannot be confidently done. A replication of Cooper's study was done in 1987 and generally concluded the same thing (Davis, Longfellow, Moody, & Moynihan, 1987). Carr (1986) in her sample of 243 Christians found a common variance between the SMI and SWB scale scores of 43%.

These findings may cast doubt on the validity of the SWB scale from a factor analysis perspective. Although the Cooper and Davis et al. studies focused on the construct validity of the Spiritual Maturity Index it is not known what it is the SWB scale and SMI are commonly measuring. Gorsuch (1984) surmises there is a general religious dimension that is consistently found in studies of religious variables. This may be an explanation for the findings from the Cooper and Davis et al. studies.

An examination of the internal consistency of an instrument applies to both reliability and validity. Internal consistency supplies evidence of construct validity by demonstrating whether items on a scale have a high index of intercorrelation. If there is substantial intercorrelation it is assumed the items are measuring the same underlying variable, that a construct is being measured. However, even if internal

consistency coefficients indicate a construct is being measured it does not necessarily mean it is the intended construct (D. Mueller, 1986). There are limits on the contribution of internal consistency data to test validation because little can be learned about what the test measures without external data (Anastasi, 1988). Internal consistency cannot stand alone as a validation measure.

Total score and correlation of subtests to total score are two applications of the internal consistency method. Subtest to total score correlations from the SWB scale are presented later in Table 7.

Ellison (1983) reports coefficient alphas, a measure of internal consistency, as .89 for the SWB fullscale, and .87 (RWB), .78 (EWB) for the subscales. The sample consisted of 100 student volunteers from the University of Idaho. Fullscale to subscale correlations were .90 for RWB and .59 for EWB. No description of the sample was given nor correlations between EWB and RWB originally reported. Paloutzian (1982) reports the correlation between the two subscales as .32, but did not mention the sample this came from.

Convergent validity refers to how well a test correlates with other variables it should theoretically correlate with (Campbell & Fiske, 1967). Many studies have correlated the SWB scale with other variables. The SWB fullscale and RWB and EWB subscales positively correlate with many areas including religious variables, indicators of physical health, psychological health, marriage and family issues, demographics, and social desirability. Each of these areas will be discussed except for social desirability which is discussed later.

Among the single item religious variables the SWB scale correlates positively with are:

-frequency of church attendance (Bufford, 1984; Colwell, 1987; Durham, 1986; Ellison & Economos, 1981; Frantz, 1985; Hawkins, 1986; Huggins, 1988; Jang, 1987; Mitchell, 1984; Mitchell & Reed, 1983; Moody, 1988; Mullins, 1986; Quinn, 1984; Sherman, 1987)

-frequency of family devotions (Bufford, 1984)

-frequency and/or duration of personal devotions (Bressem, 1986; Bressem, Waller, & Powers, 1985; Bufford, 1984; Carr, 1986; Clarke, 1987; Colwell, 1987; Davis, Longfellow, Moody, & Moynihan, 1987; Ellison &

Economos, 1981; Huggins, 1988; Jang, 1987; Jang, Paddon, & Palmer, 1985)

-importance of religion (Bufford, 1984; Carr, 1986; Carson, Soeken, & Grimm, 1988; Davis et al., 1987; Durham, 1984; Durham, 1986; Frantz, 1985; Jang, 1987)

-religious knowledge (Bressem, Colwell, Mueller, Neder, & Powers, 1985; Carr, 1986; Davis et al., 1987; Jang, 1987)

-church leadership experience (Moody, 1988)

-feeling accepted and valued by God (Ellison & Economos, 1981; Ellison, Rashid, Patla, Calica, & Haberman, 1984)

-seeing God as a causal agent (Durham, 1984; Durham, 1986)

-estimation of one's spiritual maturity (Davis et al., 1987)

-attending seminary (Bufford, Bentley, Newenhouse, & Papania, 1986)

-participation in religious activities (Bonner, 1988)

-small group participation (Huggins, 1988).

In addition, people who describe themselves as "born again" Christians (acceptance of Jesus Christ as

personal Lord) report higher levels of SWB scale scores than those who consider themselves "ethical Christians" (follow ethical and moral teachings of Jesus) or non-Christians (Bufford, 1984; Campise, Ellison, & Kinsman, 1979; Davis et al., 1987; Durham, 1984; Durham, 1986; Ellison & Cole, 1982; Ellison & Economos, 1981; Ellison & Paloutzian, 1979; Jang, 1987; Jang, Paddon, & Palmer, 1985; Moody, 1988; Paloutzian & Ellison, 1979c; Papania, 1988; Quinn, 1984; Temple, 1987).

Those doctrinal beliefs, worship orientations, and devotional practices which promote a sense of personal acceptance and communion with God also correlate positively with SWB scale scores (Ellison & Economos, 1981; Ellison, Rashid, Patla, Calica, & Haberman, 1984). Financial giving (Jang, 1987) and a high application of Biblical principles (Jang, 1987) have positive SWB scale scores correlations. Jang also found a positive relationship between SWB scale scores and years as a Christian, a finding not supported by other studies (Bressem, 1986; Davis et al., 1987; Moody, 1988).

Scales measuring various dimensions of one's spiritual life positively correlate with the SWB scale,

including low scores on the Religious Orientation Scale-Intrinsic (ROS-I) (Bufford, 1984; Ellison & Paloutzian, 1979; E. Mueller, 1986; Quinn, 1984), and high scores on the Spiritual Maturity Index (SMI) Bressem, 1986; Bufford, 1984; Carr, 1986; Jang, Paddon & Palmer, 1985; E. Mueller, 1986; Parker, 1984), the Spiritual Leadership Qualities Inventory (SLQI) (Carr, 1986; Parker, 1984), the Supernatural Locus of Control (SLOC) (Durham, 1986), and the REL (Religious Fundamentalism Content) scale from the MMPI (Frantz, 1985).

The Spiritual Maturity Index (SMI) is a 30 item Likert type scale developed by Craig Ellison to measure spiritual maturity. It is designed to measure spiritual health through behavioral and attitudinal criteria and based largely on principles taken from the Bible. Table 2 presents some correlations between the SWB scale and the SMI.

Table 2

Correlations Between the Spiritual Well-Being Scale and the Spiritual Maturity Index

Study	N	RWB	EWB	SWB
Bressem (1986)	80	.67***	.52***	.64***
Bufford (1984)	65	.82*	.39*	.62*
Carr (1986)	243	.62***	.56***	.66***
Colwell (1987)	51	.68***	.57***	.72***
Davis et al. (1987)	321	.73***	.58***	.72***
Ellison et al. (1984)	239	.62***	NA	.57***
Jang, Paddon, & Palmer (1985)	43	.80***	.48***	.75***

*p < .05 **p < .01 ***p < .001

Note. Colwell used the 20 item version of the SMI.

The Religious Orientation Scale (ROS) is a 21 item measure of a person's intrinsic and extrinsic religious orientations. The intrinsic dimension identifies people who tend to focus themselves around their religion and view other activities as instrumental in accomplishing religious goals. Table 3 presents some

correlations between the SWB scale scores and the intrinsic scale of the ROS.

Table 3

Correlations Between the Spiritual Well-Being Scale and the Religious Orientation Intrinsic Scale

Study	N	RWB	EWB	SWB
Agnor (1986)	26	-.27	-.10	-.60***
Bufford (1984)	65	-.76*	-.27*	-.58*
E. Mueller (1986)	51	-.29*	-.35**	-.37**
Paloutzian & Ellison (1979a)	137	-.80***	-.29**	-.72***
Quinn (1984)	156	-.81*	-.37*	-.71*
Temple (1987)	106	-.83***	-.35***	-.71***

*$p < .05$ **$p < .01$ ***$p < .001$

Note. Higher SWB scale scores suggest greater well-being, lower ROS-I scores suggest a more intrinsic religious orientation.

Several studies have shown modest support for a positive relationship between the SWB scale and physical health. In dialysis patients, SWB scale scores correlated positively with global adjustment and acceptance of disability (Campbell, 1983). SWB scale scores positively correlated with ideal body weight and self-ratings of past and present health (Hawkins & Larson, 1984), reduction in use of medications after treatment (Mullins, 1986), and attitude towards seeking medical help (Bufford, 1987). SWB scale scores correlated negatively with blood pressure (Hawkins, 1986). In addition EWB subscale scores positively correlated with current health (Bufford, 1987). A positive relationship was observed between SWB scale scores and using religious means of coping with pain (Bonner, 1988; Campbell, 1983; Mullins, 1986).

In the area of psychological health, SWB scale scores positively correlated with measures of assertiveness as measured by the Interpersonal Behavior Survey (Bufford & Parker, 1985; Campbell, 1983; Hawkins, 1986; Mullins, 1986; Sherman, 1987), self-esteem (Campise, Ellison, & Kinsman, 1979; Ellison & Economos, 1981; Ellison et al., 1984; Marto, 1984; Paloutzian & Ellison, 1979a), internal locus of control

(Jang, Paddon, & Palmer, 1985; Marto, 1984; Palmer, 1985), hopefulness (Palmer, 1985), self-concept as measured by the Tennessee Self Concept Scale (Colwell, 1987; Mitchell & Reed, 1983), social skill (Ellison, 1983), acculturation of Chinese people in the U.S. (Jang, 1987), and estimation of one's life satisfaction. Temple (1987) found positive correlations between all three SWB scales and the Psychological General Well-Being Scale (PGWB), especially with the existential well-being subscale.

SWB scale scores also correlate positively with lower mood disturbance in pregnant women (Mitchell, 1984) and with purpose in life (Paloutzian & Ellison, 1979a; Paloutzian & Ellison, 1979c). See Ellison (1983) for tables of SWB scale scores correlated with the Purpose in Life scale.

Palmer (1985) reports positive correlations with the Hope Index Scale (HIS). Carson, Soeken, and Grimm (1988) also found positive correlations between the SWB scale and the State-Trait Hope Scale. The Spiritual Well-Being Scale positively correlates with the Supernatural Attribution Questionnaire and the God as Causal Agent Scale (Durham, 1984).

Rotter's Internal vs. External Locus of Control Scale is a 29 item forced choice questionnaire designed to assess an individual's expectations about how reinforcement is controlled. Lower scores indicate more internality. Table 4 presents some correlations between the SWB scale and the Rotter scale.

Table 4

Correlations Between the Spiritual Well-Being Scale and the Rotter Locus of Control Scale

Study	N	RWB	EWB	SWB
Agnor (1986)	52	-.41***	-.46***	-.60***
Durham (1986)	177	-.23**	-.32**	-.30***
Jang, Paddon & Palmer (1985)	43	-.33*	-.17	-.29*
Palmer (1985)	42	-.33**	-.48*	-.46**

*$p < .05$ **$p < .01$ ***$p < .001$

The Tennessee Self-Concept Scale (TSCS) is a 100 item self report instrument that measures one's self concept across many sub-areas including physical self,

moral-ethical self, personal self, family self, and social self, besides a general self-esteem and a total positive self-concept score. Table 5 lists some correlations between the SWB scale and the two overall scores.

Table 5

Correlations Between the Spiritual Well-Being Scale and the Tennessee Self-Concept Scale

Study	N	RWB	EWB	SWB
Agnor (1986)	52			
Total Positive		.18	.39**	.44***
Identity		.20	.41***	.47***
Colwell (1987)	51			
Total Positive		.24*	.51***	.43***
Identity		.26*	.51***	.44***

*$p < .05$ **$p < .01$ ***$p < .001$

In the area of marriage and family, SWB scale scores have correlated positively with a father's self-esteem, but not his children's (Marto, 1984), the

decision to continue a pregnancy as opposed to abortion (Mitchell, 1984), and marital satisfaction or adjustment as measured by the Marital Satisfaction Index (Quinn, 1984), the Marital Satisfaction Scale (Mashburn, 1987), and the Dyadic Adjustment Scale (Roth, 1988). Estimation of one's present family closeness has correlated with all three scales, but not family closeness while growing up (Jang, 1987). However, other studies have found positive SWB scale correlations with perceived quality of parent-child relationships, memories of family togetherness as a child, and childhood peer relations (Campise, Ellison, & Kinsman, 1979; Ellison & Paloutzian, 1978; Ellison & Paloutzian, 1979). SWB scale scores were higher among couples who reported being more androgynous than those endorsing more typical masculine and feminine sex-role orientations (Mashburn, 1987).

In one study, there were no significant relationships between one's marital status and SWB scale scores (Campise, Ellison, & Kinsman, 1979), a finding not confirmed by other studies. The EWB subscale negatively correlated with number of marriages in one study (Hawkins, 1986).

Some of the demographic variables that correlate positively with the SWB scale include full-time employment (Jang, 1987), city living (Paloutzian & Ellison, 1979d), financial independence (Jang, 1987), and financial condition (Moody, 1988). Mixed results have been obtained between education and SWB scale scores (Carr, 1986; Temple, 1987).

Discriminant, or divergent validity includes correlations with those tests that should show little or no relationship to the test (Campbell & Fiske, 1967). It also includes those constructs with which it should have negative correlations (D. Mueller, 1986). These type of correlations again do not prove that a construct is measuring what it is supposed to but provides evidence in support of it.

In the area of psychopathology and "poor" mental health, the SWB scale has negatively correlated with loneliness (Ellison & Paloutzian, 1979; Paloutzian & Ellison, 1979a; Paloutzian & Ellison, 1979c; Paloutzian & Ellison, 1979d), depression as measured by the Beck Depression Inventory (Campbell, 1983), aggression as measured by the Interpersonal Behavior Survey (IBS) (Bufford & Parker, 1985; Hawkins, 1986; Mullins, 1986; Sherman, 1987), shyness and dependency as measured by

the IBS (Bufford & Parker, 1985), and MMPI clinical scales (Frantz, 1985; E. Mueller, 1986; Mullins, 1986). Hawkins (1986) found negative correlations between all SWB scales and cigarette and alcohol use. Bonner (1988) reports high scores on the SWB scale correlated significantly with lower levels of withdrawal from social contacts and responsibilities and with lower levels of despair.

Papania (1988) and Rodriguez (1988) report lower Spiritual Well-Being Scale scores in those who have a history of sexual trauma. Those in prison also report lower SWB scale scores (Agnor, 1986) as do child molesters (Papania, 1988).

The SWB scale has negatively correlated with primary value orientations such as individualism, success, and personal freedom (Campise, Ellison, & Kinsman, 1979), and with a sense of rejection (Paloutzian & Ellison, 1979a).

The Religious Orientation Scale (ROS) has an extrinsic subscale that measures the degree to which an individual uses his religion to further other goals. Table 6 presents some correlations between the extrinsic scale and the SWB scale.

Table 6

Correlations Between the Spiritual Well-Being Scale and the Religious Orientation Extrinsic Scale

Study	N	RWB	EWB	SWB
Agnor (1986)	26	-.22	-.01	-.09
Bufford (1984)	65	-.36*	.01*	-.17*
E. Mueller (1986)	51	-.09	-.01	.06
Quinn (1984)	156	-.53*	-.35*	-.52*
Temple (1987)	106	-.36***	-.36***	-.42***

*$p < .05$ **$p < .01$ ***$p < .001$

Note. Higher SWB scale scores suggest greater well-being, higher ROS-E scores indicate a more extrinsic religious orientation.

The SWB scale has been studied with other scales and produced no significant correlations. Among these are Richardson's Visualizer-Verbalizer Questionnaire (Bressem, Waller, & Powers, 1985), the Intense Ambivalence Scale (Lewis, 1986), Hood's Mysticism scale and Betts QMI Vividness of Imagery scale (Bressem,

1986). No relationship was observed between SWB scale scores and perfectionism (Ellison et al., 1984).

According to Anastasi (1988), another source of data for construct validity are experiments on the effect of selected variables on test scores. D. Mueller (1986) calls this validation method a variant of the known group method, discussed below. Mean scores of a group known to be high in the construct are compared with mean scores of a group known to be low (same group, different instance in this method). This method is workable only if it can be assumed that the experimental treatment is effective.

Upshaw (1984) used a sample of 24 volunteer Christian newlywed couples to examine the effects of communication training on marital satisfaction, commitment, spiritual well-being, and social desirability. In this pretest-posttest control group design, the couples were randomly divided into three treatment groups: communication training, film strips, and control with the independent variable being treatment group.

Results applying to the SWB scale revealed the different treatment methods did not significantly effect SWB fullscale or RWB subscale scores but did

alter the EWB subscale. The EWB subscale temporarily decreased for the communication training group. Ten weeks later no difference was found. The author proposed that the stress of learning new ways of communicating in marriage produced a short-term decrease in reported sense of well being.

Presently, other studies involving experimental variable manipulation with the SWB scale are not known.

The final area under construct validity is known group differences. Being able to distinguish between population groups it theoretically should is another evidence of validity.

Studies using the SWB scale indicate its ability to make these distinctions. Bufford, Bentley, Newenhouse, & Papania (1986) examined group means from previous studies with the scale. They found that Unitarians scored significantly lower than all other groups except for non-Christian sociopath convicts on SWB fullscale scores and the RWB subscale. These same sociopaths were lower than all other samples on the EWB subscale. Seminarians scored higher than medical outpatients, and other church attenders on all the scales. Significant differences were also found between those suffering from eating disorders and

medical outpatients (Sherman, 1987) and between those choosing to keep their unborn babies as opposed to aborting (Mitchell, 1984).

In summary, there is a lot of evidence supporting the SWB scale in terms of construct validity. The SWB scale seems to be able to correlate negatively and positively with other measures it theoretically should. The major weakness in this area is factor analysis, in understanding what the Spiritual Well-Being Scale is measuring.

Criterion Validity. According to Aiken (1979), all tests are validated by relating test scores to performance on criterion measures. These measures are standards or variables against which test performance can be evaluated. In other words, criterion validity addresses how effective a test is in predicting an individual's behavior in specified situations (Anastasi, 1988). Whenever the criterion measure, whatever it may be, is available at the time of testing, then the concurrent validity of the test is being studied. When the criterion does not become available until sometime after the test is administered, the predictive validity of the test is of interest (Aiken, 1979).

According to the Standards, when using criterion validity the sample should be described along with procedures including time elapsed between test administration and collection of the criterion data. In addition, the statistical analysis used to determine the degree of predictive accuracy and differential prediction should be explained. All criterion measures should be described accurately along with the rationale for choosing them.

Concurrent validation procedures are employed whenever a test is administered to people falling into various categories, such as diagnostic groups or socioeconomic levels. If the average score varies substantially from category to category, then the test might be used as another, perhaps more efficient, means of classifying people into these categories (Aiken, 1979).

Quinn (1984) was unable to predict marital satisfaction, as measured by the Marital Satisfaction Inventory (MSI), from SWB scale scores. This was true even when partialling out variance due to conventionalization as measured by a MSI subscale. The SWB scales ranked 8th out of ten variables, accounting

for only 1% of the variance with his sample of church attenders.

Bressem, Colwell, Mueller, Neder, and Powers (1985) asked church leaders to select four members of their congregation, two judged to be high in spiritual maturity and two thought to be low. These individuals (N = 64) were asked to complete the SWB and SMI scales. Although no difference was discovered between the SMI scores and the two groups, the EWB subscale positively correlated with leader's perception of the individual's relationship to God, spiritual maturity, religious knowledge, and Christian walk. The RWB subscale correlated positively with leader's ratings of spiritual maturity and Christian walk, and SWB fullscale scores correlated positively with present relationship to God, spiritual maturity, religious knowledge, and Christian walk.

Clarke (1987), in a sample of 298 Youth For Christ workers, used 19 predictor variables including job-related areas, Christian life, family background, and demographics to try to predict SWB scale scores. The regression equations were weak for all three scales and Clarke concluded that identifying an adequate predictive model was not achieved.

Huggins (1988) examined the effects of small group participation on SWB scale scores among Conservative Baptists in Oregon. Using an analysis of variance regression procedure he found significant main effects for small group attendance, personal devotions, and church attendance on SWB scale scores.

Predictive validity deals with how accurately test scores predict criterion scores. This relationship is expressed by the correlation between the test and some measure of performance. Predictive validity is primarily of concern with respect to aptitude or intelligence tests (Aiken, 1979). The predictive validity of an attitude measure is its correlation with a criterion behavior.

The social psychological literature is full of studies in which attitude measures tried to predict particular behaviors only to discover zero or low correlations (D. Mueller, 1986). Mueller gives three reasons for the lack of predictive success of attitude measures: (a) low reliability of the attitude measures used, (b) people don't always act in accord with their attitudes, and (c) there is sometimes dissimilarity in the attitudinal and behavioral objects studied, in other words, the attitudinal object was assumed to be

the same as the behavioral object studied when it was not. There are no known studies which have examined this area using the SWB scale.

In summary, there is much evidence which supports the validity of the SWB scale. One of the weakest areas is factor analysis, in determining just what the SWB scale is measuring. In addition, there are few true experiments or studies exploring predictive validity with the SWB scale. In support of the scale are the many studies which show a definite relationship in the predicted direction between other measures of religion, mental health, pathology and well-being.

Reliability

Reliability refers to the accuracy (consistency or stability) of measurement by a test. There are several different methods for investigating the reliability of a test including test-retest, alternate forms, split-half, internal consistency, and inter-rater or scorer reliability.

According to the Standards, evidence of reliability that permits the reader to judge whether scores are sufficiently dependable for the recommended uses of the test should be reported. If any of the

necessary evidence has not been collected, the absence of such information should be noted. These standards apply to every score, subscore, or combination of scores on the test. If there is a low reliability between scores, caution must be taken in interpretation. The minimum difference between two scores ordinarily required for statistical significance needs to be stated.

Estimates of reliability and standard error of measurement, along with procedures, samples, and conditions should be described sufficiently to permit a user to judge to what extent the evidence is applicable to the person and problems with which he or she is concerned. This includes demographic information such as age, gender, SES level, intelligence, employment, and minority group membership and the procedures used to obtain the samples and the numbers of individuals in each sample group.

Other information which should be presented includes adjusted and unadjusted reliability coefficients and standard deviations for restriction of range, standard errors of measurement at critical score levels, and reliabilities and standard errors of measurement for different populations if these differ.

To date, only a portion of this information is available for the Spiritual Well-Being Scale. Two types of reliability studies have been done so far: test-retest and internal consistency. Because of their nature, alternate (parallel) form and scorer reliabilities do not apply to the Spiritual Well-Being Scale.

If the test developers' suggest their test is a measure of a generalized, homogeneous trait, evidence of internal consistency should be stated. Since the SWB scale is seeking to measure a homogeneous trait coefficient alphas for the full scale and the two subscales are reported. One study gives data on internal reliability (Paloutzian & Ellison, 1979a). In a sample of 122 student volunteers at the University of Idaho, coefficient alphas, a measure of internal consistency, were .88 for the SWB fullscale, .87 for the RWB subscale, and .75 for the EWB subscale. Based on this study the authors concluded the SWB scale has a high internal consistency.

Besides estimates of internal consistency, when a test consists of separately scored parts or sections, the correlations between the parts or sections should be reported along with relevant means

and standard deviations. Using a sample of 206 students from three religious colleges, moderate to high correlations were found between the fullscale SWB and RWB (r = .90) and EWB (r = .59) subscales (Ellison, 1983). It should be noted that correlations between the SWB subscale and fullscale are in part artificial because of their relationship to each other. The authors did not originally report means or standard deviations for this scale.

Table 7 presents some intratest correlations between the SWB fullscale and subscales.

Table 7

Intratest Correlations for the Spiritual Well-Being Scale

Study	N	RWB-SWB	EWB-SWB	RWB-EWB
Agnor (1986)	52	.78***	.80***	.62***
Bressem (1986)	80	.89***	.93***	.69***
Bufford (1984)	65	.68*	.41*	.20*
Campise, Ellison, Kinsman (1979)	87	.90***	.70***	.32***
Carr (1986)	243	.90***	.93***	.69***
Frantz (1985)	72	.91***	.92***	.71***
Hawkins (1986)	88	.92***	.85***	.57***
Mitchell (1984)	81	.91***	.88***	.62***
E. Mueller (1986)	51	.87***	.87***	.52***
Mullins (1986)	41	.90***	.91***	.65***
Palmer (1985)	44	.89**	.78***	.43*
Quinn (1984)	156	.91*	.83*	.54*
Temple (1987)	106	.92***	.84***	.56***
Upshaw (1984)	48	.87***	.74***	.33*

*$p < .05$ **$p < .01$ ***$p < .001$

Test-retest correlations have been reported once for the SWB scale in the initial construction and testing. On the same sample of 122 student volunteers at the University of Idaho test-retest coefficients were .93 (SWB), .96 (RWB), and .86 (EWB). The time span between testings was one week. There was no mention of means, standard deviations, and standard errors of measure (Paloutzian & Ellison, 1979a).

Flesner (1981), in developing a scale to measure spiritual distress, examined test-retest reliability for her scale and the SWB scale, which was included as a measure of validity. Using 88 volunteer nursing students (83 female, 5 male) the two scales were administered together and again one week later. Test-retest reliabilities for individual scores were not compared, only the group means from the two testings. The first administration of the SWB scale yielded a mean of 99.36 with a standard deviation of 20.10 from 88 respondents. The second administration yielded a mean of 97.64 and a standard deviation of 16.87 from 83 respondents. The differences in the mean scores of the two testings were less than 1.7 percent.

Another test-retest reliability study was unintentionally done by Upshaw (1984). In his

experiment on the effects of communication training on marital satisfaction and spiritual well-being the SWB scale was administered three times to 48 subjects. Three groups of 16 subjects were formed with two groups receiving training and one group serving as a control. Measurements were taken prior to the intervention, then four weeks later immediately following the treatments, and finally ten weeks after that.

Test-retest reliability correlations for this study are presented below. These correlations were only partially reported in the dissertation. A reanalysis of the raw data yielded the following scores, presented in Table 8.

Table 8

Test-Retest Correlations for the Spiritual Well-Being Scale From Upshaw (1986)

	Four Weeks			Ten Weeks			Fourteen Weeks		
Group	RWB	EWB	SWB -	RWB	EWB	SWB -	RWB	EWB	SWB
CCS	.66*	.30	.53	.68*	.36	.59*	.97#	.81#	.84#
FFS	.98#	.95#	.96#	.98#	.95#	.96#	1.00#	1.00#	1.00#
Control	.99#	.98#	.99#	.99#	.98#	.99#	1.00#	1.00#	1.00#
Total	.89#	.62#	.76#	.89#	.65#	.78#	.99#	.96#	.98#

# $p < .001$ * $p < .01$

Note. CCS and FFS were the treatment groups, $N = 48$ (16 in each group).

While these results are promising, more needs to be done to satisfy the standards in terms of a test's reliability. Missing are additional test-retest coefficients, internal alphas, and reporting of standard errors of measurement.

Test Development and Revision

The Standards assert that tests should be developed on a sound scientific basis prior to publication. This includes the specifications used in constructing items and in designing the test instrument as a whole. The definition of the universe used for constructing or selecting items should be described and be clear enough that knowledgeable experts can judge the relations of items to the domains they represent. Item content should also be sensitive to the cultural and experiential diversity of the intended population.

As mentioned earlier, spiritual well-being is operationally defined as the affirmation of life in a relationship with God, self, community, and environment that nurtures and celebrates wholeness (NICA, 1975). SWB is theorized as having a vertical and horizontal dimension reflecting both a religious and social-psychological component. The authors made more distinctions such as differentiating between well-being and spiritual health and maturity.

It is assumed that from these distinctions and definitions, the items which constitute the SWB scale and its religious and existential subscales were selected. However, the process was not reported and it

is not possible to evaluate the items except from the already stated criteria. Since the SWB scale is an attitude scale and not a measure of aptitude, this is not an area of great concern (Anastasi, 1988).

The Standards also recommend that test taking strategies which could influence test performance significantly should be explained to the test takers themselves. This includes sensitivity to practice or coaching. The extent to which scores are susceptible to an attempt by test takers to present false or unduly favorable pictures of themselves should be examined.

Several studies have examined the relationship between scale scores of the SWB and various measures of social desirability. Social desirability is defined as trying to present oneself in a favorable light.

Carr (1986) studied 239 volunteers from several churches and a seminary. Edward's Social Desirability scale positively correlated with SWB fullscale, and EWB and RWB subscales. Clark, Clifton, Cooper, Mishler, Olson, Sampson, & Sherman (1985) examined 33 Christians selected as mature or less mature by church leaders. Using the same two scales they found positive correlations between the Edward's and SWB scales. Mitchell and Reed (1983) found a positive correlation

between the Edward's scale and the SWB fullscale with 49 single adult Christians. Wong (1989) found near zero, nonsignificant correlations between the Edward's scale and all three SWB scales with 72 Chinese Americans.

The Marlowe-Crowne Social Desirability Scale was used by Upshaw (1984) to measure its relationship with the SWB scale on 24 newly married couples participating in a communication skills training program. The correlations between the two were negligible and not significant.

Correlations with the validity scales of the MMPI have been used in two studies with the SWB scale. Frantz (1985) found significant negative correlations with the MMPI F scale and SWB fullscale and subscales using 72 outpatient counseling clients. Mullins (1986) reported modest but significant correlations with the MMPI K scale and SWB scales in a study of 41 chronic pain inpatients.

Besides these studies, others are in progress which are examining the effects of asking subjects to fake good or bad on the SWB. Moody (1988) asked a sample of church attenders to complete the SWB scale using three sets of instructions. Subjects were

randomly assigned to three groups and asked to complete the SWB scale honestly, faking good, or faking bad. Results found a significant difference between those who faked bad and those who completed it honestly. Significant differences were not found between the fake good group and the honest group scores. Moody postulated that this lack of difference may be due to the low ceiling on the scale.

In summary, the studies to date are inconclusive on the effects of social desirability and spiritual well-being. More research needs to be done in this area on examining how people are presenting themselves when they complete the SWB scale. For a more thorough discussion, see Moody (1988).

Another area involves test administration. Directions for test administration should be detailed enough so test takers can respond to the task in the manner the test developer intends. The directions for administration should be presented with sufficient clarity and emphasis so the test user can duplicate, and will be encouraged to duplicate, the administrative conditions under which the norms and the data on reliability and validity were obtained.

The directions for taking the test are given to the test taker on the test itself. At the top of the sheet these instructions are given: "For each of the following statements circle the choice that best indicates the extent of your agreement or disagreement as it describes your personal experience."

No other guidelines are given or suggested by the authors for the test users or takers. No studies have commented on the adequacy of these instructions and respondents seem to understand how to complete the scale.

The Standards recommend test revisions be made when new research data, significant changes in the domain represented, or new conditions of test use and interpretation make the current version inappropriate. So far, the original version has not been amended since the initial work was done on the scale.

Procedures for scoring tests locally should be presented by the test developer in sufficient detail and clarity to maximize the accuracy of scoring.

The SWB scale is an objective instrument and scoring is simple and straightforward as with most Likert scaled instruments. The numerical value of each item ranges from 1 to 6 with higher scores indicating

more well-being. For items numbered 3, 4, 7, 8, 10, 11, 14, 15, 17, 19, 20 the scoring is reversed (for example Strongly Disagree, which is normally assigned a 6 is scored as a 1). The sum of all 20 items make up the SWB fullscale score, the 10 odd numbered items constitute the RWB subscale, the 10 even numbered items the EWB subscale. A computer scoring program for the SWB scale is in use at Western Conservative Baptist Seminary (WCBS).

In research at WCBS missing data is typically assigned a value of "3.5" and summed with the other scores. When two responses are circled the average of the two is used. If more than 5 items are left blank the test is not scored and considered invalid. There are other ways to handle missing data. Whichever way the scoring of missing data is done should be mentioned in the reporting of the data.

Scaling and Norming

The scales used for reporting scores should be carefully described to increase the likelihood of accurate interpretation and understanding of both the test user and the test taker. The norms should be reported in terms of standard scores or percentile ranks which reflect the distribution of scores in an

appropriate reference group or groups. This is done because raw test scores are expressed in units that result from arbitrary features of the test.

The norming groups should be clearly defined groups to whom users of the test will ordinarily wish to compare the person tested. A well planned sample should be taken and reported in the manual with sufficient information about sampling method, numbers, and procedures, and the year done. Because the norming process can be a difficult and costly one, user norms or program norms that consist of descriptive statistics based on all test takers in a given period of time rather than norms obtained by formal sampling methods may be used but should be reported as such.

To date, only preliminary attempts have been made to establish norms or convert scale scores for the SWB scale. Table 9 presents some descriptive data for different samples on the SWB fullscale and RWB, EWB subscales. Included are the number of respondents who received the highest score and what percentage of the population that is. Note the means are usually within one to two standard deviations from the ceiling. The maximum score for the SWB fullscale is 120, and for both the RWB and EWB subscales, 60.

Table 9

Spiritual Well-Being Scale Descriptive Statistics

Religious Well-Being Subscale							
Study	N	Description	Mean	SD	Range	Top	Pct
A	25	sociopaths	35.6	9.2	18-60	1	4%
B	80	Bible college	55.5	4.7	40-60	12	15%
C	243	Christian	55.1	6.2	27-60	67	28%
D	298	youth workers	55.4	5.3	37-60	74	25%
E	177	Christian	53.9	7.2	31-60	59	33%
F	72	outpatients	47.3	8.9	24-60	6	8%
G	88	medical outpt	51.0	10.9	10-60	23	26%
H	169	Chinese church	53.2	7.4	29-60	37	22%
I	46	Baptists	53.5	7.4	36-60	13	28%
J	51	Unitarians	34.1	13.0	10-60	1	2%
K	51	seminarians	54.8	5.9	37-60	9	18%
L	41	chronic pain	43.9	10.9	16-60	5	12%
M	112	school mothers	53.4	8.2	25-60	31	28%
N	44	quit smoking	42.9	11.3	17-60	5	11%
O	55	molesters	43.7	12.4	18-60	6	11%
P	62	eat disorder	40.6	11.1	19-60	1	2%

(table continues)

Table 9 continued

Existential Well-Being Subscale							
Study	N	Description	Mean	SD	Range	Top	Pct
A	25	sociopaths	40.7	9.2	24-60	1	4%
B	80	Bible college	50.9	6.3	33-60	4	5%
C	243	Christian	51.1	7.3	13-60	23	10%
D	298	youth workers	55.4	5.3	26-60	22	7%
E	177	Christian	51.4	6.9	29-60	13	7%
F	72	outpatients	39.6	10.4	19-60	1	1%
G	88	medical outpt	50.3	8.4	28-60	9	10%
H	169	Chinese church	49.5	7.8	24-60	10	6%
I	46	Baptists	50.8	8.1	32-60	6	13%
J	51	Unitarians	48.7	7.6	33-60	3	6%
K	51	seminarians	51.3	5.9	17-60	3	6%
L	41	chronic pain	41.7	11.1	17-60	3	7%
M	112	school mothers	51.2	6.7	30-60	7	6%
N	44	quit smoking	45.7	8.0	26-60	1	2%
O	55	molesters	42.8	11.1	18-60	2	4%
P	62	eat disorders	38.4	8.4	13-58	0	0%

(table continues)

Table 9 continued

Spiritual Well-Being Fullscale							
Study	N	Description	Mean	SD	Range	Top	Pct
A	25	sociopaths	76.3	16.3	50-120	1	4%
B	80	Bible college	106.59	10.15	77-120	3	4%
C	243	Christian	106.12	12.43	44-120	20	8%
D	298	youth workers	106.20	10.94	67-120	17	6%
E	177	Christian	105.38	13.07	61-120	12	7%
F	72	outpatients	86.65	17.65	45-119	0	0%
G	88	medical outpt	101.37	17.11	61-120	9	10%
H	169	Chinese church	102.78	14.38	55-120	8	5%
I	46	Baptists	104.02	14.23	72-120	4	9%
J	51	Unitarians	82.81	15.02	59-118	0	0%
K	51	seminarians	106.00	10.29	74-120	2	4%
L	41	chronic pain	85.34	19.75	33-120	2	5%
M	112	school mothers	104.60	13.15	70-120	5	5%
N	44	quit smoking	88.52	16.46	57-119	0	0%
O	55	molesters	86.56	19.14	47-120	2	4%
P	62	eat disorders	78.98	16.24	39-115	0	0%

(table continues)

Table 9 continued

Note. Top = Number of respondents achieving maximum score;

Pct = Percentage of sample achieving maximum score;

The symbols for each study are:

A = Agnor (1986); B = Bressem (1986);

C = Carr (1986); D = Clarke (1987);

E = Durham (1986); F = Frantz (1985);

G = Hawkins (1986); H = Jang (1987);

I = Lewis (1986); J = Lewis (1986);

K = E. Mueller (1986); L = Mullins (1986);

M = Newenhouse (1988); N = Palmer (1985);

O = Papania (1988); P = Sherman (1987).

According to Anastasi (1988), test construction should ideally follow these steps: theoretical description of a measure, item development and selection, psychometric investigation, and normative studies. At this point, psychometric investigation on the SWB scale is still underway and needs to be sufficiently complete before the last step of developing norms is taken.

Test Publication

Test manuals are important because they communicate important information about the test so qualified users or reviewers can evaluate the appropriateness and technical adequacy of the test. The Standards suggest that when a test is published for operational use, it should be accompanied by a manual that makes every reasonable effort to follow the recommendations and meet the specific standards set forth. This manual should be updated at appropriate intervals.

Presently, no manual exists for the Spiritual Well-Being Scale. A manual should be developed and circulated when there is general consensus the SWB scale has satisfied the professional standards.

Summary

This section has looked at the technical standards recommended for all psychological tests and how the SWB scale has measured up. As presented, the evidence for reliability and validity look very promising. Most of the studies have focused on construct validity and sampled evangelical Christian populations. There are still many areas which must be examined, including reliability studies, ways to reduce high scores, factor

analysis, norming, manual publication, and others. Different populations also need to be sampled including different religious beliefs, ethnic backgrounds, and socio-economic status. Research and development should continue with this scale before it is released for general use.

The next two sections focus in on two areas of the SWB scale. These sections will look at reliability and response measurement, how they relate to the SWB scale, and what must be done with the scale considering this information.

Reliability

In any type of measurement some error is inevitable (Kaplan & Saccuzzo, 1982). This error can either be from systematic bias or from random error. An example of a systematic error would be a thermometer that always registered two degrees higher than the actual temperature. Random error would be at work if that same thermometer was accurate but the researcher misread it while making different measurements. Random errors of measurement are never eliminated, but in order to depict nature in its ultimate lawfulness, efforts can be made to reduce them as much

as possible. When these random errors are slight, a measure is said to be reliable.

Reliability concerns the extent to which measurements are repeatable. Measurements are intended to be stable over a variety of conditions in which virtually the same results are obtained. The theory of measurement error has developed mainly from the discipline of psychology, and largely by psychologists (Nunnally, 1978).

Reliability tries to answer the question: Does the instrument yield consistent results? Henerson, Morris, & Fitz-Gibbon (1978) use the analogy of a friend when describing reliability. A reliable friend is one on whom you can count on to behave the same way time and again. A test, or a questionnaire, which yields essentially the same results when readministered is an instrument that is reliable in this sense. These authors do caution that consistency does not guarantee truthfulness. Therefore, an instrument that is reliable does not necessarily mean it is a good measure of what it was created to measure.

According to Kaplan and Saccuzzo (1982), the concept of reliability conveys the extent to which individual differences in test scores are due

to "true" differences in the characteristics under consideration and the extent to which they are due to chance, or measurement errors. Since measuring instruments are imperfect, the score observed for an individual may differ from the person's true ability or characteristic.

In other words, measures of test reliability make it possible to estimate what proportion of the total variance of test scores is error variance. Any condition that is irrelevant to the purpose of the test represents error variance. So, when examiners try to maintain uniform testing conditions by controlling the testing environment, instructions, time limits, rapport, and other similar variables, they are seeking to reduce error variance and make the test scores more reliable.

Despite the best testing conditions, no test is a perfectly reliable instrument (Anastasi, 1988). Knowing this, every test should be accompanied by a statement of its reliability to help test users make better use of the test.

There are many ways of examining test reliability. Since all types of reliability address the degree of consistency or agreement between two independently

obtained sets of scores, they are expressed in terms of a correlation coefficient. Correlations between measures of abilities are nearly always positive, although often low (Anastasi, 1988).

The most common way to compute a correlation coefficient is with the Pearson Product-Moment correlation coefficient which takes into account the individual's position in the group, and the amount of their deviation above or below the group mean. The Pearson correlation coefficient will have a high positive value when corresponding scores are of equal sign and approximately equal amount in the two variables.

Sources of Variance

There are many variables that prevent measurements from being exactly repeatable, the number and nature depending on the type and use of the test. According to Anastasi (1988), any reliability coefficient may be interpreted directly in terms of the percentage of score variance due to different sources. Thus, a reliability coefficient of .85 signifies that 85% of the variance in test scores depends on true variance in the trait measured, and 15% depends on error variance.

All the errors which occur within a test can easily be encompassed by the domain sampling model (Nunnally, 1978). The domain sampling model considers the problems of using a limited number of items to represent some larger domain or construct. Since all the items that make up the construct of spiritual well-being cannot possibly be used in a test, estimates must be made from a sample of all the items. The task of reliability analysis is to estimate how much error there is by using the score from the shorter test as an estimate of someone's true spiritual well-being (Kaplan & Saccuzzo, 1982). This model conceptualizes reliability as the correlation between the observed score from the SWB scale and the longer true score.

For any test or scale the sampling of items from a domain can be thought of in terms of not only the physical collection of items, but also the sampling of the many situational factors that will influence responses to those items. Thus, not only would each person receive a random sample of items from the domain, but also each item would be accompanied by a random set of situational factors. All such sources of error will tend to lower the average correlation among items within the test, but the average correlation is

all that is needed to estimate reliability (Nunnally, 1978).

The study of measurement error relating to variation between domain items generally uses alternative forms, which are intended to approximate randomly parallel tests. One source of error with this method comes from systematic differences in the content of the two tests. Since the SWB scale does not use an alternate form, another way of estimating reliability must be used. Two ways of doing this are the test-retest and internal consistency methods described later. Another method that can be used is the standard error of measurement, also discussed later.

People can change in regard to the attribute being measured, which is another source of variation in test performance from one occasion to another. A person might feel much better on one occasion than on another, might study in the domain of content, or might change attitudes. It is reasonable to think there is some fluctuation in abilities from day to day depending on a host of physiological and environmental factors. Even more expected are variations in moods, self-esteem, and attitudes toward people and issues. These changes would make

correlations between testing sessions less than would be predicted from the average correlations among items on each test. In this case, the correlation among the tests administered would be a better estimate of reliability than an internal item correlation from one test and one administration (Nunnally, 1978).

Thorndike compiled a list of test score variance by category which Cronbach (1970) has adopted and modified. Table 10 presents this list.

Table 10

Sources of Test Score Variance

I. Lasting and general characteristics of the individual

1. General skills (e.g. reading)
2. General ability to comprehend instructions, test-wiseness, techniques of taking tests.
3. Ability to solve problems of the general type presented in the test.
4. Attitudes, emotional reactions, or habits generally operating in situations like the test situation (e.g. self-confidence).

II. Lasting and specific characteristics of the individual

1. Knowledge and skills required by particular problems in the test.
2. Attitudes, emotional reactions, or habits related to particular test stimuli (e.g. fear of high places brought to mind by an inquiry about such fears on a personality test).

(table continues)

Table 10 continued

III. Temporary and general characteristics of the individual (systematically affecting performance on various tests at a particular time)

1. Health, fatigue, and emotional strain.
2. Motivation, rapport with examiner
3. Effects of heat, light, ventilation, etc.
4. Level of practice on skills required by tests of this type.
5. Present attitudes, emotional reactions, or strength of habits (insofar as these are departures from the person's average or lasting characteristics--e.g., political attitudes during an election campaign).

IV. Temporary and specific characteristics of the individual

1. Changes in fatigue or motivation developed by this particular test (e.g., discouragement resulting from failure on a particular item).

(table continues)

Table 10 continued

2. Fluctuations in attention, coordination, or standards of judgment.
3. Fluctuations in memory for particular facts.
4. Level of practice on skills or knowledge required by this particular test (e.g., effects of special coaching).
5. Temporary emotional states, strength of habits, etc., related to particular test stimuli (e.g., a question calls to mind a recent bad dream).
6. Luck in the selection of answers by guessing.

Note. From Cronbach (1970).

Most of the potential factors affecting test scores listed in Table 10 apply to the SWB scale. The factors which do not apply are those that have to do with abilities. Since the SWB scale is an attitudinal measure there are no right or wrong answers. Therefore, guessing, problem solving ability, memory for facts, and similar variables are not a concern with this scale.

Models of Reliability

There are several ways of examining the effects of unsystematic errors on a measurement instrument. Some examples are test-retest, alternate forms, inter-rater, and internal consistency. Each method takes into account the different conditions that can produce unsystematic changes in scale scores and affect the error of measurement (Aiken, 1979).

The two most appropriate models of estimating reliability for the SWB scale are test-retest and internal consistency. Because the SWB scale is an objective measure and does not have a parallel form, alternate form and inter-scorer reliability are not usable.

Test-Retest. The most direct way of estimating a test's reliability is to readminister the same test to the same group of examinees. After the second administration, a test-retest reliability coefficient, sometimes called a coefficient of stability, may be computed by correlating the scores from the two administrations (Aiken, 1979). Test-retest reliability indicates the extent to which scores on a test can be generalized over different occasions. The higher the reliability the less susceptible the scores are to the

random daily changes in the condition of the examinee or of the testing environment (Anastasi, 1988).

Random fluctuations from the different administration times are potential sources of variance or measurement error in test-retest studies. Fluctuations can occur in testing conditions, distractions, changes in the examinee in terms of motivation, health, practice effects, attention, emotional state, and reactions to testing and environmental conditions (Cronbach, 1970).

Two problems with test-retest reliability are practice effects, especially if the interval is short, and recall by the examinees of their former responses. Practice, or carry-over effects occur when the first testing session influences scores on the second one. Because of these effects, the scores from the two administrations of the test are not independently obtained and the correlation between them may be spuriously high (Anastasi, 1988). The interval between testing sessions must be selected and evaluated carefully (Kaplan & Saccuzzo, 1982).

D. Mueller (1986) notes that retest scores of attitudinal scales may legitimately differ from the original scores because of a real change in attitude

between testings. To counteract these problems it is desirable for the time between testing to be long enough for examinees to forget the details of the first testing but short enough so that little or no real change in attitude occurs between the testings. Mueller suggests a few weeks as a good compromise. Henerson, Morris, & Fitz-Gibbon, (1978) contend that a good rule of thumb is to wait one month between administrations.

Internal Consistency. The internal consistency method is useful for tests or scales measuring a homogeneous construct like spiritual well-being. An advantage of this type of reliability is that it requires only one administration of the instrument. A procedural difference also exists in this type from other estimation procedures because it does not use correlational statistics directly (D. Mueller, 1986). A conceptual difference is that an internal consistency coefficient describes similarity in measurement across items rather than stability over time or across forms.

The two most common internal consistency formulas are the Kuder-Richardson formula 20 (K-R 20) and the Cronbach alpha. The K-R 20 formula is for tests that have dichotomously scored items, such as aptitude tests

(right or wrong), or affective scales with items having only two response categories that are scored "1" and "0" (D. Mueller, 1986). Tests using items along a continuum, such as a Likert scale, require the use of the Cronbach alpha.

The alpha method examines the consistency of responses to all items in the test. This method divides the test into two halves using all the possible half-splits and taking the average of the reliability coefficients. In effect, alpha treats each item as an alternate test form and establishes the consistency of measurement across forms (D. Mueller, 1986). This can be a very time consuming task. For example, a test of 50 items requires computing 1,225 split-half reliability coefficients and then averaging them. The Cronbach alpha usually underestimates test reliability (Aiken, 1979).

Two sources of error variance can influence interitem consistency: content sampling, and heterogeneity of the behavior domain sampled. The more homogeneous the domain, the higher the interitem consistency (Anastasi, 1988). In internal consistency studies it is important for all items to be measuring the same construct. For other types of tests, such as

achievement or predictive tests, internal consistency is less useful.

Estimates of reliability based on the average correlation among items concern the internal consistency. This is partly a misnomer, because the magnitude of the reliability coefficient depends on both the average correlation among items (internal consistency) and the number of items. If the coefficient alpha is low, either the test is too short or the items have very little in common. If this is true, there is no need to make other estimates of reliability because they will be even lower (Nunnally, 1978).

Standard Error of Measurement

As stated earlier, when discussing measurement error the most useful model is one which considers any particular measure as composed of a random sample from a hypothetical domain of items (Nunnally, 1978). Thus, the 10 items of the RWB subscale and the 10 from the EWB subscale are thought of as a random sample of the many possible items that could be composed for that measurement domain, or universe.

According to Nunnally, it is not realistic to believe that every item in the universe had a chance to

be randomly selected for the scale. However, this is assumed for purposes of theory. Another assumption is that the purpose of any particular measure is to estimate the measurement that would be obtained if all the items in the universe were employed. The score that would be obtained from the whole universe is called the true score, or universe score. An individual would ordinarily have a different universe score for each universe measured.

When a single observation is taken and used as if it represented the universe, generalizing is used (Cronbach, 1970). If the observed scores from a particular procedure agree closely with the true score then it can be said the observations are accurate, or reliable. However, if there is a difference between a person's universe, or true score, and the score on one observation, an error of measurement has occurred. Since the tester does not know the person's universe score, a determination of the error in any particular observed score cannot be done. However, an estimate of how large the error tends to be can be made. This estimate is called the standard error of measurement.

Measurement theory assumes that each person has a true score for a particular universe, one that would be

obtained if there were no errors of measurement (Nunnally, 1978). Since there is some random error in the score obtained for a person on a particular occasion, obtained scores should differ from true scores on a random basis.

If it were possible to give many alternative forms of the same test the average score on the tests would closely approximate true scores. These scores would be randomly distributed above and below the true score, reflecting a normal distribution. Since it is expected these distributions of random errors are normally distributed, it is assumed that distributions of obtained scores will be normally distributed about true scores.

The wider the spread of obtained scores about true scores, the more error there is in employing the type of instrument. The standard deviation of the distribution of errors for each person is an index of the amount of error (Gravetter & Wallnau, 1988). The standard deviation of errors is assumed to be the same for all persons and is called the standard error of measurement. Presuming that all groups of people have similar standard error of measurement is a risky assumption and several scores representing different

groups should be reported based on the variable measured (Cronbach, 1970).

The standard error of measurement is an estimate of the standard deviation of the normal distribution of test scores that an examinee would obtain if she or he took the test many different times. The mean of this hypothetical distribution is the examinee's true score on the test.

The standard error of measurement (SE) increases as reliability decreases. When the reliability coefficient is +1.00 there is no error at all in estimating an individual's true score form the observed score; when the coefficient is .00 the error of measurement is at the maximum and equal to the standard deviation of observed scores in the sample (Aiken, 1979).

The standard error of measurement is particularly well suited to the interpretation of individual scores and sometimes more useful than the reliability coefficient (Anastasi, 1988). For example, an individual receives a score of 100 on the SWB scale and the SE is 1. Remember the assumption is that the individual's score is influenced by chance errors which fluctuate in a normal distribution about the mean of

their true score. Therefore, the probability the individual's true score lies somewhere between 99 and 101 is 68% (68% of scores will fall within one standard deviation above and below the mean). The probability the individual's true score is between 98 and 102 is 98% (the total of two standard deviations above and below).

The SE and the reliability coefficient are two ways of expressing test reliability. Unlike the reliability coefficient, the standard error of measurement is independent of the variability of the group on which it is computed (Anastasi, 1988). Expressed in terms of individual scores, it remains unchanged when found in homogeneous or heterogeneous groups. So, when comparing the reliability of different tests, the reliability coefficient is the better measure. To interpret individual scores, the standard error of measurement is more appropriate (Anastasi, 1988).

The SE is also useful when determining how accurate someone wants the test scores to be. In other words, if the test user wants to make fine discriminations among test takers then a smaller SE is more desirable.

Using the Reliability Estimate

The major use of reliability coefficients is in communicating the extent to which results obtained from a measurement method are repeatable. The reliability coefficient is one index of the effectiveness of an instrument, reliability being a necessary but not sufficient condition for any type of validity. In addition, the reliability coefficient is useful for making corrections for attenuation and confidence zones.

Measurement theory allows for corrections of correlations between two measures, treating them as if they had not been measured with error. The only information needed is the reliabilities of the two tests and the correlation between them (Kaplan & Saccuzzo, 1982).

The correction formula for attenuation must be used with caution because poor reliabilities and small samples can inflate correlations over +1.00. If the two variables have good reliability estimates, the attenuation formula can estimate more accurately the relationship between two traits, for example spiritual well-being and intrinsic religious orientation (Nunnally, 1978).

Establishing confidence zones for obtained scores is another use of the reliability coefficient. Using the standard error of measurement alone is an inaccurate way to establish the confidence zone because the SE zone lies symmetrically about the obtained score. It is inaccurate because obtained scores tend to be biased; high scores biased upward and low scores downward (Nunnally, 1978).

According to Nunnally, before establishing confidence zones, one must obtain estimates of unbiased scores. Unbiased scores are the average scores people would obtain if they were administered all possible tests from a domain, holding constant the number of items randomly drawn for each.

For example, if a person received a score on the SWB scale of 110 with the standard deviation 10 and the reliability .90 the estimated true score would be 9 in standard deviation units of SWB. If the mean was 100 then the estimated true score would be 109 for that individual and the 68% confidence zone would extend from 99-119. This information is rarely used by researchers because there seldom is a need for true scores or confidence zones. However, it is useful to have this data available if the need arises.

Interpreting Reliability Coefficients

At some point the question of what constitutes a satisfactory reliability coefficient level needs addressing. The answer depends on the proposed use for the measure (Nunnally, 1978). As a rule, desirable reliability estimates should be in the .80's or .90's (Anastasi, 1988). Nunnally (1978) argues that reliabilities should be modest, .70 or higher, when beginning to develop a measure of a construct. If significant correlations exist, corrections for attenuation will estimate how much the correlations will increase if the reliabilities of the measures are increased. Henerson, Morris, & Fitz-Gibbon (1978) consider reliability coefficients above .70 respectable, but lower coefficients are sometimes tolerated. The confidence of making decisions based on the results of the measurement is reduced when reliabilities are low.

For basic research, Nunnally argues efforts to increase reliabilities beyond .80 is often wasteful of time and funds. At that level measurement error affects correlations very little. To obtain a higher reliability coefficient might require strenuous efforts at standardization and increasing the number of items.

Thus, a more reliable test could be excessively time consuming to construct, administer, and score.

There are special problems associated with establishing a scale's reliability in attitude measurement because attitudes tend not to be as stable as skills. Test-retest coefficients must be interpreted with internal consistency data to determine whether to attribute differences to problems in the instrument or changes in the respondents attitude over time (Henerson, Morris, & Fitz-Gibbon, 1978). The importance of cross-validating reliability coefficients across groups is an important part of test development (D. Mueller, 1986).

Increasing Reliability

If increasing reliability is desired, there are at least three ways to do so: (a) altering the difficulty level of the items, (b) increasing test length, and (c) altering the heterogeneity of the group sampled (Aiken, 1979). Items of moderate difficulty have more variance than very difficult or very easy items.

Increasing the number of items also increases a scale's reliability. The Spearman-Brown formula can be used to estimate how much the reliability would increase by adding x amount of items. In reverse,

decreasing items will decrease reliability coefficients. That is why the split-half reliability coefficients are generally lower (because half of the test items are correlated with the other half). The Spearman-Brown prophecy formula attempts to correct for this phenomenon.

Another major influence on reliability coefficients involves the characteristics of the sample measured. The more varied a group of examinees are on a given variable the greater the variance on the test scores. With a homogeneous population, say a highly religious church group, the restricted range would show little relationship between two variables (Anastasi, 1988). This is also true for test-retest reliabilities. Respondents who are already very close together on the construct being measured are likely to reorder themselves on a later administration owing to unpredictable factors. However, when the ordering of the respondents on the construct is quite different they will be more likely to maintain a similar rank order because the unpredictable influences on the score are much smaller than the real, consistent differences of where they stand in relation to the construct (Henerson, Morris, & Fitz-Gibbon, 1978).

That the reliability of a test varies with the nature of the group tested is reflected in the practice of reporting separate reliability coefficients for different age, grade, sex, and occupational groups. Studying religious constructs may require separate coefficients for different beliefs and practices.

Summary

In summary, Cronbach (1970) suggests several general principles that apply to the interpretation of the reliability coefficient:

1. The coefficient tells what proportion of the observed score variance is non-error variance.

2. The coefficient depends on the spread of scores in the group studied.

3. The coefficient depends on the number of observations entering the person's score.

4. Other things being equal, a less accurate score is less valid.

The reliability studies done so far on the SWB scale were presented earlier. They consist of one internal consistency study and one test-retest study over a one week time span. Initial coefficient alphas for the fullscale and two subscales and test-retest

correlations were very good, tentatively suggesting a high reliability for the scale.

Many areas have not yet been explored with this scale including longer test-retest periods, standard errors of measurement, and cross-validation. So, while the initial studies are promising, more can be done in this area to increase confidence in the reliability of the scale.

The priorities for additional reliability studies with the SWB scale are as follows:

1. A test-retest study with a time span between one and two months.

2. Additional internal consistency alphas for the fullscale and subscales.

3. An examination of standard errors of measurement at the mean and standard deviations.

4. Test-retest, standard errors of measurement, and internal consistency coefficients for different demographic populations including age, gender, marital status, and religious variables.

The next section considers the issues involved in measuring responses and how to construct scales to measure attitudes.

Response Measurement

This section presents a brief overview of the issues and techniques underlying the construction of measuring instruments. The emphasis of the review is not to be exhaustive but to highlight the variety of considerations and decisions that must be made in this process. Topics include measurement and attitude theory, scaling techniques, and properties of rating scales. The specific properties of the present Spiritual Well-Being Scale will be highlighted and another response scale proposed for purposes of research to try to overcome problems of response distribution.

Measurement Theory

In administering assessment instruments, the goal is to gather information about the characteristics of the individual. These characteristics may be directly observable, such as height, or hair color and assessment may involve simply recording observations on a form. Usually, the characteristic is not directly observable, such as intelligence or spiritual well-being. In these cases, information obtained from assessment instruments is used to infer these characteristics.

Besides gathering information about an individual's characteristics, there is usually interest in discovering the particular amount of each characteristic that a person has (Reckase, 1984). This implies that the recording scheme used must quantify observations in some way. The resulting numerical scores not only indicate the level of each characteristic, but allows for comparisons among persons. Numerical scores give a convenient procedure for summarizing observations. They also lend themselves to further analysis that may help discover relationships that exist among different characteristics of a person.

Scaling is the process of assigning numbers to observations. If a particular scaling is successful, the numerical score obtained from an assessment instrument can be used to accurately infer the characteristics of a person.

According to Reckase (1984), scaling is the assignment of psychological meaning to a set of numbers. The basic concept in scale formation theory is that of a property. A property is defined by a set of entities, any set of entities can define a property. For example, the set of flowers defines the property

"flower." If an entity is a flower it belongs to that set. Psychological properties are defined by sets of people having the same level of the trait of interest. For example, a set of the people who all have the same amount of spiritual well-being defines that property of that level of spiritual well-being. Another set of people defines another, different level of spiritual well-being. A different set of people exists for each different level of spiritual well-being, and each of these defines a property.

All persons who have a property are equivalent on the trait of interest and different from those who do not have this same property. If a procedure can be developed to determine whether two individuals are equivalent on the trait studied, then the first step toward scale formation has been taken.

Natural variables are another issue in scaling theory. A natural variable is a collection of properties in which every entity is included in a property and no entity is in more than one property. It is called "natural" because it exists in the real world and does not have anything to do with abstract symbols such as numbers (Reckase, 1984).

All the variables commonly dealt with in psychology are natural variables. With the variable "spiritual well-being," it is assumed that at any moment in time many groups can be formed, each of which contain individuals equivalent in their level of spiritual well-being. All persons have some level of spiritual well-being, and no one has more than one level of spiritual well-being at any given time.

In order for the concept of a variable to be of use, some means must be determined to identify the particular set a person belongs to without going through the sorting process. The general procedure described by Reckase is to assign an abstract label to each property set and then develop a set of rules for determining the label that goes with each person. In other words, individuals can be grouped according to a number which has no natural connection to any underlying trait. This type of variable is a scaled variable. If a scaled variable can be linked to a natural variable, a very powerful relationship results. However, even though a variable is continuous in theory, the process of measurement always reduces it to a discrete one (Minium, 1978). Recorded measurements form a discrete scale.

Recorder measurements may be exact or approximate numbers. Exact numbers have no margin of error. Approximate numbers are not as exact and may result from rounding off or estimating. It can also result from making a continuous variable a discrete one. It is up to the investigator to determine the degree of accuracy appropriate to the problem.

Usually, more information about an individual is desired than whether he or she has a certain property. Generally, what is sought is the magnitude of the level. Therefore, some type of ordering is needed (Reckase, 1984). There are different levels of measurement available for use, commonly labeled nominal, ordinal, interval, and ratio.

Most of the variables in the behavioral sciences do not have the full properties of interval or ratio scales (Minium, 1978). For example, a person with an IQ of 150 is not thought to be twice as bright as one with an IQ of 75. In measurement, this problem may be particularly critical when a test does not have enough "top" or "bottom" to make adequate differentiation among the population being studied. A person who gets a top score on a scale may demonstrate his or her maximum level of attainment, while another who receives this

score may be capable of a much higher level but the measuring instrument is incapable of showing it.

The classification of scales as ordinal or interval takes on importance because psychometric theorists have pointed out that many common statistical procedures (such as the mean and standard deviation) require interval level data for proper application (Reckase, 1984). These procedures use the difference between scores to compute the descriptive statistics. Since the distance between scores is not clearly defined for ordinal scales, the meaning of the statistics for these scales is questionable.

Reckase (1984) asserts most psychologists consider ordinal scales as giving a reasonable approximation of an interval scale unless severe distortions in the scale proportion occur. According to Adams, Fagot, and Robinson (1965), statistical operations on measurements of a given scale are not appropriate or inappropriate per se, but only relative to the kinds of statements made about them.

Labovitz (1970) identifies three advantages of treating ordinal variables as if they were interval: (a) the use of more powerful, sensitive, better developed and interpretable statistics with known

sampling error; (b) the retention of more knowledge about the characteristics of the data; and (c) greater versatility in statistical manipulation, such as partial and multiple correlation and regression, analysis of variance and covariance, and most pictorial presentations.

Measuring Attitudes

An attitude is a psychological construct and like all psychological constructs, is hypothetical (D. Mueller, 1986). Since attitudes cannot be observed or measured directly their existence is inferred from their consequences. Psychological constructs must be observable and measurable by some means in order to be useful to researchers.

According to Henerson, Morris, & Fitz-Gibbon (1978), there are four approaches to evaluating attitudes. The first approach involves self-report measures and is the most direct type of attitude assessment. This is the method to employ unless there is some reason to believe the people studied are unable or unwilling to provide the necessary information. Examples of self-report measures are interviews, surveys, polls, questionnaires, attitude rating scales, logs, journals, and diaries.

A second technique uses the reports or assessments of others to measure a certain individual's feelings, beliefs, or behavior. Some examples of these types of measures are interviews, questionnaires, logs, journals, diaries, and observation procedures.

A third approach to evaluating attitudes involves sociometric procedures. This occurs when members of a group report on their attitudes toward one another and gives a picture of the social patterns within a group. Examples of this method are peer ratings and social choice techniques.

The final method takes records into account. It is useful when you have access to records that provide information relevant to the attitude in question and when these records are complete. Examples include counselor files and attendance records.

Since the focus is on the SWB scale the emphasis will be on the self-report measures and the techniques directly applicable to rating scales. The following section deals with some of the various models developed in the scaling field.

Models for Scaling Attitudes

The purpose of psychological scaling techniques is to assign numbers to individuals in such a way that a

scaling results (Reckase, 1984). That is, a rule must be developed for assigning numbers in such a way that most of the persons in the same property set of the natural variable receive the same number. Many different techniques have been developed for this purpose. Four of these techniques will be presented with emphasis on the Likert scale, the technique used with the SWB scale.

Guttman Scales. In deterministic models, such as the Guttman scalogram, each item is assumed to have a perfect relationship, of one kind or another, with a hypothetical trait. In the scaling of attitudes, the trait in question is the set of true scores for subjects on a particular dimension of an attitude (Nunnally, 1978).

The only deterministic model that has received widespread attention for the scaling of verbalized attitudes is the monotone deterministic model, usually referred to as the Guttman scale (Guttman, 1944). In Guttman's approach, the properties in a natural variable are ordered so that individuals in a higher level property include all the characteristics of those in lower level properties plus at least one more. The primary task is to find a series of behaviors such that

all those persons who exhibit a particular set of behaviors belong to the same property, and those in the next higher property exhibit at least one additional behavior. The emphasis is on the unidimensionality of the construct (D. Mueller, 1986).

The classic example of a Guttman scale is the measure of fear developed for use with soldiers in World War II. For that scale, those who did not experience "violent pounding of the heart" formed the lowest property set, while those who had formed the next higher property in the natural variable. If a sinking feeling in the stomach and violent pounding of the heart were reported, the person belonged in the next higher property. In all, 10 fear properties were defined in this way (Reckase, 1984). The scaled variable corresponding to the natural variable occurs by counting the number of characteristics present.

There are several criticisms directed at this type of measurement rendering it less suitable for measuring human traits. First, this scale is usually of the ordinal level because of its cumulative nature and deals only with dichotomous responses. Second, it is limited in its application because of the requirement of cumulative and unidimensional properties in the

natural variable, not taking into account the amount of unique variance in each item. Consequently, this makes it very difficult to find items to fit the model. Finally, deterministic models are mainly useful as theoretical reference points, for developing practical models for the actual scaling of attitudes (Nunnally, 1978).

Thurstone Scales. Similar to the Guttman scale, the nonmonotone model uses dichotomous responses to statements about attitudes. Each item represents, in a statistical sense, one point on an attitude continuum. According to Nunnally (1978), only persons in a narrow zone about that point should agree with an item; persons having either more positive or negative attitudes should disagree with the item. Ideally, one would expect the curve showing the probability of agreeing to the item as a function of the underlying trait to be a normal distribution.

The Thurstone scale (Thurstone, 1928) is the representative for this model, also called the method of equal intervals (Edwards & Kenney, 1946). Thurstone believed that properties in a natural variable are distinguishable and he developed a model of interaction between a person and statements

describing positive attitudes toward an object. A person who is a member of a particular property set will endorse some attitude statements and not others. Persons in a different property set will endorse a different, although possibly overlapping, set of alternatives. Those persons who endorse similar sets of statements belong to the same property.

By merely sorting persons into categories on the basis of the responses to a set of attitudes, a variable can be defined. However, this variable does not contain any information about the level of an attitude toward an object. In order to add the information about the relative level of attitude in the scaling, Thurstone suggested first scaling the attitude statements themselves.

The first step in this method is to identify and carefully delineate the attitudinal object (D. Mueller, 1986). A large number of statements about the attitude object are generated and given to judges who rate and classify the statements into the 11 piles ranging from highly unfavorable (scored a 1) through neutral (6) to highly favorable (11). If statements are in widely differing piles they are eliminated. The remaining statements are averaged according to their pile and

given a mean value. Approximately 25 items are selected which cover the entire continuum. These items are listed in random order on a scale. The score for the respondent is the mean of the scale values selected (Henerson, Morris, & Fitz-Gibbon, 1978).

The major advantage of this type of scale is that it permits a direct interpretation of the attitude of an individual, or the average attitude of a group of people, without recourse to general norms for the attitude in question (Nunnally, 1978). However, since researchers usually like to make comparisons and correlations with other measures, there is little need for a direct interpretation of the attitude of any one person in an absolute sense.

Another advantage of Thurstone scales is the existence of a zero or neutral point which allows for an "absolute" interpretation of scale scores rather than only "relative" interpretations. In Likert scaling there usually is no neutral point (D. Mueller, 1986).

The biggest drawback of Thurstone scales is the amount of effort required. The necessity for administration to a group of judges, totally separate from the administration to scale respondents, is enough

to tip the balance in favor of the Likert scaling method (D. Mueller, 1986). Other potential problems include the subjective ratings of the judges and the dichotomous nature of the measure.

Semantic Differential Scales. The semantic differential approach was originally developed by Charles Osgood and his colleagues but not for attitude measurement (D. Mueller, 1986). As Osgood studied the nature of meaning he observed the thousands of adjectives humans use to describe the world seemed to have considerable overlap. Using factor analysis he found that a large proportion of all meaning could be accounted for with three cognitive dimensions: evaluation, potency, and activity.

The semantic differential is Osgood's instrument for measuring the extent to which respondents attribute each of the several meaning dimensions to particular objects. For purposes of attitude measurement a special form of the semantic differential, consisting entirely of adjective pairs representing the evaluation dimension, is constructed (D. Mueller, 1986).

The semantic differential scale consists of a series of adjectives and their antonyms listed on opposite sides of the page with seven "attitude

positions" in between. At the top of the page, the attitude object is used as the heading. The attitude object may be a word, a phrase, or it may even be a picture. Each adjective pair contributes a score from 1 to 7 (the higher the score the more positive the response) and the sum of all the items represents the total score.

In developing this scale five steps are necessary (Henerson, Morris, & Fitz-Gibbon, 1978). First, determine the attitude object or objects to investigate. Next, select appropriate adjective pairs (approximately 10). Third, place them on a page under the attitude-object word or phrase, providing random polarity. Fourth, instruct respondents to rate the object quickly using their first impressions. Finally, compute the score by summing the responses.

The semantic differential is easy to construct, quick to administer, and usually highly reliable (D. Mueller, 1986). These scales correlate highly with Likert and Thurstone attitude scales. According to Mueller, one drawback is administration rapport. This is a concern if respondents don't like certain adjective pairs. Another drawback is validity since these scales can be faked very blatantly by the

respondents because of the transparent nature of the scale. Finally, since this scale uses only the evaluative dimension, only one dimension of the semantic differential theory is measured. This isolates the semantic differential scale from the theory behind its development.

Likert Scales. The Likert, or summative, model is the most useful type of scale with respect to psychological traits (Nunnally, 1978). It assumes only that individual items are monotonically related to underlying traits and that a summation of item scores is approximately linearly related to the trait. The total score comes from adding scores on individual items, whether they are dichotomous are multipoint (Nunnally, 1978).

The Likert procedure begins by assuming the existence of a natural variable with properties that can be ordered according to the magnitude of the trait possessed by the persons in each property set. The form of the Likert procedure is a statement about the concept in question, followed by five answer choices ranging from strongly agree to strongly disagree. An assumption is the five answer choices divide the

natural variable into five classes ordered with respect to the attitude toward the concept (Reckase, 1984).

If there is only one item in the measuring instrument, the five categories would be numbered from one to five and each person assigned the score corresponding to the response selected. If the statement rated has a negative connotation, scoring is reversed. The score assignment forms the scaled variable for this procedure.

In reality, there is usually more than one item in the Likert scale. It is assumed each item divides the natural variable in a similar, yet different, way. As more items are added to the instrument, the score for each person is obtained by summing the numbers assigned to each response category.

According to Nunnally (1978), when constructing a summative attitude scale there are five stages to follow. The first is the generating of an item pool which should contain no more than forty items evenly divided between positive and negative statements. Likert (1967) gives guidelines for the selection of these items. Some of these guidelines are: statements should be expressions of desired behavior and not statements of fact; use clear, concise,

straight-forward statements; word the statements so modal reactions spread the responses along the continuum; use positive and negative statements; and use only a single attitude variable for each item.

The second step in Nunnally's stages involves taking the item pool and administering it to a group of people similar to the group the measure is intended for. The sample size should be ten times the number of items or larger.

After administering the items, the third step involves correlating individual items to total scale scores, rank ordering according to these correlations, and eliminating items with low correlation. The fourth step is an optional weighting of the items. Nunnally believes this step does not add much to the end result. Finally, factor analysis is done to examine common factors.

Summative scales have many advantages over all the other models. According to Nunnally (1978), summative scales: (a) follow from an appealing model, (b) are rather easy to construct, (c) usually are highly reliable, (d) can be adapted to the measurement of many different kinds of attitudes, and (e) have produced meaningful results in many studies to date. In

addition, the ease of adding more items to summative scales contributes to the confidence of using these ordinal measurements as interval level.

Currently, the most widely used and easily constructed type of rating scale is the summative model. Edwards & Kenney (1946) in comparing the Likert (summative) and Thurstone techniques for scale construction concluded that Likert scales yield higher reliability coefficients with fewer items than scales constructed by the Thurstone method. They also concluded the Likert technique scale construction is less time consuming and less laborious.

Poppleton & Pilkington (1963) compared four methods of scoring an attitude scale (Likert, Thurstone, scale-product, and weighted proportions) and concluded that summation scores give higher reliabilities than limen scores (limen scores indicate a subject's central response tendency towards the attitude). Likert methods provide as high a reliability as the method of weighted proportions and better than the scale-product method. The Likert method also provided a good indication of validity.

Gorsuch (1984), in his review of research in the area of religion, claims that it does not matter which

scaling method is used. He cited research by Fishbein and Ajzen (1974) who used five different methods for measuring religion: Likert, Guttman, Thurstone, Self-Report, and Semantic differential. When the different methods were compared the intercorrelations among them were all about as high as reliabilities generally are. Gorsuch concluded the several different methods for questionnaire measurement of religion can be considered parallel forms of each other.

In summary, several techniques have been developed to measure attitudes, including Guttman's scalogram, Thurstone's equal-appearing intervals, Osgood's semantic differential, and Likert's summative model. Currently, there is debate over the most appropriate one to use. Likert scales are easy to construct and generally demonstrate reliability and validity coefficients as good as or better than the others. The Spiritual Well-Being Scale uses a Likert response format and there are no convincing arguments or reasons to change to one of the other major scales.

The next section highlights some of the issues to consider when constructing a Likert scale.

Properties of Rating Scales.

Anchor Words. According to Nunnally (1978), before respondents can use a rating scale, steps on the scales must be defined. The definitions of scale steps are called anchors. Among the different types of anchors available are numerical, degrees of agreement and disagreement, adjectives, actual behavior, and product scales.

Scales which use adjectives for anchors commonly employ bipolar adjectives, such as valuable-worthless, effective-ineffective. The semantic differential scale uses this type of anchor. A second type of anchor uses actual behavior. Actual behavior anchors are more useful for the rating of people rather than their attitudes. They are difficult to construct and judges often disagree on rating the behaviors.

A third type of scale, called a product scale, uses comparison of stimuli for anchors. For example, a product scale may be used for the judgment of handwriting. A six-step scale is employed, with each of the numbers 1 through 6 illustrated with samples of handwriting at different levels of legibility. Raters compare responses to the examples and mark the appropriate level.

Numerical scales, like their title suggests, use numbers for anchors. Numerical scales are advantageous because they remind the respondents of the meanings of the scale steps and facilitate the analysis of data. According to Nunnally, numerical anchors are often used simultaneously with other types of anchors, like words. A special type of numerical anchor is a percentage scale. On scales in this class, subjects rank themselves on a continuum ranging from 0 to 100 percent. The scale can be divided by 10 or 20 step intervals. Percentage scales usually are highly meaningful to subjects and make it very easy to formulate and communicate to the subject rating scales with many steps. This is frequently difficult with other forms of verbal anchors (Nunnally, 1978).

Another type of anchor is the kind the SWB scale uses. This type employs degrees of agreement or disagreement. These anchors are generally easy to work with, easily understood, and easily interpretable by the researchers. Superficially these scales may appear to be measuring judgments rather than sentiments. This is not the case because by responding to agreement scales respondents indicate their sentiments by

agreeing or disagreeing with favorable and unfavorable statements (Nunnally, 1978).

Number of Categories. In constructing rating scales the question arises of how many scoring categories and numbers to use. In a discussion of this problem, Guilford (1954) notes that if too few steps are used the scale is a coarse one and the discriminative power of the rater lost. However, if the scale is graded too finely it may be beyond the rater's limited powers of discrimination.

Many studies have been done to see if there is an optimum level of categories for a scale. McKelvie (1978), in a review of the research, observed that a very small number of categories, five or six, should be used. The five category scale was the most reliable for attitude judgment as measured by internal consistency. They also observed that a large number of scale categories (greater than 9 to 12) held no psychometric advantage over the five category scale and a small number of scales (less than five) evidenced a loss of discriminative power and validity.

McKelvie also conducted a study investigating continuous scales without categories or anchors and category scales. The subjects seemed to prefer the

continuous scales stating it allowed them to be more consistent and accurate when in reality it did not offer any advantage in terms of reliability or validity.

In choosing the number of rating categories the amount of discrimination needed is clearly a consideration (Garner, 1960). The SWB scale has been shown to have a low ceiling especially when administered to highly religious samples. This was demonstrated in Table 9.

The SWB scale uses six categories: Strongly Agree (SA), Moderately Agree (MA), Agree (A), Disagree (D), Moderately Disagree (MD), and Strongly Disagree (SD).

The scales are presented as follows:

1. (item statement) SA MA A D MD SD

Meyers (1986) discussed the lack of ceiling on the SWB scale and compared the effects on the present scale with a response scale like this:

Always true Never true

1 2 3 4 5 6

Using an available sample of 171 students from a Christian college who were randomly given one form of the scale or the other he found significant differences between the two scales on group means with the original

scale having higher means. The difference was not large enough to be of any practical value, however, and Meyers concluded a change was not warranted.

According to D. Mueller (1986), when considering the number of response categories, reducing the number reduces the spreading out of scores and thus tends to reduce reliability. Adding categories increases reliability because it adds more variance. However, there comes a point when respondents can no longer distinguish between adjacent categories. At this point, random error variance becomes a consideration.

The question of how many steps to use on a rating scale is very important when dealing with only one scale item. The number of steps issue is less important if scores are summed over many item statements (Nunnally, 1978).

Midpoints. Another issue regarding the number of steps or categories on ratings scales is whether to have an even or odd number. An argument for the odd number is that it permits the use of a middle step meaning "neutral," "neither," or "neither agree or disagree." This is thought to make subjects more comfortable in their ratings and allows the measurement of truly neutral responses (Nunnally, 1978). Arguments

against the use of a neutral midpoint are that it allows for an undesirable response style. Nunnally concludes the use of a neutral point apparently is not important, particularly if scores are summed over many items. The final decision is left largely to the judgment of the test constructor for the particular situation in which the ratings scales are employed.

Response Sets. A confounding variable affecting the validity of a measurement device is making sure the instrument is not measuring what it shouldn't be measuring (D. Mueller, 1986). One of the most pervasive measurement problems in this regard is response sets. Response sets are systematic response patterns based on considerations other than the content of the items. Two response sets which pose particular problems in affective measurement are acquiescence and social desirability.

The tendency to acquiesce, or agree, carries over into test taking behavior. If test scores are the result of an acquiescence response set rather than opinions about the statements, there is a reduction in validity. This type of response set is easy to control in a Likert scale by wording half the item statements positively and half negatively. This does not

eliminate acquiescence responding or even reduce it, but rather cancels out its effect.

Controlling for social desirability is a more difficult task. A social desirability response set is the tendency for test takers to make socially desirable responses to test items at the expense of responses based on true beliefs and preferences (D. Mueller, 1986). On many affective tests, identifying the most socially desirable response is easy.

The most direct way to reduce this effect is to establish rapport with the respondents and try to make them feel unthreatened by the measurement process. Assurances of anonymity or confidentiality will also reduce the threat.

Instructions. According to the Standards, directions for the test taker should be detailed enough so that test takers can respond to the task in the way which the test developer intends. If appropriate, sample material and practice or sample questions should be provided.

For the SWB scale these are the instructions stated on the test form:

"For each of the following statements circle the choice that best indicates the extent of your agreement or disagreement as it describes your personal experience."

These instructions are followed by the six categories and their labels and then the 20 item statements which make up the scale. Several studies using the SWB scale (Cooper, 1986; Davis et al., 1987; Ellison & Economos, 1981) have provided additional instructions which ask the participant to answer the questions as accurately as they can and to answer them as they are, not as they would like to be or think they should be.

General Appearance. Physical appearance is one of the least important considerations regarding the construction and selection of rating scale but one that does merit consideration (Nunnally, 1978). One choice is whether to place the scale horizontally or vertically on the page. Some have argued that the vertical scale is more familiar to the average person, as in reading a thermometer. In Nunnally's opinion, these and other variations on the physical appearance of rating scales apparently make little difference in the important psychometric properties of ratings. Such

differences are based more on esthetic preferences than on psychometric considerations.

According to Henerson, Morris, & Fitz-Gibbon (1978), the appearance of a questionnaire is very important because the first impression will effect the response rate. They suggest making the scale look easy to fill out, including as few questions as possible, making the response mode clear, and to not cram everything together.

Summary

In summary, the SWB scale in its current version is a Likert type scale with six verbal anchors of the agree-disagree type. No midpoint is given. One problem with the current rating scale is the tendency of religious samples to score near or at the top of the scale. This phenomenon does not allow for the finer discrimination among subjects desired. Research with a different rating scale is warranted to see if this effect can be minimized.

As this section has indicated, there are many decisions to make when constructing a response scale. Three basic decisions are the level of measurement, measurement approach, and scale model. The level of measurement involves choosing the kind of data desired.

The current SWB scale uses ordinal data for each item statement. When summed, this data approaches interval level and is treated as such for statistical purposes (Reckase, 1984). Ordinal data is a limitation in attitude measurement and not much can be done about it.

Another basic decision involves the measurement approach. The SWB scale relies on self-report which is the most direct approach of assessment (Henerson, Morris, & Fitz-Gibbon, 1978). No change is necessary here. Although other approaches are valuable, the SWB scale should stay with this approach.

The third basic decision involves which scaling model to use. The current version of the SWB scale uses the Likert technique. As presented earlier, there are strong arguments for the use of the Likert scale. The Guttman model has major flaws with its cumulative and unidimensional nature making it difficult to measure the construct of spiritual well-being. The semantic differential scale has its merits but the item statements of the SWB scale would have to be changed. It would not work with the present statements. The Thurstone scale is similar to a Likert scale but the extra amount of effort it requires and little advantage gained does not make it desirable (Edwards & Kenney,

1946). The Likert scale model appears to be the best one to use and should be retained.

Changes can be made with the Likert response scale to give it more variability while retaining the same anchor words. One modification can be made by changing the two middle anchors (agree and disagree) to add more clarity. The word "slightly" should be added in front of those anchors.

Because the current six-point scale does not seem to be as discriminating as it needs to be, using a scale that allows for more variability is warranted. A numerical scale that ranges from 0 to 100 would add variability. A potential problem with using too many categories is that respondents may not be able to distinguish between them. This should not be a problem if the six verbal anchors are retained and spread out on the scale. The respondents would make discriminations based on the six anchors with the ability to make more accurate responses within each of these categories.

Other changes can be made for the benefit of the test taker. For instance, respondents prefer midpoints (Nunnally, 1978). Giving them this option may help to build rapport. Also, using a percentage type of

numerical scale and making it easy to read by placing it like a thermometer can further this goal. Finally, presenting a response example and clearly labeling how and where to respond can serve to clarify questions and avoid missing responses.

Many modifications can be done with the response scale that retains the original intent and scaling model of the Spiritual Well-Being Scale while allowing for greater discriminations and variability with each item statement.

Purpose of the Study

This chapter has focused on the Spiritual Well-Being Scale, its psychometric properties, and psychometric theory. Three research questions were asked. The first question asked if a system could be devised for the SWB scale that presents the previous research conducted with the scale in an orderly manner according to some criteria. This was done using the Standards for Educational and Psychological Testing (1985).

The review indicated several technical areas of the SWB scale could use further research. Two of those

areas were selected for this study: reliability and response measurement. The second research question asked if additional data on the scale's reliability, consistent with earlier studies, could be generated. The literature revealed additional data would be useful to have for the SWB scale. These include a longer test-retest reliability study, more internal alphas, additional intratest correlations, and standard errors of measurement (SE).

The third research question addressed a specific problem with the SWB scale. The scale does not have a high enough ceiling for highly religious populations. The third question asked if the rating scale could be modified to minimize these ceiling effects and produce scores approximating a more normal distribution. A review of the literature highlighted the variety of decisions that go into measuring attitudes. The review also indicated changes could be made with the scale to allow for more variability and discrimination.

Considering this information, three separate studies were proposed. The first study created a test version of the Spiritual Well-Being Scale that modified the rating scale while retaining the same item statements. This scale was tested with a population of

highly religious people and examined for response distribution (measures of central tendency and variability), correlations with other measures of religiosity, internal consistency and intratest correlations. The original SWB scale was included for comparison purposes.

A second study examined test-retest reliability with the scale. Both versions of the scale were used but not with the same respondents. Each participant received one version of the scale and two other measures of religiosity. At a later time they received the same measures. The two scale versions were compared for response distribution (central tendency and variability), correlations with other measures, and measures of reliability. Two samples were selected for the study, a religious population and a non-religious population. A religious population targets the people the scale has the most difficulty with and the non-religious group brings more heterogeneity and spreads out the distribution (Anastasi, 1988).

A third study compared the findings from the above two studies with other samples. Since there are no norms available for the SWB scale, comparisons with other studies can reveal patterns and

contribute to the confidence of the results. Measures of response distribution (central tendency and variability), and internal consistency from three other samples were computed (Davis, Longfellow, Moody, & Moynihan, 1987; Huggins, 1988; Wong, 1989).

CHAPTER 2

METHODS

The present investigation consists of three separate studies, each of which examined the effectiveness of the rating scale and/or the reliability of the SWB scale. The first study compares the original and a test SWB scale with each other and with some other religious measures. Study two examines test-retest reliabilities for each SWB scale and correlates it with other religious measures. The third study uses archival data to examine internal consistency and other descriptive statistics with the original SWB scale.

The methods for collecting and statistically analyzing the data needed to achieve the research objectives are set forth in this chapter along with a description of a test rating scale developed for the SWB scale. Each of the studies are discussed in turn, in three sections: (a) participants, (b) research instruments, and (c) procedures.

Study One

The purpose of the first study is to gather additional information on the reliability and response measurement of the SWB scale. In this study a test version of the SWB scale, which retains the original items but uses a modified response scale, is introduced and examined using a religious sample. This study also used several other measures of religiosity including the original version of the SWB scale. Internal consistency correlations and measures of response distribution are computed.

Participants

The participants for this study were taken from a separate longitudinal study on spiritual growth begun in the spring of 1988 (Brinkman, 1989). The subjects were an available sample of volunteers from three churches. Two denominations were represented: Conservative Baptist and Evangelical Free Church of America, along with an independent church.

The number of participants in the study was dependent on how many volunteered. As a rule of thumb, a sample greater than 30 was sought in order for the

shape of the sample distribution to approach that of the population distribution (Gravetter & Wallnau, 1988). In all, 72 people completed the surveys. Ages ranged from the mid 20's to over 80, and 30 were male, 42 female. They were mostly middle-class and Caucasian, and regarded their faith, or religion, as very important in their lives.

Research Instruments

This study used five different measures of religiosity including two versions of the SWB scale, the original and a test version constructed by the researcher. In addition, the Concept of God scale (COG), the Spiritual Maturity Index (SMI), and the Religious Orientation Scale (ROS) provided evidence for construct validity. Since the original version of the SWB scale was thoroughly examined in the previous chapter it will not be presented here. The other four scales are discussed in this section along with four single item measures.

Spiritual Well-Being Scale (Form B)

The first chapter presented the rationale behind the modification of the SWB scale. The original

version of the scale is not useful for discriminating between highly religious people, a population heavily studied with the scale. Research with different versions is warranted.

The test version of the SWB scale is similar to the original scale in that it retains the 20 original items in the same order with the same wording. It is different in appearance, initial instructions, and rating scale.

The major change with the test version is the rating scale. Instead of six discrete categories (Strongly Agree, Moderately Agree, Agree, Disagree, Moderately Disagree, and Strongly Disagree) a continuous numerical percentage scale is given with six verbal anchors placed alongside. The anchor words are distributed along the scale and evenly spaced with the two numerical extremes (0, 100) anchored by the first and last anchors. Strongly Agree is alongside the number 100 at the top of the scale, Moderately Agree placed next to 80, Slightly Agree next to 60, Slightly Disagree next to 40, Moderately Disagree next to 20, and Strongly Disagree at the bottom next to 0. The two middle categories reflect a slight change in the verbal anchors. The word "slightly" was added

before each one. This was done to help clarify the categories.

A midpoint of 50 was placed on the scale without a verbal anchor. A copy of the scale and scoring instructions are in Appendix B.

The item statements remain unchanged, thus nine of the items are still worded negatively in order to avoid an acquiescence response style by the respondent.

The test version retains the same instructions that appear on the original version but adds a new paragraph to add clarity. The new version states:

> "For each of the following statements select a number between 0 and 100 from the scale shown on the left which best indicates the extent of your agreement or disagreement as it describes your personal experiences."
>
> "For example, if you read the statement "I like the color blue" and blue is your favorite color, you might record a 95. If you dislike the color blue, recording a 3 would be more appropriate; if blue was an o.k. but non-important color to you a score of 60 might reflect your

opinion best. Do the same with each of the items listed below:"

In addition, underneath the first item statement appears this instruction:

"(Select a number and write it on the blank line)."

Several changes were made in an attempt to improve the appearance of the SWB scale. For instance, it was printed on a laser printer to give it a more professional look. Another variation was the vertical aligning of the response spaces. The response spaces are immediately to the left of each item number and vertically aligned. This is to make it easier to see the response blanks and to reduce the probability of item omissions.

The response scale is located on the left side of the page and labeled as such. It is of a different print type and darker shading to distinguish it from the rest of the scale. It is also vertical and easy to glance at as each item is read and pondered. The numbers are on one side of the scale, like a thermometer, and the verbal anchors on the opposite side for guidance. This is a change from the original rating scale which was at the top with the rating categories for each item placed after the statement.

After constructing this version, it was pilot tested on a group of married couples (n = 12) in a Sunday school class. They were given the scale with the explanation it was a new scale and their feedback was desired. They completed the scale and commented it was very understandable and readable. The only negative feedback concerned the bad grammar of the item statements.

Concept of God Scale

The Concept of God scale (COG) is a 75 item self-report measure of one's concept of God as described in adjective ratings. Respondents rate their view of God by reporting whether they think that each adjective presented is strongly like God to strongly unlike God.

The development of the COG scale is still in process. It began when Spilka, Armatas, and Nussbaum (1964) administered 63 adjectives to a sample of religious undergraduate students and Catholic girls and found four or five common factors: stern father, the omni-ness of God, the impersonal God, the kindly father, and possibly the supreme ruler. They were

hesitant to match up these concepts, so Richard Gorsuch (1968) sought to replicate and expand on this study.

Gorsuch administered 91 adjectives using a three point scale (1 = does not describe God, 2 = describes God, and 3 = describes God particularly well) to 585 college undergraduates. Using factor analysis he discovered 11 factors from 75 items. Loading with an absolute value of .30 or better were 51 items for the Traditional Christian factor (TRA), 12 items under Benevolent Deity (BEN), 7 items under Companionable (COM), 12 items under Kindliness (KIN), 13 items under Wrathfulness (WRA), 5 items under Deisticness (DEI), 4 items under Omni-ness (OMN), 5 items under Evaluation (EVL), 4 items under Irrelevancy (IRR), 4 items under Eternality (ETR), and 3 items under Potently Passive (PAS).

Gorsuch reports no other validity studies or reliability coefficients for this scale. Other studies have used modified versions of the scale by altering the adjectives presented.

Lewis (1986) modified the rating scale for the COG by expanding it to six categories and reversing the order so that "1" equaled strongly like God and "6" equaled strongly unlike God. As a result, he had to

change all the correlation signs from negative to positive. He also inadvertently left three items off the scale. Lewis found strong correlations between the subscales of the COG and the SWB full and RWB subscale. He also found denominational differences in his sample, with Baptists obtaining significantly different scores than Unitarians on ten of the eleven subscales.

For this present study, the original 75 items identified by Gorsuch were used and the rating scale developed by Lewis retained, but the direction again reversed so that high scores indicated respondents felt this attribute was like God. A copy of the scale and scoring instructions are in Appendix C.

Spiritual Maturity Index

The Spiritual Maturity Index (SMI) is a 30 item scale developed by Craig Ellison in the early 1980's. According to the developer (Ellison, Rashid, Patla, Calica, & Haberman, 1984), the SMI was designed to measure spiritual maturity similar to the Spiritual Maturity Scale with the addition of attitudinal and behavioral criteria. Spiritually mature persons are autonomous, adhering to conventional beliefs on the

basis of their own experience and reflection. These individuals are able to transcend themselves, are self-principled, and able to cope with suffering. They find their own identity in relationship to God and religious beliefs and practices are an integral part of their daily activity. The SMI is supposed to measure degree of maturity rather than spiritual health (Bufford, 1984).

Using a rational process, Ellison developed criteria for spiritual maturity and then formed questions which individuals could respond to on a six point Likert scale similar to the SWB scale. The SMI began as a 20 item measure, but an additional 10 items were later added. Clarke, Clifton, Cooper, Mueller, Sampson, and Sherman (1985), in a study of the two versions, demonstrated the 10 extra items highly correlate with the 20 item version and load on the same factors.

Bressem (1986) reports a Guttman split-half reliability coefficient of $\underline{r}$ = .78, with a correlation between forms equal to $\underline{r}$ = .66, and an internal consistency coefficient alpha of $\underline{r}$ = .82 for the scale. The population Bressem drew from were 80 randomly

selected students at a private Bible college in Portland, OR.

Much of the investigative work so far with the SMI is in the area of validity, especially construct validity. The SMI correlates positively with many attitudinal and behavioral measures including self-esteem, feeling valued by God, perceiving the church as a caring community, and feeling there is a God given purpose in life (Ellison et al., 1984). It also correlates in the expected direction with the ROS intrinsic and extrinsic scales (Bufford, 1984). In one study, it did not correlate with measures of social desirability (Clarke, Clifton, Cooper, Mishler, Olson, Sampson, & Sherman, 1985).

Some of the behavioral single item measures the SMI has positively correlated with include devotional frequency and duration, frequency of church attendance, and involvement in Christian and non-Christian ministry (Bressem, 1986).

Bressem (1986) conducted factor analysis on the SMI items and found 10 factors with eigenvalues greater than +1.00. He used the principle components technique and four factors in a forced factor oblique rotation. However, commonality among the items was not evident.

Other factor analytic studies have found the SWB scale and SMI scale to be measuring the same general factor (Bufford, 1984; Cooper, 1986; Davis et al., 1987). Bufford reports the SMI positively correlated with the three SWB scales, sharing 68% common variance with the RWB subscale. There is some doubt the SMI is measuring something different from other measures of religiosity.

The SMI is constructed similar to the SWB scale with the same six agree/disagree categories. Half of the items are negatively worded. The total score is the sum of all the items after correction for reversal. Higher scores indicate greater spiritual maturity. No norms are available for the scale. See Appendix E for a copy of the SMI along with scoring instructions.

Religious Orientation Scale

The Religious Orientation Scale (ROS) is a 21 item self-report instrument originally designed to distinguish between a person's intrinsic (I) and extrinsic (E) religious orientations. It evolved at Harvard University under the guidance of Feagin (1964) and Allport and Ross (1967). The motivation behind the

ROS scale originally came from Gordon Allport's disturbing finding that religious people tended to be more prejudiced than non-religious individuals (Allport & Kramer, 1946). As he studied the phenomenon further he discovered a difference in prejudice between the intrinsically oriented church attender and the extrinsic church attender. A dichotomy was proposed between these two traits.

An extrinsic scale was designed to measure the degree to which a person's external social environment has influenced his or her personal religion. An intrinsic scale was created to measure the degree to which internal needs for creativity, strength, and direction shape an individual's religion.

Although one can obtain a single total score, it is customary to score the intrinsic and extrinsic subscales separately because for many respondents these constructs appear to be independent (Hunt & King, 1971). Studies done with a revised 20 item version indicate that it probably distinguishes among four types of religious orientations (Allport & Ross, 1967). These are intrinsic, extrinsic, indiscriminately pro-religious, and indiscriminately anti-religious. An intrinsically motivated person is more likely to live

his religion than use it. An extrinsically motivated person tends to view his religion as an activity which is instrumental in accomplishing other goals. Individuals high on both the I and E dimensions are labeled indiscriminately pro-religious, while individuals low on both dimensions are called indiscriminately anti-religious.

The internal consistency of this scale was examined on several occasions. Feagin (1964) reports item to scale correlations ranging from .22 to .54 when the whole scale was given one score. Item to intrinsic scale correlations ranged from .54 to .71 and item to extrinsic subscale .48 to .68. For Allport and Ross (1967) item to subscale correlations ranged from .18 to .58.

Several validity studies have been conducted with the ROS. The ROS has been able to distinguish prejudice in people, with respondents labeled prejudiced reporting more of an extrinsic orientation (Allport & Ross, 1967; Feagin, 1964). Allport and Ross also found that people who endorsed both extrinsic and intrinsic items were the most prejudiced of all. McClain (1978), in his study on personality and religious orientation, found intrinsically religious

persons scored significantly higher on self-control, personal and social adequacy, and stereotyped femininity. Individuals scoring low on the intrinsic scale (indicating a more intrinsic orientation) have higher scores on the SWB scales. High extrinsic scores (indicating a more extrinsic orientation) are associated with lower scores on the SWB fullscale and subscales (Ellison & Paloutzian, 1979). Robinson and Shaver (1973), in their review of the ROS, conclude that the Intrinsic-Extrinsic scale appears to consistently demonstrate its construct validity.

The ROS is administered individually to individuals who respond to the items as it applies to themselves. Items are scored from 1 to 5. In scoring, a 4 or 5 indicates an extrinsic orientation, a 1 or 2 an intrinsic orientation, and 3 assigned to any items omitted. Originally, the total score was simply the sum of the 21 items. However, most researchers now score separate items for the Intrinsic and Extrinsic subscales.

Several versions of the ROS exist with slightly different wording and responses. The version used for this study came from the review by Robinson and Shaver (1973). Norms are not available for this scale.

Low I scores indicate a more intrinsic orientation while high E scores suggest a more extrinsic orientation. A difference of more than 12 points between the two scales places respondents in the category of "indiscriminately proreligious." See Appendix F for a copy of the ROS and scoring instructions.

Single Item Measures

The four single item measures came from a previous study on religious variables (Davis, Longfellow, Moody, & Moynihan, 1987). These items inquired about how important respondent's considered their religious beliefs, how extensive was their religious knowledge, current degree of life satisfaction, and estimation of spiritual maturity. The purpose of including these items was for the study on spiritual growth from which this data was taken (Brinkman, 1989).

Procedures

Participants from three churches were asked through bulletin announcements and personal appeals to participate in a longitudinal study of spiritual

growth. They were informed the study would involve completing some surveys at that time and again a year later. In addition, following the second information gathering time, they would receive feedback on both sets of data so they could compare their individual spiritual growth as measured by these scales. Initial data was collected from January to May of 1988.

Of the 72 who volunteered to participate and completed the package, 30 were from a Conservative Baptist church in Vancouver, WA; 30 were from an Evangelical Free Church in Seattle, WA; and 12 were from two Bible study groups in Washington, DC. These churches participated because of contacts known to this researcher.

Of those who volunteered, the participants from Vancouver were given the questionnaires at a Sunday morning service which they filled out and returned to the church. Others received the material by mail with a cover letter (see Appendix J) and a stamped, self-addressed return envelope.

The church from Vancouver became involved in the study as part of a program for members to read the Bible through in one year. The senior pastor was contacted about trying to measure the anticipated

change in the members who completed the reading. Using bulletin announcements and announcements from the pulpit, volunteers were asked to participate and given the packet during a morning service. A box was provided at the church to place the completed surveys. Participants placed their names on the cover letter which was later numbered and separated from the surveys to protect confidentiality.

At the same time several other churches were contacted and asked to participate. One church in Seattle agreed and volunteers were solicited through a bulletin insert. In addition, members of two Bible study groups in Washington, DC also agreed to participate. From the names and addresses provided, a numbered survey was mailed for them to return. They were instructed not to place their names on the surveys. For those who did not mail back the survey within a few weeks, a reminder postcard was sent (Dillman, 1978).

The order of the measurement instruments was mixed so that some received the original SWB scale first and the test SWB scale later in the survey, while others received the test scale first and the original one later. There was no systematic procedure

to assure a truly random mix of the surveys. Besides the five measures, some single item questions concerning their religious life were included. The items inquired about the importance of religion, current religious knowledge, life satisfaction, spiritual maturity, and hours a week spent in ministry. It was planned to collect demographic information at the second data gathering. See Appendix G for a copy of the single item religious questions.

The SPSS/PC+ statistical software system was used to analyze the data (Norusis, 1986). Pearson product-moment correlations with two-tailed probabilities were computed using both the original and test SWB scale versions and the other measures (Gravetter & Wallnau, 1988). The statistical significance level was set at $p < .001$ (Norusis, 1986). Reliability coefficients were computed on the two forms of the SWB scale. These include the Cronbach alpha for internal consistency, the Pearson "r" for intratest correlations, and the standard error of measurement (SE).

Several statistics were calculated on the distribution of the individual scores. Measures of central tendency include the mean, median, mode, and

skewness. Also the number of respondents who received the maximum score and what percentage of the sample this is were noted. Measures of variability include the range, interquartile range (distance between the first and third quartile), and standard deviation.

Study Two

The purpose of the second study is to gather additional information about the reliability and response measurement of the SWB scale. The main objective is to conduct a test-retest reliability study of the test SWB-B and original SWB scale. In addition, construct validity coefficients, internal consistency correlations, and score distributions are examined. Three measures of religiosity were given to two samples, a religious group and a community college students.

Participants

The subjects for this study came from a Baptist church in Vancouver, WA and a community college in Gresham, OR. The number of participants in the study was dependent on how many volunteered. A sample

greater than 30 for each version of the SWB scale was desired so that the sample score distribution would approximate the population distribution (Gravetter & Wallnau, 1988).

The portion of the sample who completed both sessions consisted of 70 female and 50 male participants, mostly Caucasian, middle-class, and Christian. A total of 197 people completed at least one session. Approximately one-third of the sample was under age 20, one-third between 30 and 40. Roughly 40 percent were single, 45 percent married, and 40 percent had completed some college.

Research Instruments

Each participant was given three measures plus a demographic questionnaire to complete. Approximately half received the original SWB scale and the other half the test SWB scale version. In addition, each one was given the Concept of God scale and another measure of spiritual well-being, the Spiritual Distress scale. This scale was included because it was another measure in the domain of spiritual well-being and comparisons with the SWB scale were desired. Additional data on the reliabilities of both the Spiritual Distress scale

and Concept of God scale was gathered for other studies.

The SWB and COG scales have already been described. Following is a review of the Spiritual Distress scale and the demographic questionnaire.

Spiritual Distress Scale

The Spiritual Distress scale (SDS) is a 22 item self-report attitudinal survey designed to measure distress of the human spirit. Ruby Flesner developed the scale as part of her studies at Marquette University.

According to Flesner (1981), there is general agreement in the nursing profession that there is a relationship between unmet needs of the human spirit and the total well-being of an individual. Many nurses believe it is important to attempt to meet the spiritual needs of their patients but little research has been done in this area. In order to try to fill this void, Flesner developed her scale.

Spiritual distress is defined as "the painful and/or damaging effects of the stress that occurs to the mind and body of man when he is unable to adapt to an unmet need of the spirit" (p. 11). The very basic

universal need of the human spirit is to experience a dynamic relationship with God and through this relationship to experience forgiveness, love, hope, trust, and meaning and purpose in life.

Flesner, using these five dimensions, developed an item pool of questions that would be indicators of spiritual distress in relation to each dimension. Four statements from each of the five areas were eventually chosen, plus two additional statements which judges felt should be included as measuring or preventing distress. A total of 22 items make up the scale, half of which are negatively worded. The rating scale is a six point Likert type identical to the Spiritual Well-Being scale.

Evidence of the scale's reliability was examined with a test-retest study. The sample consisted of 88 first year nursing students (83 female, 5 male) who were asked to participate. The SDS, along with the SWB scale were given to this group twice with a one week interval. Eighty-three of those individuals participated in the second administration. Mean scores only were compared. There was a difference between means of about 1.7 percent for the SDS.

Evidence for construct validity was examined through a correlation with the SWB scale. Correlations reported from the first administration were (r = -.45), and from the second (r = -.90). Both were significant at the p <.001 level. The SDS did not significantly correlated with age or gender. A modest correlation (r = .22) existed between the SDS and religious participation.

Following this administration Flesner reworded some items and shifted the order of presentation. No studies are known to have been done with this revised version.

The test is administered in the same manner as the SWB scale with the same instructions listed at the top. Items are scored from 1 to 6 with higher scores indicating greater spiritual distress.

No norms are available for this test, nor are any other studies with this scale known. The mean score the original sample received was 49.2, with a standard deviation of 9.8 on the first administration. The sample had a mean of 49.2, and a standard deviation of 12.6 the second time.

In summary, the SDS has shown promise as another indicator of spiritual well-being. However, the author

suggests several tasks that should be done with the scale including factor analysis, internal consistency studies, additional test-retest correlations, and correlations with other scales. This present study can contribute towards these goals. See Appendix D for a copy of the scale along with scoring information.

Demographic Questionnaire

A demographic questionnaire compiled by the researcher was included for both testings. The first administration contained a one page questionnaire asking for data on age, gender, marital status, education, income, ethnic origin, religious affiliation, and personal estimates of spiritual maturity and well-being. The items were constructed in accordance with standards given by Dillman (1978). The item concerning estimation of one's spiritual maturity was taken from a previous study (Davis, Longfellow, Moody, & Moynihan, 1987). The item on estimation of one's own spiritual well-being was constructed for this study.

The one page questionnaire included with the second administration inquired about religious beliefs and practices, including beliefs about God, Jesus, and

the Bible. The first two belief questions came from a questionnaire used by Agnor (1986) and adapted from the Orthodoxy Index (Glock & Stark, 1966). The item about the Bible was a modified version of one used by Papania (1988). This item attempts to make a clearer distinction among individuals regarding their view of the Bible as the source for their belief. Other items include a rating concerning their identification with Christianity, how many years they have been a Christian (if they consider themselves one), and how often they participate in religious activities. The purpose for these additional questions was to try to make finer distinctions between a person's spiritual beliefs for analysis and understanding of results. Appendix H contains a copy of these items.

Procedures

The data collection took place on two separate occasions approximately six weeks apart. Several sources were contacted to participate in the study. Two psychology professors at one community college agreed to make announcements in their introductory psychology classes and give extra course credit (as one choice in their regularly offered extra credit program)

for those students who participated in both sessions. A room on campus was available for use during the lunch hour on October 19, 20, 26, and 27th, 1988.

Students interested in participating were instructed to come to the room and complete the assignment (see Appendix K). There was a sign posted outside the room informing students that research was in progress and to enter quietly. Two weeks prior to the second session, the professors were contacted and reminded. Another announcement was available for the professors to give to the class. The same room was used during the lunch hour on November 30th, December 1, and 7th, 1988 and the students again came in and completed the surveys. Following their completion they were given a handout explaining the study and given the opportunity to receive individual feedback (see Appendix M).

Besides the students from the college, the pastoral staff of a Baptist church in Vancouver, WA agreed to participate through their Sunday school program in the study. Each of the Sunday school class leaders were contacted by a pastor and the researcher to assure their participation and understanding and answer any questions. All the classes from high school age

on up participated except for the senior citizens class. Class members were informed of the study during the regular Sunday class time where they also completed the packet. The first session was on October 23, 1988 and the second on December 4, 1988. For those who missed the second session addresses were looked up in the church directory and they were mailed a copy with a cover letter and stamped, self-addressed envelope. During the second session, people who had missed the first session were given one packet to complete and the data included in some of the analysis.

At both sites, each person who agreed to participate received a manila envelope that contained a four page survey packet and an index card. Each packet contained, in order, the SWB scale, the COG scale, the SDS, and the first demographic sheet. The envelopes were arranged using a random number table so that a random distribution of original and test scales of the SWB occurred.

Instructions were given verbally (see Appendix L). The participants were asked to open the envelope and place their names on the index card so surveys could be matched for the second testing. They were instructed not to put their names on the surveys themselves.

The participants were then asked to complete the surveys and when finished to put it back into the envelope and turn it in along with the index card. All were informed this study involved a second session a few weeks later but were not told it involved completing the same tests. If someone did not understand an item they were told to leave it blank.

Between sessions, the surveys were numbered with a number placed on the index card and the face sheet. The data was entered into a data base and scored and the scores placed on the tests.

At the second session the participants were given a manila envelope with the index card they had completed stapled to the outside. The second set of instruments in the same order were inside this packet. The final page was the second demographic sheet. Each survey was numbered to match the first survey. Again subjects were asked not to put their name on them and to remove the index cards from the envelope.

For the church sample, a sealed envelope with the scored scales from the first administration along with a sheet explaining the purpose of the study was in the envelope. After the subjects had completed the surveys they were given the opportunity to compare them with

the first session and to ask questions. Group data was also available for their information. They were instructed to keep the index card with their name and number if they wished to discuss the results later because no master list existed that had this information on it.

For the community college sample, they were given the same sheet explaining the purpose of the study after they had finished the second session. They were also given an opportunity to sign up for an individual appointment or to give their name and phone number to discuss results of the study. Names of those completing both sessions were submitted to the professors for credit. Results from the study were made available to the professors and pastors for their use.

The data was analyzed using the SPSS/PC+ statistical software system (Norusis, 1986). Pearson product-moment correlations with two-tailed probabilities were computed using both the original and test SWB scale versions and the other measures (Gravetter & Wallnau, 1988). The level of statistical significance was set at $p > .001$ (Norusis, 1986). Reliability coefficients were computed on the two

versions of the SWB scale. These include the Cronbach alpha for internal consistency, the Pearson "r" for intratest correlations and test-retest reliabilities, and the standard error of measurement (SE).

Several statistics were calculated on the distribution of the individual scores for both the SWB fullscale and two subscales. Measures of central tendency include the mean, median, mode, and skewness. Also examined were the number of respondents who received the maximum score and the percentage of the sample this is. Measures of variability include the range, interquartile range (distance between the first and third quartile), and standard deviation.

A one-way analysis of variance was calculated looking for differences at the .05 level between the samples. The church sample was divided into students (college and high school classes) and adults for this analysis.

Study Three

The purpose of the third study is to gather additional data about the internal consistency and response distribution of the original SWB scale.

Archival data was used from three previous studies (Davis, Longfellow, Moody, & Moynihan, 1987; Huggins, 1988; and Wong, 1989). The samples were selected because they represented large religious populations. The Wong data represented an ethnic group other than Caucasian.

Participants

The Davis et al. sample consisted of 330 participants who were attending the Sunday school of a Missionary Alliance church one Sunday morning in the fall of 1987. The study involved a factor analysis of the SWB scale and SMI together. Of the 350 surveys completed, 331 were usable for this present study. The sample consisted of 148 males, 170 females. The average age was 38 with the range from 18-83. A large percentage (98%) reported being High School graduates, 46 percent from college and 75 percent of the respondents were married, 16% never married. The majority professed to be born again Christians.

The Huggins sample was randomly taken from the population of people who attended Conservative Baptist churches in Oregon and had their names listed in their church directories. They were part of a study

examining the effects of small groups on SWB scale scores. The sample consisted of 297 subjects (response rate of 65%). Of those, 112 were male and 172 female. The average age was 47 and the range was from 19-88. Of the respondents, 69 percent reported being married, 11 percent never married, and 90 percent were High School graduates, 37 percent from college.

The Wong sample came from a mail survey conducted in the fall of 1988. The sample consisted of 72 ethnic Chinese Americans who attended churches in the Pacific Northwest. Wong studied the relationship between spiritual well-being, social desirability, and self-esteem. Thirty-seven were female, 35 male. The average age was 39, ranging from 19 to 76. Over half of the sample reported some college education and 64% were married.

Research Instruments

The only instrument studied was the original version of the SWB scale which was discussed earlier.

Procedures

The data for these studies had already been collected and was available for secondary analysis.

The data for the Davis et al. sample was collected during the Sunday school hour in the fall of 1987. The Huggins sample was a mail survey conducted about the same time and the Wong sample a mail survey from the fall of 1988.

Using the SPSS/PC+ statistical program (Norusis, 1986), an internal consistency (Cronbach alpha) was run on the SWB full scale and subscales from both samples. Other measures of reliability computed were the Pearson "r" for intratest correlations, and the standard error of measurement.

Several statistics were calculated on the distribution of the individual scores. Measures of central tendency include the mean, median, mode, and skewness. The number of respondents who received the maximum score and what percentage of the sample this is was noted. Measures of variability include the range, interquartile range (distance between the first and third quartile), and standard deviation.

CHAPTER 3

RESULTS

This chapter will present an overview of the findings from each study separately and then combine the data under descriptive statistics, reliability, and response measurement. For each study, the demographic and descriptive information are presented along with correlations between the other measures.

Study One

The sample consisted of volunteers for a longitudinal study of spiritual growth (Brinkman, 1989). Of those who agreed to participate, 72 returned completed surveys. There were 30 males and 42 females in the sample. The majority were married, Caucasian, and middle-class. More specific demographic data was not collected in the first administration. The demographic information was part of the second phase of the study which has not yet been completed. Descriptive data for both versions of the SWB scale,

the Spiritual Maturity Index, the Concept of God scale, the Religious Orientation Scale, and some single item measures is presented in Table 11. The descriptors include two measures of central tendency: the mean and median; two measures of variability: standard deviation (SD) and range (minimum and maximum); and a measure of reliability: the standard error of measurement (SE).

Table 11

Descriptive Data for Study One Measures

Test	Mean	Median	SD	SE	Min	Max
RWB	54.2	55.5	5.9	.69	40	60
EWB	51.1	52.0	6.6	.77	36	60
SWB	105.3	108.0	11.3	1.34	77	120
RWB-B	88.6	91.0	10.5	1.23	54	100
EWB-B	83.5	84.8	12.3	1.45	44	100
SWB-B	172.1	176.3	20.2	2.38	98	199
COG						
Traditional	292.4	298.0	16.1	1.95	218	306
Benevolent	68.5	71.0	5.1	.62	48	72
Companionable	39.5	41.0	3.8	.46	24	42
Kindliness	69.2	71.3	5.5	.66	40	72
Wrathful	45.2	45.0	13.7	1.66	18	72
Deisticness	7.8	6.0	3.6	.44	5	17
Omni-ness	23.5	24.0	1.9	.23	13	24
Evaluation	28.1	29.5	2.8	.34	19	30
Irrelevancy	4.3	4.0	1.0	.12	4	11
Eternality	23.9	24.0	.5	.06	22	24

(table continues)

Table 11 continued

Test	Mean	Median	SD	SE	Min	Max
COG (continued)						
Passive	11.1	11.0	2.9	.36	3	18
SMI	142.5	142.0	16.4	1.95	100	171
ROS-I	13.4	13.0	3.3	.39	9	25
ROS-E	20.7	20.0	5.1	.60	12	35
RELB	6.4	7.0	.8	.09	4	7
RELK	4.9	5.0	1.2	.15	2	7
LIFSAT	5.0	5.0	1.1	.13	2	7
SPMAT	4.7	5.0	.8	.09	2	6
MINHOURS	9.1	4.5	13.4	1.67	0	70

Note. N = 72. RWB, EWB, SWB = Spiritual Well-Being Scale; RWB-B, EWB-B, SWB-B = Test SWB scale; COG = Concept of God scale (subscales listed) SMI = Spiritual Maturity Index ROS-I, ROS-E = Religious Orientation Scale RELB = Importance of Religious Beliefs RELK = Amount of Religious Knowledge LIFSAT = Life Satisfaction; SPMAT = Estimation of Spiritual Maturity; MINHOURS = Hours/week in ministry

Several scales correlate significantly with the SWB fullscale. Pearson product-moment correlations over .40 with two-tailed significance levels greater than .001 include the SWB subscales, SMI, ROS-I (negative), ROS-E (negative), COG subscales Traditional Christian, Benevolent Deity, Companionable, Kindliness, Deisticness (negative), and Evaluation. Single item measures with high correlations include Importance of Religious Beliefs, Life Satisfaction, and Estimation of Spiritual Maturity. Similar correlations were obtained with the test version of the SWB fullscale.

The SWB Religious Well-Being subscale (RWB) correlates greater than $r = .40$ ($p > .001$) with the SMI, ROS-I (negative), Importance of Religious Beliefs, Life Satisfaction, Estimation of Spiritual Maturity, and COG subscales Traditional, Benevolent, Companionable, Kindliness, Deisticness (negative), and Evaluation. The test RWB subscale (RWB-B) significantly correlates with the same scales as the original RWB and also the COG Omni-ness subscale and ROS-E (negative) scale.

The Existential Well-Being (EWB) subscale has correlations greater than .40 ($p < .001$) with the SMI, and the single item measures of Importance of Religious

Beliefs, and Life Satisfaction. The test EWB subscale (EWB-B) significantly correlates with the same measures as well as the COG subscale Benevolent.

Table 12 lists the Person product-moment correlations between the original and test SWB scales and the other measures with two-tailed significance levels of $\underline{p} < .01$ (*) and $\underline{p} < .01$ (**).

Table 12

Correlations Between SWB Scales and Data from Study One

Test	RWB	EWB	SWB	RWB-B	EWB-B	SWB-B
COG						
Traditional	.59**	.26	.46**	.69**	.34*	.56**
Benevolent	.59**	.39*	.53**	.66**	.41**	.59**
Companionable	.57**	.27	.45**	.70**	.33*	.57**
Kindliness	.56**	.27	.45**	.63**	.34*	.54**
Wrathful	-.08	-.17	-.14	-.12	-.27	-.22
Deisticness	-.48**	-.39**	-.48**	-.52**	-.36*	-.48**
Omni-ness	.37**	.14	.28	.39**	.12	.27
Evaluation	.48**	.34*	.45**	.49**	.38**	.49**
Irrelevancy	-.30*	-.13	-.23	-.27	-.13	-.22
Eternality	.24	.06	.16	.28*	.09	.20
Passive	-.03	.04	.01	-.07	-.16	-.13

(table continues)

Table 12 continued

Test	RWB	EWB	SWB	RWB-B	EWB-B	SWB-B
SMI	.76**	.54**	.71**	.73**	.54**	.69**
ROS-I	-.54**	-.39**	-.51**	-.49**	-.39**	-.47**
ROS-E	-.37**	-.23	-.32*	-.45**	-.23	-.39**
RELB	.66**	.52**	.64**	.67**	.53**	.67**
RELK	.26	.15	.22	.17	.03	.11
LIFSAT	.49**	.64**	.62**	.43**	.53**	.54**
SPMAT	.41**	.35*	.42**	.44**	.32*	.42**
MINHOURS	.31*	.06	.19	.27	-.13	.07

*p. < .01 **p. < .001

Note. N = 72. RWB, EWB, SWB = Spiritual Well-Being Scale; RWB-B, EWB-B, SWB-B = Test SWB scale

COG = Concept of God scale (subscales listed)

SMI = Spiritual Maturity Index

ROS-I, ROS-E = Religious Orientation Scale

RELB = Importance of Religious Beliefs

RELK = Amount of Religious Knowledge

LIFSAT = Life Satisfaction

MINHOURS = Hours/week in ministry

SPMAT = Estimation of Spiritual Maturity

Study Two

The second study focuses on the test-retest reliability of the SWB scale. In addition, internal consistency alphas, measures of central tendency and variability, and frequency distributions for both the original and test version of the scale are examined.

There were two samples for this study. The community college sample consists of volunteers from two introductory psychology classes who were invited to participate by their professors and given extra credit for completing both sessions. Sixty-six students came for the initial session and 42 returned approximately six weeks later to complete the second session. Twenty-seven students came from one of the psychology classes, 35 from the other one, and four students came from other classes. For two of the four students from other classes, their professor agreed to give them extra credit. The professor for the other two participants refused to offer credit and they did not return for the second session.

The church sample consists of five Sunday school classes at a Baptist church. The classes were high school (n = 35), college (n = 8), ladies (n = 9), young

adults (n = 33), and middle age (n = 46). No one refused to participate in the study. The total number of participants from each class to complete both sessions are as follows: high school (n = 26), college (n = 3), ladies (n = 6), young adults (n = 19), middle age (n = 25). The total number from the church to complete both sessions was 79.

Of the 131 people from the church who completed the packet at least once, 28 did so at the second session. Twenty-four were present at the first session but not at the second. For these absentees, names were looked up in the church directory and those who had addresses listed (n = 14) were mailed the second packet with instructions and a stamped, return envelope. Four of those were retuned.

A one-way analysis of variance was calculated using the original SWB scale as the dependent variable and testing for differences between three groups. For this analysis, the church sample was divided into two groups: students (the high school and college classes) and adults. Results indicate the groups are significantly different for the SWB fullscale at the .01 level: F (2, 104) = 4.08, p < .019. A Scheffe

a posteriori test concluded no two groups are different at the $p > .05$ level.

Table 13 presents demographic data for three separate groups: (a) those taking the original SWB scale both times, (b) those completing the test SWB scale both times, and (c) combined data for everyone completing at least one session. The last group includes those at the first session who missed the second and those at the second session who missed the first.

The demographic data from all of the questionnaire items is included in this table which takes up the next several pages. Under each of the three groupings the number and percentage of the sample who circled that category are listed. One of the questionnaire pages was given out during the second session and a large number of people did not have the opportunity to respond to those items. This is reflected in the missing data category. For a copy of the demographic items see Appendix H.

Table 13

Demographic Data from Study Two

	Total		Original SWB		Test SWB	
Category	Freq	Pct	Freq	Pct	Freq	Pct
Age						
Under 20	68	35%	18	34%	25	37%
20-29	29	15%	6	11%	7	10%
30-39	58	29%	19	36%	27	40%
40-49	21	11%	7	13%	8	12%
Over 50	8	4%	1	2%	1	1%
Missing	5	3%	2	4%	0	0%
Gender						
Female	116	59%	32	60%	38	56%
Male	78	40%	21	40%	29	43%
Missing	2	1%	0	0%	1	1%
Marital Status						
Single	78	40%	22	41%	27	40%
1st Marriage	88	45%	22	41%	31	46%
Sep/Divorced	9	5%	3	6%	3	4%
Remarried	13	7%	5	10%	4	6%

(table continues)

Table 13 continued

Category	Total Freq	Total Pct	Original SWB Freq	Original SWB Pct	Test SWB Freq	Test SWB Pct
Marital Status (continued)						
Live Together	5	3%	1	2%	3	4%
Family Income						
< $10,000	20	10%	8	15%	6	9%
$10-20,000	29	15%	7	13%	11	16%
$20-30,000	34	17%	11	21%	6	9%
$30-40,000	51	26%	9	17%	25	37%
$40-50,000	21	11%	4	8%	6	9%
Over $50,000	20	11%	7	13%	8	12%
Missing	22	11%	7	13%	6	9%
Education						
< High School	23	12%	7	13%	12	18%
High School	37	19%	8	15%	11	17%
Trade/Bus	10	5%	1	2%	5	8%
Some college	80	41%	26	50%	22	32%
College Grad	17	9%	3	6%	9	13%
Some Graduate	6	3%	1	2%	2	3%

(table continues)

Table 13 continued

Category	Total Freq	Total Pct	Original SWB Freq	Original SWB Pct	Test SWB Freq	Test SWB Pct
Education (continued)						
Grad Degree	13	7%	6	11%	4	6%
Missing	11	6%	1	2%	3	4%
Ethnic Heritage						
Black	3	2%	0	0%	1	2%
Native Amer	1	1%	1	2%	0	0%
Oriental	1	1%	0	0%	1	2%
Caucasian	185	94%	52	98%	64	94%
Other/Missing	7	4%	0	0%	2	3%
Religious Identification						
Catholic	5	3%	1	2%	1	2%
Jewish	1	1%	1	2%	0	0%
Protestant	114	58%	31	59%	44	65%
Other	53	27%	13	25%	15	22%
None	18	9%	6	11%	8	12%
Missing	6	3%	1	2%	0	0%

(table continues)

Table 13 continued

Category	Total Freq	Total Pct	Original SWB Freq	Original SWB Pct	Test SWB Freq	Test SWB Pct
Belief In God						
Don't believe	3	2%	1	2%	2	3%
Higher Power	1	1%	0	0%	1	2%
Sometimes	3	2%	3	6%	0	0%
More/less	16	8%	6	11%	9	13%
No Doubts	110	56%	42	80%	51	75%
None/Missing	64	32%	1	2%	5	7%
Belief In Jesus						
Don't believe	3	2%	1	2%	2	3%
Only a man	1	1%	0	0%	1	2%
Basically	11	6%	5	9%	4	6%
Divine Son	118	60%	43	81%	59	87%
None/Missing	64	32%	4	8%	2	3%

(table continues)

Table 13 continued

Category	Total Freq	Total Pct	Original SWB Freq	Original SWB Pct	Test SWB Freq	Test SWB Pct
Christian Profession						
Not Christian	10	5%	5	9%	5	7%
Moral/ethical	7	4%	5	9%	1	2%
Christ Savior	14	7%	4	8%	9	13%
Follow Christ	105	53%	37	70%	52	77%
Missing	61	31%	2	4%	1	2%
Years A Christian						
1-4	11	6%	3	6%	8	12%
5-9	23	12%	11	21%	9	13%
10-19	37	19%	16	30%	17	25%
20-30	35	18%	10	19%	22	32%
Over 30	12	6%	2	4%	4	6%
Missing	79	40%	11	21%	8	12%

(table continues)

Table 13 continued

Category	Total Freq	Total Pct	Original SWB Freq	Original SWB Pct	Test SWB Freq	Test SWB Pct
Belief In Bible						
Not needed	5	3%	2	4%	3	4%
Ultimate	101	51%	35	66%	49	72%
Experience	6	3%	2	4%	4	6%
Chrch hierarchy	3	2%	1	2%	2	3%
Other sayings	3	2%	3	6%	0	0%
Don't know	9	5%	3	6%	6	9%
None/Missing	70	36%	7	13%	4	6%
Religious Participation						
< 1/year	7	4%	5	9%	2	3%
1-2/year	8	4%	4	7%	4	6%
3 to 11/year	7	4%	3	6%	4	6%
1 to 3/month	5	3%	2	4%	2	3%
Weekly	24	12%	6	11%	12	18%
> weekly	87	44%	32	60%	43	63%
Missing	59	30%	1	2%	1	2%

(table continues)

Table 13 continued

Note. Some of the demographic questions (From "Belief in God" to "Religious Participation") were only included in the second session so there is a large number of participants who did not have the opportunity to answer them.

There were three measures of religious belief given in this study. Respondents had the opportunity to answer one of two versions of the Spiritual Well-Being Scale (SWB), the Concept of God (COG) scale, and the Spiritual Distress Scale (SDS). In addition some single item measures of religious belief and behavior were included.

Table 14 presents descriptive data for these variables. The table is separated into three sections: descriptive data for the (a) total sample, (b) for the first session of those completing both sessions, and (c) for the second session of those completing both. The descriptors include the mean and median, the standard deviation (SD) and range (minimum and maximum), and the standard error of measurement (SE).

Table 14

Descriptive Data for Study Two Measures

	Total Sample						
Test	Mean	Median	SD	SE	Min	Max	N
RWB	50.8	55.0	10.1	.98	10	60	107
EWB	49.0	50.0	7.4	.71	25	60	107
SWB	99.8	103.0	15.1	1.46	53	120	107
RWB-B	79.1	84.0	21.7	2.29	60	100	90
EWB-B	77.9	80.5	14.1	1.48	31	100	90
SWB-B	156.9	161.3	30.4	3.20	65	200	90
SDS	53.5	51.0	15.3	1.11	29	103	191
SPMAT	4.6	5.0	1.2	.08	1	7	193
SPWB	4.9	5.0	1.2	.09	1	7	191
COG							
Traditional	280.3	294.5	37.7	2.83	51	306	177
Benevolent	65.1	69.0	9.4	.70	27	72	180
Companionable	66.9	71.0	9.3	.69	12	72	182
Kindliness	38.1	40.0	5.8	.43	7	42	180
Wrathful	45.5	45.5	13.9	1.06	13	78	175
Deisticness	10.0	7.5	6.1	.45	5	30	179

(table continues)

Table 14 continued

Test	Mean	Median	SD	SE	Min	Max	N
Total Sample (continued)							
COG (continued)							
Omni-ness	21.7	24.0	3.9	.30	4	24	171
Evaluation	26.9	28.0	4.3	.32	5	30	180
Irrelevancy	5.4	4.0	2.9	.22	4	24	184
Eternality	23.1	24.0	2.8	.21	4	24	183
Passive	10.7	11.0	3.6	.27	3	18	181
First Session of Test-Retest							
RWB	52.0	56.0	9.6	1.32	10	60	53
EWB	51.1	51.0	5.3	.73	38	60	53
SWB	103.2	106.0	12.5	1.71	65	120	53
RWB-B	81.0	85.2	20.4	2.48	6	100	68
EWB-B	77.8	80.5	14.0	1.69	31	100	68
SWB-B	158.8	161.8	28.6	3.47	85	200	68
SDS	50.4	48.8	12.9	1.19	29	85	116
SPMAT	4.7	5.0	1.2	.11	1	7	119

(table continues)

Table 14 continued

First Session of Test-Retest (continued)							
Test	Mean	Median	SD	SE	Min	Max	N
SPWB	5.1	5.0	1.2	.11	1	7	119
COG							
Traditional	282.0	295.0	39.6	3.79	51	306	109
Benevolent	65.8	69.5	9.1	.86	27	72	111
Companionable	67.4	71.0	9.5	.89	12	72	112
Kindliness	38.6	41.0	5.9	.57	7	42	111
Wrathful	45.5	46.0	14.4	1.39	13	78	108
Deisticness	9.6	7.5	5.6	.54	5	30	110
Omni-ness	22.1	24.0	3.7	.37	4	24	102
Evaluation	27.1	29.0	4.3	.41	5	30	111
Irrelevancy	5.2	4.0	2.9	.28	4	24	112
Eternality	23.1	24.0	3.2	.30	4	24	112
Passive	10.8	10.5	3.8	.36	3	18	111

(table continues)

Table 14 continued

Test	Second Session of Test-Retest Mean	Median	SD	SE	Min	Max	N
RWB	53.4	57.0	7.8	1.08	28	60	53
EWB	52.2	53.0	5.9	.81	40	60	53
SWB	105.7	110.0	10.9	1.50	84	120	53
RWB-B	80.2	88.0	22.1	2.68	5	100	68
EWB-B	78.7	81.0	14.9	1.81	40	100	68
SWB-B	158.9	163.5	30.3	3.68	85	200	68
COG							
Traditional	279.1	296.0	47.3	4.56	54	306	108
Benevolent	64.3	69.0	11.1	1.07	13	72	108
Companionable	66.1	71.0	11.7	1.11	12	72	111
Kindliness	37.9	41.0	7.1	.67	7	42	111
Wrathful	47.2	47.0	14.9	1.45	13	77	106
Deisticness	10.4	8.0	6.3	.61	5	30	106
Omni-ness	21.9	24.0	4.6	.46	4	24	101
Evaluation	26.9	29.0	5.1	.49	5	30	108
Irrelevancy	5.6	4.0	3.9	.37	4	24	109
Eternality	22.8	24.0	3.8	.36	4	24	113

(table continues)

Table 14 continued

Second Session of Test-Retest (continued)							
Test	Mean	Median	SD	SE	Min	Max	N
COG (continued)							
Passive	11.1	11.0	3.6	.35	3	18	107
SDS	49.4	48.0	14.3	1.31	26	96	119

Note: RWB, EWB, SWB = Spiritual Well-Being Scale

RWB-B, EWB-B, SWB-B = Experimental SWB scale

COG = Concept of God scale (subscales listed)

SDS = Spiritual Distress Scale

SPMAT = Estimate of Spiritual Maturity

SPWB = Estimate of Spiritual Well-Being

SE = Standard error of measurement

Correlations between the two measures (COG, SDS) and the two versions of the SWB scale are presented in Table 15. In addition, correlations between the two SWB scales and some of the demographic data are listed. The correlations are from the total sample only. Sample size for the original SWB scale is 107, and for the test version, 90.

Table 15

Correlations Between SWB Scales and Data from Study Two

Test	RWB	EWB	SWB	RWB-B	EWB-B	SWB-B
COG						
Traditional	.78**	.22	.64**	.76**	.32*	.68**
Benevolent	.79**	.36**	.72**	.78**	.27*	.68**
Companionable	.76**	.23	.63**	.71**	.26*	.63**
Kindliness	.72**	.22	.60**	.68**	.30*	.62**
Wrathful	.09	-.26*	-.06	.09	-.02	.05
Deisticness	-.57**	-.40**	-.58**	-.71**	-.31*	-.65**
Omni-ness	.66**	.29*	.59**	.63**	.26	.57**
Evaluation	.75**	.24	.63**	.64**	.26*	.58**
Irrelevancy	-.44**	-.23	-.41**	-.48**	-.36**	-.51**
Eternality	.62**	.12	.48**	.47**	.23	.44**
Passive	.00	-.29*	-.14	.00	-.11	-.05
SDS	-.59**	-.78**	-.78**	-.66**	-.69**	-.78**
SPMAT	.21	.18	.23	.15	.22	.21
SPWB	.42**	.50**	.53**	.47**	.45**	.54**

(table continues)

Table 15 continued

Test	RWB	EWB	SWB	RWB-B	EWB-B	SWB-B
AGE	.15	.10	.15	.27*	.13	.26
INCOME	-.10	-.03	-.08	.11	.07	.11
EDUC	.14	.17	.17	.12	.25	.19
GOD	.43**	.30	.41**	.20	.29	.27
JESUS	.46**	.14	.34*	.39*	.18	.30*
PROF	.44**	.17	.34*	.41**	.33*	.40**
YEARS	.05	.00	.03	.22	.11	.18
RPART	.47**	.17	.37*	.43**	.42**	.46**

* p. < .01 ** p. < .001 (two-tailed)

Note. RWB, EWB, SWB = Spiritual Well-Being Scale (N = 107)

RWB-B, EWB-B, SWB-B = Experimental SWB scale (N = 90)

COG = Concept of God scale (subscales listed)

SDS = Spiritual Distress Scale; SPMAT = Estimate of Spiritual Maturity; SPWB = Estimate of Spiritual Well-Being; EDUC = Level of Education; GOD = Belief in God

JESUS = Belief in Jesus; PROF = Profession of Christian

YEARS = Years a Christian; RPART = Participation in Religious Activities.

Several scales correlate significantly with the SWB fullscale. Pearson product-moment correlations over .40 with two-tailed significance levels greater than .001 include the SWB subscales, SDS (negative), COG subscales Traditional Christian, Benevolent Deity, Companionable, Kindliness, Deisticness (negative), Evaluation, Irrelevancy (negative), and Eternality. Single item measures with moderate positive and significant correlations include Estimation of Spiritual Well-Being and Belief in God.

The test version of the SWB fullscale yields similar results. The exceptions between the two versions were that the test version correlates significantly with Profession of Christianity and Religious Participation and not Belief in God.

The SWB Religious Well-Being subscale (RWB) correlates greater than $r = .40$ ($p > .001$) with SDS, Estimation of Spiritual Well-Being, all COG subscales except Wrathful and Passive, Belief in God, Belief in Jesus, Profession of Christianity, and Religious Participation.

The test RWB subscale (RWB-B) significantly correlates with the same scales as the original RWB except for Belief in God, and Belief in Jesus.

The Existential Well-Being (EWB) subscale had correlations greater than .40 ($p < .001$) with SDS, the COG subscale Deisticness (negative), and the single item measures of Estimation of Spiritual Well-Being. The test EWB subscale (EWB-B) significantly correlates with the same measures except the COG subscale Deisticness and Religious Participation.

Table 15 lists the Pearson product-moment correlations between the original and test SWB scales and the other measures with two-tailed significance levels of $p < .01$ (*) and $p < .01$ (**).

Study Three

Some of the demographic and descriptive data for the third study is archival and has been presented elsewhere (Davis, Longfellow, Moody, & Moynihan, 1987; Huggins, 1988; Wong, 1989). Table 16 lists demographic data for the three samples. Additional descriptive statistics including measures of central tendency and variability, internal consistency alphas, and standard errors of measurement from these samples are presented in subsequent sections.

Table 16

Demographic Data from Study Three Samples

	Davis et al.		Huggins		Wong	
Category	Freq	Pct	Freq	Pct	Freq	Pct
Age						
Under 20	16	5%	1	1%	1	1%
20-29	65	20%	40	14%	20	28%
30-39	122	37%	69	24%	27	38%
40-49	78	24%	49	17%	6	8%
50-59	17	5%	43	15%	7	10%
Over 60	29	9%	78	27%	11	15%
Gender						
Female	174	53%	172	60%	37	51%
Male	154	47%	112	40%	35	49%
Marital Status						
Single	53	16%	32	11%	22	31%
Married	248	75%	196	69%	46	64%
Divorced	17	5%	8	3%	NA	NA
Widowed	6	2%	27	10%	4	6%
Remarried	NA	NA	21	7%	NA	NA

(table continues)

Table 16 continued

	Davis et al.		Huggins		Wong	
Category	Freq	Pct	Freq	Pct	Freq	Pct
Marital Status (continued)						
Live Together	4	1%	NA	NA	NA	NA
Education						
< High School	5	2%	31	11%	NA	NA
High School	75	23%	61	21%	12	17%
Some college	108	33%	89	31%	39	54%
College Grad	90	27%	75	26%	NA	NA
Post College	52	16%	28	10%	21	30%

Descriptive Statistics

The following two tables (Tables 17 and 18) combine the data from all the studies for purposes of comparison. Table 17 presents statistics from the original version of the Spiritual Well-Being Scale and Table 18 the test version. The tables separate the SWB fullscale and RWB and EWB subscales and give a variety of descriptive statistics. Included are the common measures of central tendency and variability: mean, median, mode, range, standard deviation (SD), and interquartile range (IQ). For the interquartile range, the score value at the 25th and 75th percentile is given. The interquartile range is the difference between these two scores. The tables also list the standard error of measurement (SE), skew, and the number (Top) and percentage (Pct) of respondents who received the highest score. Study Two is separated into three groups: (a) the total number completing the scale at least once (labeled "0"), (b) the first session of those completing both sessions (labeled "1"), and (c) the second session for those competing both sessions (labeled "2").

Table 17

Descriptive Statistics for the Original SWB Scale

Religious Well-Being Subscale							
Study	N	Mean	Median	Mode	Min	Max	Range
One	72	54.2	55.5	60	40	60	20
Two (0)	107	50.8	55.0	60	10	60	50
Two (1)	53	52.0	56.0	60	10	60	50
Two (2)	53	53.4	57.0	60	28	60	32
Davis	330	53.6	55.0	60	32	60	28
Huggins	285	54.8	58.0	60	34	60	26
Wong	72	54.3	56.0	60	39	60	21

Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	5.9	.69	- .77	49	60	11	22	31%
Two (0)	10.1	.98	-1.50	45	59	14	20	10%
Two (1)	9.6	1.32	-2.00	47	59	12	11	21%
Two (2)	7.8	1.08	-1.32	47	60	13	14	26%
Davis	6.2	.34	-1.06	50	59	9	69	21%
Huggins	6.1	.36	-1.26	51	60	9	90	32%
Wong	5.7	.67	- .86	51	59	8	16	23%

(table continues)

Table 17 continued

Existential Well-Being Subscale							
Study	N	Mean	Median	Mode	Min	Max	Range
One	72	51.1	52	47	36	60	24
Two (0)	107	49.0	50	46	25	60	35
Two (1)	53	51.1	51	49	10	60	50
Two (2)	53	52.2	53	55	40	60	20
Davis	330	49.4	51	59	26	60	34
Huggins	285	51.2	53	60	17	60	43
Wong	72	50.9	52	57	40	60	20

Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	6.6	.77	-.56	47	57	10	6	8%
Two (0)	7.4	.71	-.67	44	55	11	4	2%
Two (1)	5.3	.73	-.19	47	56	9	2	2%
Two (2)	5.9	.81	-.65	49	57	8	4	3%
Davis	7.4	.41	-.62	45	55	10	17	5%
Huggins	7.3	.43	-.85	46	57	11	41	14%
Wong	6.2	.74	-.37	45	57	12	3	4%

(table continues)

Table 17 continued

Spiritual Well-Being Fullscale							
Study	N	Mean	Median	Mode	Min	Max	Range
One	72	105.3	108	114	77	120	43
Two (0)	107	99.8	103	115	53	120	67
Two (1)	53	103.2	106	115	65	120	55
Two (2)	53	105.7	110	115	84	120	36
Davis	330	103.0	105	115	68	120	52
Huggins	285	105.9	110	120	51	120	69
Wong	72	105.2	107	112	79	120	41

Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	11.3	1.34	- .63	96	115	19	3	4%
Two (0)	15.1	1.5	- .79	90	113	23	3	2%
Two (1)	12.5	1.71	- .66	94	115	21	2	4%
Two (2)	10.9	1.50	- .61	96	115	19	2	4%
Davis	12.3	.68	- .66	94	113	19	13	4%
Huggins	12.6	.75	-1.03	99	116	17	32	11%
Wong	10.8	1.28	- .56	96	115	19	2	3%

(table continues)

Table 17 continued

Note. Study Two (0) = those taking test only once

Two (1) = first session for those completing both

Two (2) = second session for those completing both

SE = Standard error of measurement

25 = 25th percentile score

75 = 75th percentile score

IQ = Interquartile range (range between 25th and 75th percentile).

Top = number of respondents receiving top score

Pct = percentage of respondents receiving top score.

Table 18

Descriptive Statistics for the Test SWB Scale

Religious Well-Being Subscale								
Study	N	Mean	Median	Mode	Min	Max	Range	
One	72	88.6	91.0	100	54	100	46	
Two (0)	90	79.1	84.0	100	6	100	94	
Two (1)	68	81.0	85.2	100	6	100	94	
Two (2)	68	80.2	88.0	100	5	100	95	
Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	10.5	1.23	-1.48	86	96	10	6	8%
Two (0)	21.7	2.29	-1.47	68	97	29	8	9%
Two (1)	20.4	2.48	-1.88	73	98	25	6	9%
Two (2)	22.1	2.68	-1.75	73	96	23	6	9%

(table continues)

Table 18 continued

Existential Well-Being Subscale								
Study	N	Mean	Median	Mode	Min	Max	Range	
One	72	83.5	84.8	94	44	100	55	
Two (0)	90	77.9	80.5	84	31	100	69	
Two (1)	68	77.8	80.5	84	31	100	69	
Two (2)	68	78.7	81.0	81	40	100	60	
Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	12.3	1.45	- .97	75	94	19	1	1%
Two (0)	14.1	1.48	-1.02	72	86	14	3	3%
Two (1)	14.0	1.69	-1.13	73	86	13	2	3%
Two (2)	14.9	1.81	- .82	69	90	21	3	3%

(table continues)

Table 18 continued

Spiritual Well-Being Fullscale								
Study	N	Mean	Median	Mode	Min	Max	Range	
One	72	172.1	176.3	185	97	199	101	
Two (0)	90	156.9	161.3	180	65	200	135	
Two (1)	68	158.9	161.8	180	85	200	115	
Two (2)	68	158.9	163.5	176	85	200	115	
Study	SD	SE	Skew	25	75	IQ	Top	Pct
One	20.2	2.38	-1.18	160	187	27	0	0%
Two (0)	30.4	3.20	- .87	139	180	41	3	3%
Two (1)	28.6	3.47	- .89	146	180	34	2	3%
Two (2)	30.3	3.68	- .78	140	183	43	2	3%

(table continues)

Table 18 continued

Note. Study Two (0) = those taking test only once

Two (1) = first session for those completing both

Two (2) = second session for those completing both

SE = Standard error of measurement

25 = 25th percentile score

75 = 75th percentile score

IQ = Interquartile range (range between 25th and 75th percentile).

Top = number of respondents receiving top score

Pct = percentage of respondents receiving top score.

Reliability

Intratest

Intratest correlations are another indicator of internal consistency. Table 19 presents intratest correlations from the studies for both versions of the SWB scale. The correlations are Pearson product-moment with two-tailed significance levels reported for $p < .01$ (*) and $< .001$ (**).

The correlations between the original SWB fullscale and RWB subscale range from $r = .85$ to .92. Fullscale-EWB subscale correlations range from $r = .69$ to .94. The correlations between the two subscales range from $r = .26$ to .73.

The test version has similar intratest correlations as the original. SWB-B fullscale to RWB-B subscale correlations range from $r = .87$ to .91 while SWB-B fullscale EWB-B subscale correlations range from $r = .72$ to .90. RWB-B subscale correlations with the EWB-B subscale range from .31 to .57.

Table 19

Intratest Correlations for the SWB Scale

Study	N	RWB-SWB	EWB-SWB	RWB-EWB
One	72	.90**	.92**	.66**
Two (0)	107	.90**	.81**	.48**
Two (1)	53	.92**	.69**	.35*
Two (2)	53	.85**	.72**	.26
Davis	330	.88**	.92**	.63**
Huggins	285	.91**	.94**	.73**
Wong	70	.89**	.92**	.65**
		Test SWB Scale		
One	72	.87**	.90**	.57**
Two (0)	90	.91**	.76**	.41**
Two (1)	68	.89**	.74**	.36*
Two (2)	68	.88**	.72**	.31*

* p. < .01 ** p. < .001 two-tailed

Note. Two (0) = those taking test only once; Two (1) = first session; Two (2) = second session.

Test-Retest

Table 20 presents test-retest correlations from the second study. The Pearson product-moment correlation with two-tailed significance levels at $p < .01$ and $< .001$ are reported. Reliability coefficients for the SWB fullscale and both subscales are given for the original and test version. In addition, correlations are given for different subgroups including gender (male, female), marital status (single, first marriage), and the two samples (community college, church). The church sample is broken down into students (high school and college classes) and adults.

As a reference point, the table lists test-retest coefficients from the original study by Paloutzian and Ellison. No significance levels were reported for the original correlations. The time span for the original study was one week, and for this study, six weeks.

Test-retest coefficients for the SWB fullscale range from .75 to .90. Coefficients for the RWB subscale range from .78 to 94, and for the EWB subscale from .61 to .84.

Table 20

Test-Retest Reliability Coefficients for the SWB Scale

	N	RWB	EWB	SWB
Total Sample	53	.88**	.73**	.82**
Females	32	.88**	.61**	.82**
Males	21	.90**	.84**	.90**
Singles	22	.83**	.70**	.75**
First Marriage	22	.94**	.61*	.89**
Community College	21	.88**	.66**	.78**
Baptist Students	13	.76*	.73*	.78**
Baptist Adults	19	.81**	.72**	.81**
Ellison (1983)	122	.96	.87	.93
		Test SWB Scale		
Total Sample	68	.84**	.73**	.78**

* p. < .01 ** p. < .001

Note. Time span approximately six weeks. For Ellison study time span one week (no significance levels).

Internal Consistency

Internal consistency alphas provide evidence for the homogeneity of a trait or construct. The Cronbach alpha is based on the average correlation of items within a test (Norusis, 1988). Table 21 lists Cronbach alphas for both versions of the SWB scale and its subscales from all the studies. Also included are the original alphas reported by Paloutzian (1982).

Alphas for the original SWB fullscale range from .86 to 91. The RWB subscale alphas range from .85 to .94, and EWB subscale alphas from .86 to .91. For the test version of the SWB scale, fullscale alphas range from .89 to .91, RWB subscale alphas from .83 to .94, and EWB subscale alphas from .81 to .84.

Table 21

Internal Consistency Alphas for the SWB Scale

Study	N	RWB	EWB	SWB
One	71	.86	.85	.91
Two (0	100	.94	.83	.92
Two (2)	50	.89	.76	.86
Davis	305	.82	.84	.89
Huggins	285	.86	.86	.91
Wong	70	.85	.81	.89
Paloutzian (1982)	100	.87	.78	.89
	Test Version			
One	72	.83	.84	.89
Two (0)	85	.92	.81	.90
Two (2)	67	.94	.84	.91

Note. Two (0) = Subjects completing first session.

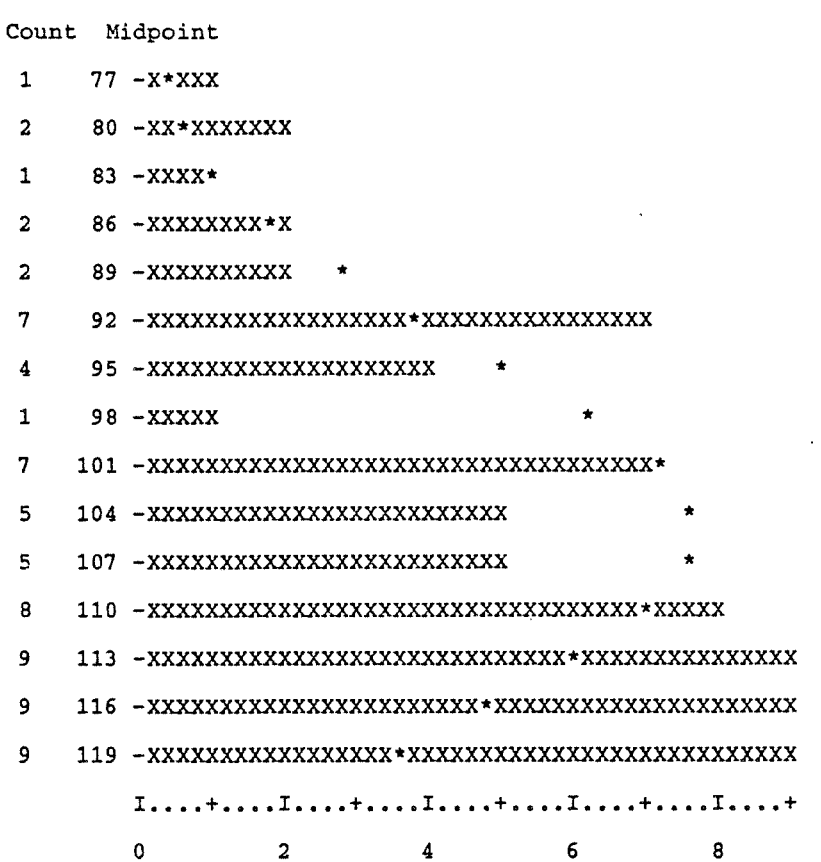
Two (2) = Subjects completing second session.

Response Measurement

One of the purposes of this study was to try to find a solution to the ceiling problems of the SWB scale. Towards that goal an alternative response scale was constructed and tested. The previous tables have already presented many statistics for both versions of the scale. These tables presented descriptive data including measures of central tendency and variability along with reliability coefficients and correlations with other measures.

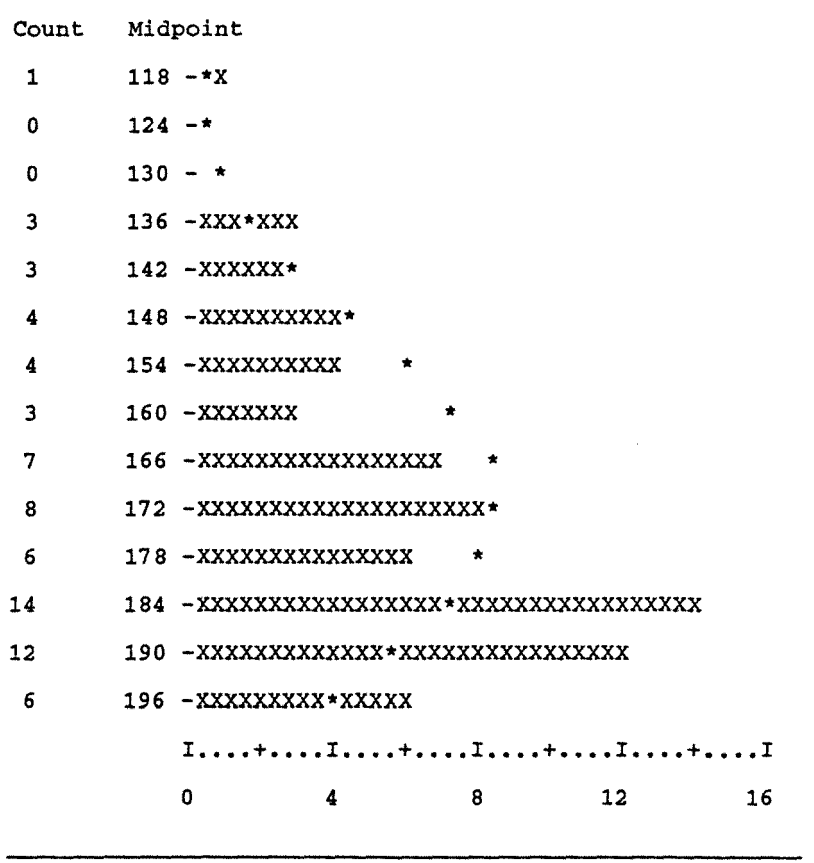
Study one used both of the scales together. The Pearson product-moment correlations between the original and test SWB scales are .84 ($p < .001$) for the fullscale, .84 ($p < .001$) for the RWB subscale, and .82 ($p < .001$) for the EWB subscale.

Figures 1-4 give a visual picture of the distribution of scores for both scales from the two studies in which they were used. The figures are a histogram of the score distribution. Different interval levels are used for visual presentation so the figures are not raw data distributions. The levels combined a range of scores, for example raw scores 95, 96, and 97 together. Superimposed on each figure is a normal curve for comparison purposes.



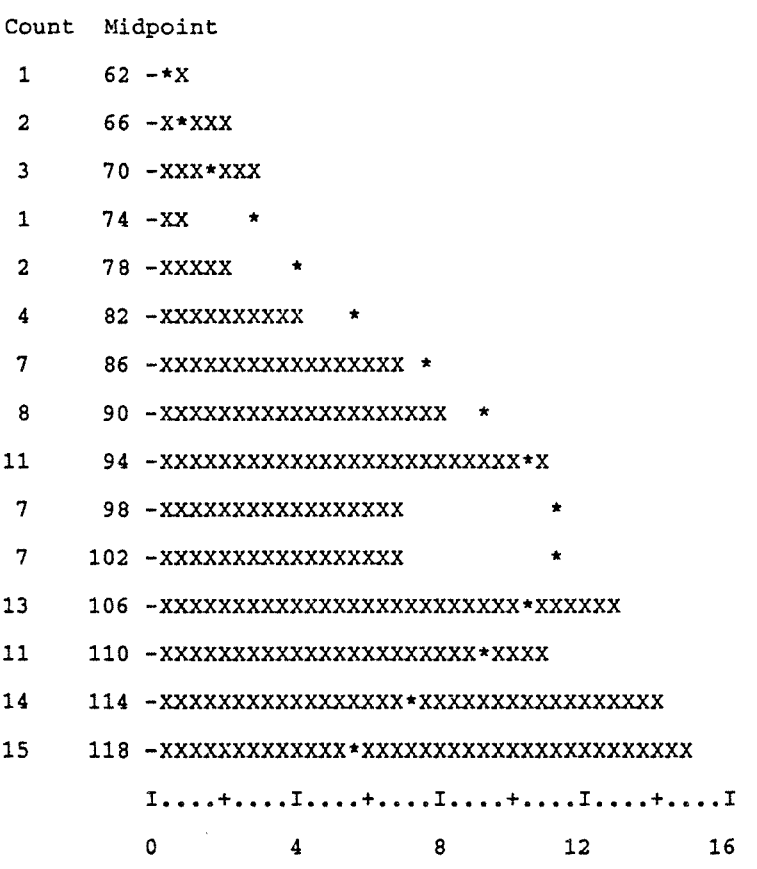
Note. N = 72. Normal curve superimposed. Score interval is 3.

Figure 1. Score Distribution for Original SWB Scale--Study One



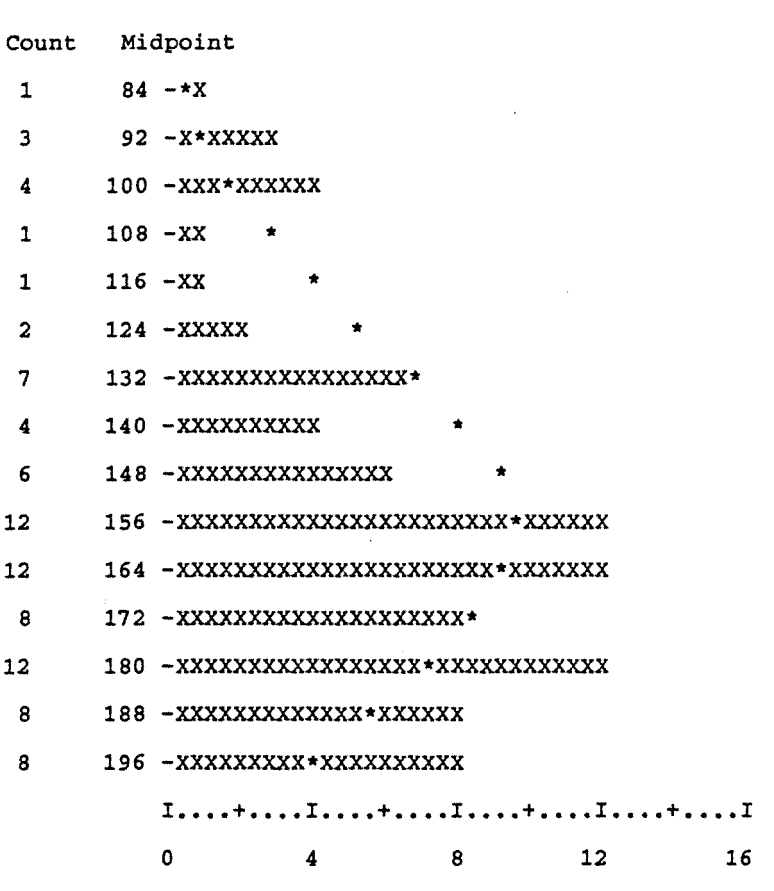
Note. N = 72. Normal curve superimposed. Score interval is 6.

Figure 2. Score Distribution for Test SWB Scale--Study One



Note. N = 107. Normal curve superimposed. Score interval is 4.

Figure 3. Score Distribution for Original SWB Scale--Study Two



Note. N = 90. Normal curve superimposed. Score interval is 8.

Figure 4. Score Distribution for Test SWB Scale--Study Two

CHAPTER 4

DISCUSSION

The purpose of this study was to conduct an evaluation of the Spiritual Well-Being Scale and contribute original research on its psychometric properties. Three research questions were asked:

1. Can a system of evaluation be devised that presents the research conducted with the Spiritual Well-Being Scale in a manner that is understandable and allows for comparison with some standard?

2. Can additional reliability coefficients be generated that are consistent with the original studies and defensible by professional standards?

3. Can the rating scale of the SWB scale be modified to minimize ceiling effects and produce scores approximating a normal distribution?

The first question was answered affirmatively in the first chapter. Using the Standards for Educational and Psychological Testing (1985) as the criteria, five technical areas of the SWB scale were examined: validity, reliability, test development, scaling and

norming, and publication. Most of the previous studies provided evidence for the construct validity of the scale; specifically, correlations with other measures. The SWB scale correlates positively with a variety of measures of religiosity, mental health, and physical well-being as it theoretically should.

As a result of the evaluation several "psychometric needs" became evident to this researcher. For instance, the scale needed more research in the area of reliability. There was only one test-retest study with a time span of one week and one set of internal alphas reported.

Another problem area was the response distribution. Populations that are highly religious tend to score at or near the top of the scale. While in theory this should occur, the scale ceiling is too low and does not provide adequate discrimination among individuals in these populations. Other deficient areas include factor analysis to understand what construct the scale is measuring, experimental manipulation, norming, and publication of a manual.

In conducting the evaluation, it was not possible to include all the previous research with the SWB scale. Research from other institutions has been

harder to locate. In addition, research at Western Conservative Baptist Seminary with the SWB scale continues. There presently are at least a half dozen studies in various stages not included.

As evidenced by the research questions, this study chose to contribute data on reliability and response measurement for the scale. After reviewing the literature in these two areas, three separate studies were done.

The first study introduced a test version of the SWB scale that had a different response scale, a continuous percentage scale from 1 to 100. This scale was tried on a religious sample along with the original SWB scale and three other measures of religiosity. The test scale correlations were in the same range as the original scale with the other measures. In addition, no one in that sample received the top score on the test SWB scale. As a result, further research on the test scale was warranted.

In regard to this study, the use of volunteers always raises the question of generalization. For the original SWB scale, the means and standard deviations were similar to studies by Huggins (1988) and Davis, et al. (1987). Since the results obtained were similar to

other studies with similar samples, some degree of confidence can be placed in the data in representing the relationships between the scales for these populations.

Using both versions of the SWB scale may have influenced the results because of practice effects. Having to respond to the same items twice could have affected the second taking of the scale. To minimize this, the order of presentation was mixed. Some received the original scale first and test scale second, and others the test scale first and the original scale last. By doing this (and by ordering the survey so all the other tests were between), the practice effects should have canceled each other out.

One improvement with this particular study would have been to include the demographic data the first time. Although that data will be gathered at the second session, there will be some who dropout and that information will be missing.

The second study examined test-retest reliability. Two samples, one consisting of community college students and the other attenders of a Baptist church, agreed to participate in a six week study. Volunteers were randomly assigned to complete either the original

SWB scale or the test scale along with two other religious measures. Six weeks later they completed the same three scales. The validity and reliability coefficients were consistent with previous studies as were measures of central tendency and variability.

In doing a test-retest study with a long time span, there were limitations on the type of subjects available. In order to control some of the possible extraneous variance, the study used groups that met together regularly. This enabled the same facilities to be used, time limitations imposed, and standard instructions to be given. However, this prevented random sampling. Volunteers were needed who were available.

This study sampled from two populations, evangelical Christians, like so many studies had used, and non-religious persons. A non-religious sample was sought in order to spread out the distribution and reduce the homogeneity of the sample (Anastasi, 1988).

The Christian sample targeted the population the scale is most often used with to see if the test scale could make the discriminations the six-point Likert scale could not. With this goal in mind, several

institutions were contacted and those which agreed to participate were used.

It can be argued that the data produced in this study only represents the people who participated. However, since the correlations and other results are in the same range as other studies it can also be argued the results are generalizable. Each study contributes another piece of evidence. There are many studies that have produced similar results despite using non-random samples (Davis et al., 1987; Flesner, 1981; Lewis, 1986).

To improve this study, it would have been better to give both demographic questionnaires the first session instead of giving one the first session and the other the second session. Those who completed both sessions had the opportunity to answer all the questions but data was missing on the respondents who only took it once.

The third study used archival data to contribute additional data on the original SWB scale. Descriptive statistics included standard errors of measurement, internal alphas, and other measures of central tendency and variability.

The three samples were chosen because of their size and because item level data was available for secondary analysis. The Wong (1989) study contributed statistics from an ethnic group (Chinese) that was not Caucasian. All three of the studies used religious samples, comprised mainly of middle class, evangelical Christians. One limitation of this study was the unavailability of data from a non-religious population, Christians from different socio-economic backgrounds, and non-evangelical faiths. These populations would provide information on how representative the results are.

Results from the three studies are reported in the previous chapter. The section on descriptive statistics places the samples together and compares them on several levels. For the original fullscale the means range from 99.8 to 105.9, with standard deviations ranging from 15.1 to 10.8. The medians for the fullscale are all higher than the means and the modes were the highest of all. Fifty percent of the scores are within twenty points of each other.

The measures of central tendency and variability highlight the raw score distribution problems with the scale. Means are less than two standard deviations

from the maximum, scores cluster together, and the score distributions are negatively skewed.

Some of this score distribution can be explained from the samples chosen. Religious individuals are a homogeneous group and theoretically should score higher on a scale measuring religious constructs. However, the second study included a non-religious sample and results still revealed a clustering of scores at the top. This may have been due to the community college population. There were a large number of them who described themselves as religious. This, added to the Baptist sample, may explain the distribution.

Another possible explanation may be the scale itself measures a general religious factor for Americans. Since other samples of non-religious populations (Frantz, 1985; Mullins, 1986; Palmer, 1985; Sherman, 1987) yield lower SWB fullscale scores, this is not as likely.

Construct Validity

Although the main purpose of this study was not to look at evidence for validity, this study did provide some. The Spiritual Well-Being Scale correlates

positively and significantly with several religious scales and single item measures. In study one, the original version of the SWB scale correlates positively ($p < .001$) with the Spiritual Maturity Index ($r = .71$), and negatively with the Religious Orientation Intrinsic scale ($r = -.51$). These correlations are similar to other studies (Bufford, 1984; Carr, 1986; Colwell, 1986; Davis et al. 1987; Quinn, 1984) which also studied religious populations.

In the two studies using the Concept of God scale, the SWB fullscale correlates moderately and significantly ($p < .001$) with the COG subscales Traditional Christian ($r = .46$ and $.64$), Benevolent Deity ($r = .53$ and $.72$), Companionable ($r = .45$ and $.63$), Kindliness ($r = .45$ and $.60$), Evaluation ($r = .45$ and $.45$), Deisticness ($r = -.48$), Omni-ness ($r = .59$), Eternality ($r = .48$) and Irrelevancy ($r = -.41$). COG subscales with low or slight correlations with the SWB scale include Wrathful, and Potently Passive.

The COG scale has not been studied very much and caution must be used in interpreting the results. However, what is clear is a relationship between scores on the SWB scale and what one believes about God.

There is a high negative correlation between the SWB fullscale and the Spiritual Distress Scale (r = -.78 p < .001). In the original study with the SDS, the correlations between the two scales were r = -.45 and -.90. Since that was a one week test-retest study, the difference did not make sense. More stable correlations emerged from this study. The second administration correlation between the two scales was r = -.77 (p < .001).

Although the SDS is a relatively new scale, it shows potential as another useful measurement tool. In a previous study (Brinkman, Capes, Kunkel, & Tackett, 1988) the SDS did not suffer from the score distribution problems the SWB scale did with religious populations. Both the SDS and SWB scale purportedly measure in the same domain. Therefore, more research should be done with this scale because it may be a better measure of the spiritual well-being construct than the current version of the SWB scale.

The SWB scale did not significantly (p < .01) correlate with age, income, education, or years as a Christian. That the scale did not correlate highly with age and education was consistent with Diener's (1984) review of subjective well-being studies.

However, income was a predictor of well-being in his review. It may be possible that spiritual well-being is not influenced by income. There may be a distinction between this area of subjective well-being in contrast to other areas, for example, emotional. That the SWB scale did not correlate with years as a Christian is consistent with other studies (Davis et al., 1987) and with Ellison's conceptualization of SWB as distinct from spiritual maturity.

Also contributing evidence for construct validity are the correlations of the SWB scale with single item measures estimating one's spiritual maturity and well-being. In the second study, the SWB fullscale correlates with well-being (r = .53, p < .001) but not maturity (r = .23, p > .01). In study one, estimation of one's spiritual maturity correlates moderately and positively (r = .42 p < .001), with the SWB scale. The study one result may have been due to the highly religious nature of the population and their willingness to respond to questions about their spiritual life. They may have been feeling good about their religious life at the time. The study two participants did not know in advance about the nature of the testing.

The ability of the respondents in study two to distinguish between spiritual maturity and spiritual well-being suggests there is a difference between these constructs. It also suggests that the single item measure of SWB may be useful as an indicator of well-being when the SWB scale cannot be used.

Reliability

As documented in the first chapter, the original reliabilities for the SWB scale were encouraging, but not enough so to satisfy professional standards for psychological tests. The data from this study adds to the knowledge base by providing a six-week test-retest reliability study, additional internal alphas, additional intratest correlations, and standard errors of measurement.

Test-retest reliabilities range from $\underline{r}$ = .73 (EWB), $\underline{r}$ = .82 (SWB) to $\underline{r}$ = .88 (RWB), all significant at the $\underline{p} < .001$ level. The correlations from the six-week study are lower but comparable to the original one-week test-retest study.

As discussed earlier, Anastasi (1988) believes reliability estimates should be in the .80's or .90's,

while Nunnally (1978) considers coefficients above .70 to be respectable. Since the SWB scale is measuring attitudes which are not as stable as skills, lower reliability coefficients can be tolerated. These coefficients are acceptable for most purposes in research but not in a clinical or other setting where decisions are made that could influence someone's future (Kaplan & Saccuzzo, 1982).

Intratest correlations are very high between RWB-SWB and EWB-SWB, ranging from .85 to .94. RWB-EWB correlations ranged from .26 to .73. The RWB-EWB correlations have generally been the lowest of the intratest correlations (see Table 6).

The RWB-EWB relationship has some possible explanations. One is that subscale to fullscale correlations are partly artificial because the subscale makes up part of the fullscale. For that reason, only the RWB-EWB relationship should be considered because it doesn't suffer from the problem of having the same items on both sides of the correlation.

Another reason for the lower correlation is that religious well-being is theoretically different from existential well-being. It should be possible to feel good about one's relationship with God and be

dissatisfied with life. The biblical character Job may be an example of this.

In highly religious samples, the EWB scale generally shows more variation than RWB, a finding that is reversed for non-religious samples. This suggests a relationship with God contributes to more stable RWB scores.

Cronbach's internal alphas for the RWB scale range from .82 to .94, for EWB .76 to .86, and for SWB .86 to .92. These numbers are very consistent with the original alphas reported for the scale: .87 (RWB), .78 (EWB), .89 (SWB). The internal coefficients suggest the SWB scale and subscale items are measuring a homogeneous construct, as hypothesized.

No standard errors of measure (SE) were originally reported so there is no reference point for them. The SE ranges from .68 to 1.71 for the samples on the SWB fullscale. The sample size was too small to report separate standard errors for the mean and standard deviation or for different demographic variables. The SE is useful mainly for individual interpretation, something the scale is not currently being used for (Nunnally, 1978). Since the SWB scale is a research

instrument, this may be the reason these statistics have not been reported before.

Standard errors of measurement seem to be frequently overlooked when discussing a scale's reliability. When the SWB scale develops to the point of being used for decision making purposes, this information will be more valuable.

The second research question asked, "can additional reliability coefficients be generated that are consistent with the original studies and defensible by professional standards?" The results and discussion indicate the reliability of the SWB scale is adequate for its present use as a research instrument. There needs to be more evidence accumulated before using the scale for individual evaluation.

Response Measurement

There are still problems in the area of response measurement. Changing the response scale did not adequately solve the problem of a low ceiling nor did it change the basic shape of the frequency distribution. The test scale demonstrates similar validity coefficients to the original scale with

religious scales and single item measures. This was somewhat surprising because introducing more variability should have influenced the correlations (D. Mueller, 1986).

The measures of reliability are also similar. The test-retest reliability for the test SWB fullscale was $r = .78$ ($p < .001$). Internal alphas for the test fullscale range from .89 to .91.

The test scale, which provided more opportunity for variability, did reduce the number of people achieving the maximum score by about half. However, for the RWB subscale, the mode was the maximum score, as it was for the original scale. In addition, the shape of the score distribution (Figures 1-4) was not altered very much by the different scale. The mean still stayed within two standard deviations from the top. Changing the response scale to a continuous percentage scale does not seem to be the answer to this problem.

The third research question asked, "can the rating scale of the SWB scale be modified to minimize ceiling effects and produce scores approximating a normal distribution?" This study did not answer that question as the response scale tested was unable to achieve this

goal. The score distributions did not change nor did the difference between the correlations with other measures. No advantage of using the continuous scale was evident. A different type of response scale may be able to do what a continuous numerical scale could not. Alternatively, the problem may be in the item statements and not in the response scale.

Contributions

The Spiritual Well-Being Scale has been a very popular research tool since its introduction in the late 1970's. It has been used in over 50 studies (Moody, 1988). The present study sought to step back from using the tool as a research instrument and integrate previous research on the scale, comparing it with standards published by various professional organizations for tests. This evaluation identified several needs, two of which were addressed, reliability and response measurement.

This study has made several contributions. The first is the identification and systematic organization of past research efforts under the framework of professional standards. This simplifies the evaluation

process and eliminates the need to repeat it for each new study. As more research is done, it can easily be added to the data base. If there are mistakes or omissions, they can easily be corrected. The framework is there to use and build on.

The second contribution is the additional data on the SWB scale's reliability. More support was provided for the scale in this area by examining a longer test-retest time span, internal alphas, and standard errors of measurement. Intratest correlations were available and this study collected them together for comparisons. As a result, the reliability of the SWB scale has moved from a psychometric area of weakness to one in which there is more confidence with the scale. Reliability cannot be set aside, though, because more cross-validation studies should be done.

A third contribution from this study is the information on response measurement. So far, two studies have altered the response scale--the present one and Meyers (1986)--without solving the low ceiling problem. The next logical area to address is attempting to reword or substitute item statements that will produce more variable responses. The Standards state that when a test is shown to need revising that

it should be done. It is becoming increasingly clear that the SWB scale is not able to discriminate very well among highly religious persons and needs revision. Discrimination among religious individuals was one of the proposed uses for the scale (Ellison 1982a) and the scale is often used in research with these populations. At this writing, exploratory studies are being done in this area.

A final contribution is the religious demographic questions. These items were compiled in response to dissatisfaction with classifying people solely by their religious denomination. It is hoped these questions can be studied further to see if they can make the discriminations desired.

Suggestions for Future Research

Currently, the two most pressing needs with the Spiritual Well-Being Scale are score distribution and construct analysis. Something needs to be done about the scale's limited usefulness with highly religious populations. At this point, rewording or substituting new items seem to be the next option. With revision comes the task of proper scale construction (Nunnally,

1978). This involves generating item statements and testing them, computing item to scale correlations, accepting and rejecting items, and doing factor analysis.

Another need is trying to decide what the SWB scale is actually measuring. Recent studies are questioning the religious and existential well-being construct theory of the scale. The SWB scale is measuring something well but the construct it represents could be Gorsuch's (1984) general religious factor, or spiritual maturity, or even subjective well-being. Moberg (1984) has identified seven aspects of spiritual well-being: Christian faith, self-satisfaction, personal piety, subjective SWB, optimism, religious cynicism, and elitism. Ellison (1983) highlights two aspects, religious well-being and existential well-being, while Flesner (1981) adds spiritual distress as another dimension. These all need more study to see how they relate to each other and developed into a model of spiritual well-being.

Research in this area should include a non-religious sample to better deal with the problem of the homogeneity of the sample. Also different

populations other than middle-class evangelicals should be studied.

Once these two areas are addressed, research with the scale can continue in developing norms, publishing a manual, and making the scale available for general use.

Overall, the Spiritual Well-Being Scale has been the subject of much research and scrutiny. It has shown great promise as a reliable measure in the domain of religion. The criticisms leveled at the scale should not be cause to reject the scale but to focus efforts at improving it and to make the scale a stronger instrument.

References

Adams, E. W., Fagot, R. F., & Robinson, R. E. (1965). A theory of appropriate statistics. Psychometrika, 30, 99-127.

Agnor, D. W. (1986). Christian and non-religious sociopaths compared: Self-concept, locus of control, guilt, and quality of religious experience (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0423.

Aiken, L. R. (1979). Psychological testing and assessment (3rd ed.). Boston: Allyn and Bacon.

Allen, M. J., & Yen, W. M. (1979). Introduction to Measurement Theory. Monterey, CA: Brooks/Cole.

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. Journal of Psychology, 22, 9-39.

Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. Journal of Personality and Social Psychology, 5, 432-443.

Anastasi, A. (1988). Psychological testing (6th ed.). New York: Macmillan.

Bonner, C. M. (1988). Utilization of spiritual resources by patients experiencing a recent cancer diagnosis. Unpublished master's thesis, University of Pittsburgh.

Bressem, M. R. (1986). The relationship between individual differences in imaginal ability, Christian imaginal frequency, and Christian spirituality (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0412.

Bressem, M. R., Colwell, J., Mueller, E., Neder, J. R., & Powers, H. (1985). Validity of the SMI. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Bressem, M. R., Waller, D. C., Powers, H. (1985). The verbalizer-visualizer cognitive styles and spiritual well-being of church attenders. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Brinkman, D. D. (1989). [Longitudinal study on spiritual growth indicators]. Unpublished raw data.

Brinkman, D. D., Capes, R. L., Kunkel, E. D., & Tackett, C. W. (1988). Item content and variance of the Spiritual Well-Being Scale. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Bufford, R. K. (1984). Empirical correlates of Spiritual Well-being and Spiritual Maturity scales. Paper presented at Christian Association for Psychological Studies annual meeting, Dallas, TX.

Bufford, R. K. (1987, June). The relationship between spiritual wellbeing and physical health. Paper presented at the Christian Association for Psychological Studies, Western Regional Meeting, Seattle, WA.

Bufford, R. K., Bentley, R. H., Newenhouse, J. M., & Papania, A. J. (1986). The relationship among groups using the Spiritual Well-Being Scale. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Bufford, R. K., & Parker, T. G. (1985, August). Religion and wellbeing: Concurrent validation of the Spiritual Well-Being Scale. Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Campbell, A. (1981). The sense of well-being in America: Recent patterns and trends. New York: McGraw-Hill.

Campbell, C. D. (1983). Coping with hemodialysis: Cognitive appraisals, coping behaviors, spiritual well-being, assertiveness, and family adaptability and cohesion as correlates of adjustment (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1983). Theological Research Exchange Network, 002-0199.

Campbell, D. T., & Fiske, D. W. (1967). Convergent and discriminant validation by the multitrait-multimethod matrix. In M. Fishbein (Ed.), Attitude theory and measurement, (pp. 282-289). New York: John Wiley & Sons.

Campise, R. L., Ellison, C. W., & Kinsman, R. (1979, September). Spiritual well-being: Some exploratory relationships. In R. F. Paloutzian (Chair) Spiritual well-being, loneliness, and perceived quality of life. Symposium presented at the annual meeting of the American Psychological Association, New York.

Carpenter, G., & Dean, T. (1985). The relationship between life-span development and spiritual well-being in "Christian" women. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Carr, S. A. (1986). A theoretical and empirical examination of the construct validity of the Spiritual Leadership Qualities Inventory (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0428.

Carson, V., Soeken, K. L., & Grimm, P. M. (1988). Hope and its relationship to spiritual well-being. Journal of Psychology and Theology, 16, 159-167.

Clarke, D. E. (1987). Predictors of spiritual well-being among full-time Youth for Christ/USA staff members (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04710.

Clarke, D., Clifton, C., Cooper, R., Mishler, T., Olson, M., Sampson, J., & Sherman, D. (1985). The effect of social desirability on the Spiritual Maturity Index. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Clarke, D., Clifton, C., Cooper, R., Mueller, E., Sampson, J., & Sherman, D. (1985). A comparison of the 20 and 30 item Spiritual Maturity Index. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Colwell, J. C. (1987). A correlational study of self-concept and spirituality in seminarians (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04711.

Cooper, R. L. (1987). An empirical examination of the construct validity of the Spiritual Maturity Index (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04712.

Cronbach, L. J. (1970). Essentials of Psychological Testing (3rd ed.). New York: Harper & Row.

Davis, W. W., Longfellow, D., Moody, A., & Moynihan, W. (1987). Spiritual Maturity Index: Construct validation. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Diener, E. (1984). Subjective well-being. Psychological Bulletin, 95, 542-575.

Dillman, D. A. (1978). Mail and telephone surveys: The total design method. New York: John Wiley.

Durham, M. L. (1984). Denominational differences in attribution to supernatural causation. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Durham, M. L. (1986). Denominational differences in supernatural locus of control and spiritual well-being (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1985). Theological Research Exchange Network, 002-0422.

Edwards, A. L., & Kenney, K. C. (1946). A comparison of the Thurstone and Likert techniques of attitude scale construction. Journal of Applied Psychology, 30, 72-83.

Ellison, C. W. (1982a, January). Spiritual well-being. Paper presented at the Medical Conference on Spiritual Disease, Duke University, Durham, NC.

Ellison, C. W. (1982b). Spiritual well-being. Paper presented at Western Conservative Baptist Seminary Winter Lectureship, Portland, OR.

Ellison, C. W. (1983). Spiritual well-being: conceptualization and measurement. Journal of Psychology and Theology, 11, 330-340.

Ellison, C. W., & Cole, K. C. (1982) Religious commitment, television viewing, values and quality of life. Journal of Psychology and Christianity, 1, 21-32.

Ellison, C. W., & Economos, T. (1981, April). Religious orientation and quality of life. Paper presented at the annual meeting of the Christian Association for Psychological Studies, San Diego.

Ellison, C. W., & Paloutzian, R. F. (1978). Assessing quality of life: Spiritual well-being. Paper presented at the annual meeting of the American Psychological Association, Toronto.

Ellison, C. W., & Paloutzian, R. F. (1979, September). Religious experience and quality of life. In R. F. Paloutzian (Chair) Spiritual well-being, loneliness, and perceived quality of life. Symposium presented at the annual meeting of the American Psychological Association, New York.

Ellison, C. W., Rashid, I., Patla, P., Calica, R., & Haberman, D. (1984). Personality, religious orientation, and spiritual well-being. Unpublished manuscript.

Feagin, J. R. (1964). Prejudice and religious types: A focused study of southern fundamentalists. Journal for the Scientific Study of Religion, 4, 3-13.

Fishbein, M., & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavioral criteria. Psychological Review, 81, 59-74.

Flesner, R. S. (1981). Development of a measure to assess spiritual distress in the responsive adult. Unpublished master's thesis, Marquette University, Milwaukee, WI.

Frantz, J. L. (1985). MMPI and DSM-III diagnosis related to selected measures of religious and demographic variables in adult outpatients (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1985). Theological Research Exchange Network, 002-0398.

Garner, W. R. (1960). Rating scales, discriminability, and information transmission. The Psychological Review, 67, 343-352.

Glock, C., & Stark, R. (1966). Christian belief and anti-semitism. New York: Harper and Row.

Gorsuch, R. L. (1968). The conceptualization of God as seen in adjective ratings. Journal for the Scientific Study of Religion, 7, 56-64.

Gorsuch, R. L. (1984). Measurement: The boon and bane of investigating religion. American Psychologist, 39, 228-236.

Gravetter, F. J., & Wallnau, L. B. (1988). Statistics for the behavioral sciences (2nd ed.). St. Paul: West Publishing.

Guilford, J. P. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill.

Guttman, L. (1944). A basis for scaling qualitative data. American Sociological Review, 9, 139-150.

Hawkins, D. B. (1986). Interpersonal behavior traits, spiritual well-being, and their relationship to blood pressure (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0413.

Hawkins, D. B., & Larson, R. (1984). The relationship between measures of health and spiritual well-being. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Henerson, M. E., Morris, L. L., & Fitz-Gibbon, C. T. (1978). How to measure attitudes. Beverly Hills: Sage.

Huggins, S. M. (1988). The effect of small group attendance, personal devotions, and church attendance on spiritual well-being (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1988). University Microfilms International, (in press).

Hunt, R. A., & King, M. B. (1971). The intrinsic-extrinsic concept: A review and evaluation. Journal for the Scientific Study of Religion, 10, 339-356.

Jang, S. T. C. (1987). The effects of acculturation and age on spiritual well-being among ethnic Chinese church-goers (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04714.

Jang, S., Paddon, T., & Palmer, W. (1985). Locus of control in relation to spiritual well-being and spiritual maturity. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Kaplan, R. M., & Saccuzzo, D. P. (1982). Psychological testing: Principles, applications, and issues. Monterey, CA: Brooks/Cole.

Labovitz, S. (1970). The assignment of numbers to rank order categories. American Sociological Review, 35, 515-524.

Lewis, G. G. (1986). The correlations among ambivalence, one's concept of God, and spiritual well-being as measured on two diverse religious groups (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0420.

Likert, R. (1967). The method of constructing an attitude scale. In M. Fishbein (Ed.), Attitude theory and measurement, (pp. 90-95). New York: John Wiley & Sons.

Marto, R. P. (1984). A father's locus of control, spiritual well-being, and self-esteem and their relationship to his child's self-esteem in a Catholic parochial high school population (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1983). Theological Research Exchange Network, 002-0103.

Mashburn, D. J. (1987). Effects of sex-role combinations in couples on marital satisfaction and spiritual well-being (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04717.

McClain, E. W. (1978). Personality differences between intrinsically religious and nonreligious students: A factor analytic study. Journal of Personality Assessment, 42, 159-166.

McKelvie, S. J. (1978). Graphic rating scales - How many categories? British Journal of Psychology, 69, 185- 202.

Messick, S. (1975). The standard problem: Meaning and values in measurement and evaluation. American Psychologist, 30, 955-966.

Meyers, R. W. (1986). The effect of two likert labeling formats on Spiritual Well-being scale scores. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Minium, E. W. (1978). Statistical reasoning in psychology and education. New York: John Wiley.

Mitchell, E. A. (1984). The relationship of spiritual well-being and mood state during pregnancy: An analysis of the decision-making process when accepting or rejecting the fetus (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1984). Theological Research Exchange Network, 002-0363.

Mitchell, E. A., & Reed, S. K. (1983). A study on the interaction between spiritual well-being, self-concept and social desirability. Unpublished manuscript, Western Conservative Baptist Seminary, Portland, OR.

Moberg, D. O. (1971). Spiritual well-being: Background and issues. Washington, DC: White House Conference on Aging.

Moberg, D. O. (1979). Spiritual well-being: Sociological perspectives. Washington, DC: University Press of America.

Moberg, D. O. (1984). Subjective measures of spiritual well-being. Review of Religious Research, 25, 351-364.

Moberg, D. O. (1986). Spirituality and science: The progress, problems, and promise of scientific research on spiritual well-being. Journal of the American Scientific Affiliation, 38, 186-194.

Moody, A. V. (1988). The effect of deliberate faking good and faking bad on Spiritual Well-Being Scale scores in a church sample (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1988). University Microfilms International, (in press).

Mueller, D. J. (1986). Measuring social attitudes: A handbook for researchers and practitioners. New York: Teachers College Press.

Mueller, E. E. (1986). The relationship between religious beliefs/attitudes and psychopathology in an evangelical seminary sample (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0429.

Mullins, W. H. (1986). Chronic pain: A study of treatment outcome as it relates to coping behaviors, assertiveness, spiritual well-being, and MMPI scores (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). Theological Research Exchange Network, 002-0406.

Newenhouse, J. M. (1988). Maternal spiritual well-being and the social adaptational status of first grade children (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1987). University Microfilms International, 88-01200.

National Interfaith Coalition on Aging. (1975). Spiritual well-being: A definition. Athens, GA: NICA.

Norusis, M. J. (1986). SPSS/PC+ for the IBM PC/XT/AT. Chicago: SPSS Incorporated.

Norusis, M. J. (1988). SPSS/PC+ advanced statistics V2.0 for the IBM PC/XT/AT and PS/2. Chicago: SPSS Incorporated.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Palmer, W. C. (1985). Generalized hope, expectancies, locus of control, and spiritual wellbeing in relation to quitting smoking (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1985). Theological Research Exchange Network, 002-0376.

Paloutzian, R. F. (1982, October). Research directions on spiritual well-being. Paper presented at the annual meeting of the Society for the Scientific Study of Religion, Providence, Rhode Island.

Paloutzian, R. F., & Ellison, C. W. (1979a, September). Developing a measure of spiritual well-being. In R. F. Paloutzian (Chair) Spiritual well-being, loneliness, and perceived quality of life. Symposium presented at the annual meeting of the American Psychological Association, New York.

Paloutzian, R. F., & Ellison, C. W. (1979b). Loneliness and spiritual well-being as functions of living environment and professional status in adult women. Paper presented at the annual meeting of the Western Psychological Association, San Diego.

Paloutzian, R. F., & Ellison, C. W. (1979c). Religious commitment, loneliness, and quality of life. Paper presented at the annual meeting of the Christian Association for Psychological Studies, Minneapolis.

Paloutzian, R. F., & Ellison, C. W. (1979d). Loneliness, spiritual well-being and quality of life. In L. A. Peplau and D. Perlman (Eds.), Loneliness: A sourcebook of current theory, research and therapy, (1982, pp. 224-237). New York: Wiley Interscience.

Papania, A. J. (1988). The effect of religious orientation, history of sexual trauma, and typology on spiritual well-being and interpersonal behavior among adult male child molesters (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1987). University Microfilms International, 88-01201.

Parker, T. G., Jr. (1984). An empirical examination of the construct validity of the Spiritual Leadership Qualifications Inventory (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1984). Theological Research Exchange Network, 002-0353.

Poppleton, P. K., & Pilkington, G. W. (1963). A comparison of four methods of scoring an attitude scale in relation to its reliability and validity. British Journal of Social and Clinical Psychology, 2, 36-39.

Quinn, J. B. (1984). The relationships between religiosity and marital satisfaction: Correlations among the Religious Orientation Scale, the Spiritual Well-Being Scale and the Marital Satisfaction Scale (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1984. Theological Research Exchange Network, 002-0354.

Reckase, M. D. (1984). Scaling Techniques. In G. Goldstein & M. Hersen (Eds.), Handbook of Psychological Assessment (pp. 38-53). New York: Pergamon.

Robinson, J. & Shaver, P. (1973). Measures of Social Psychological Attitudes. Ann Arbor: University of Michigan.

Rodriguez, K. D. (1988). Predictors of self-esteem and spiritual well-being among sexually abused women (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1988). University Microfilms International, 88-11007.

Roth, P. D. (1988). Spiritual well-being and marital adjustment. Journal of Psychology and Theology, 16, 153-158.

Sherman, D. B. (1987). A comparison of interpersonal behavior traits and spiritual well-being among eating-disordered patients and medical outpatients (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04722.

Spilka, B., Armatas, P., & Nussbaum, J. (1964). The concept of God: A factor-analytic approach. Review of Religious Research, 6, 28-36.

Spilka, B., Hood, R. W. Jr., & Gorsuch, R. L. (1985). The psychology of religion: An empirical approach. Englewood Cliffs, NJ: Prentice-Hall.

Standards for educational and psychological tests and manuals. (1966). Washington, DC: American Psychological Association.

Standards for educational and psychological testing. (1985). Washington, DC: American Psychological Association.

Temple, D. B. (1987). Psychological well-being, life events, and religiosity (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1986). University Microfilms International, 87-04723.

Thurstone, L. L. (1928). Attitudes can be measured. American Journal of Sociology, 33, 529-554.

Upshaw, F. G. (1984). The effects of a communication skills training program on marital satisfaction, commitment, social desirability, and spiritual well-being (Doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR, 1984). Theological Research Exchange Network, 002-0366.

Wong, E. (1989). Correlations between SWB, self-esteem, and social desirability for Chinese Americans in the Northwest. Unpublished doctoral dissertation, Western Conservative Baptist Seminary, Portland, OR.

Appendix A

Spiritual Well-Being Scale (Original Version)

Appendix B

Spiritual Well-Being Scale (Test Version)

Appendix C

Concept of God Scale

Appendix D

Spiritual Distress Scale

Appendix E

Spiritual Maturity Index

Appendix F

Religious Orientation Scale

Appendix G

Single Item Measures (Study One)

For each of the following circle the number which best describes you:

1. How important are your religious beliefs and practices?

No importance; have no religion	1 2 3 4 5 6 7 .	Extremely important; religious faith is the center of my life

2. How would you describe your current religious knowledge and development?

Limited; need help and instruction from others	1 2 3 4 5 6 7	Extensive; able to help and instruct others

3. To what degree are you satisfied with your life at the present time?

Not at all satisfied	1 2 3 4 5 6 7	Completely satisfied

4. How would you evaluate your own spiritual maturity?

Very Immature	1 2 3 4 5 6 7	Very Mature

5. How many hours per week have you spent (averaged over the last year) in some form of ministry or service (for example, teaching, serving on elder or deacon boards, visitation, personal evangelism, counseling, discipleship, child care, preparation of food, providing practical help to others, and the like)?

____ HOURS PER WEEK

Appendix H

Demographic Questions (Study Two)

Q1 What is your present age? _____ YEARS Today's date: _____

Q2 Your gender: (circle number of your answer)

1 FEMALE
2 MALE

Q3 Your present marital status: (circle number)

1 SINGLE (NEVER MARRIED)
2 FIRST MARRIAGE
3 SEPARATED OR DIVORCED
4 REMARRIED
5 LIVING TOGETHER
6 OTHER (PLEASE SPECIFY) _____

Q4 What was your approximate total family income from all sources, before taxes, in 1987? (circle number)

1 LESS THAN $10,000
2 $10,001 TO $20,000
3 $20,001 TO $30,000
4 $30,001 TO $40,000
5 $40,001 TO $50,000
6 OVER $50,000

Q5 What is the highest level of education that you have completed? (circle number)

1 DID NOT FINISH HIGH SCHOOL
2 COMPLETED HIGH SCHOOL (OR G.E.D.)
3 ATTENDED OR COMPLETED TRADE OR BUSINESS SCHOOL
4 SOME COLLEGE
5 COMPLETED COLLEGE
6 SOME GRADUATE WORK
7 A GRADUATE DEGREE

Q6 Which of the following best describes your racial or ethnic identification? (circle number)

1 BLACK (NEGRO)
2 CHICANO (MEXICAN AMERICAN)
3 NATIVE AMERICAN (AMERICAN INDIAN)
4 ORIENTAL
5 WHITE (CAUCASIAN)
6 OTHER (PLEASE SPECIFY) _____

Q7 Which religion, or faith, do you most closely identify with? (circle number)

1 CATHOLIC
2 JEWISH
3 PROTESTANT (PLEASE SPECIFY) _____
4 OTHER (PLEASE SPECIFY) _____
5 I DON'T IDENTIFY WITH ANY ORGANIZED RELIGION

Q8 How would you evaluate your own spiritual maturity? (circle number)

VERY IMMATURE 1 2 3 4 5 6 7 VERY MATURE

Q9 How would you evaluate your own spiritual well-being? (circle number)

VERY LOW 1 2 3 4 5 6 7 VERY HIGH

Today's date:_____

Q1 Which of the following statements comes closest to expressing what you believe about God? (circle the number of the response which best describes your beliefs)

1 I don't believe in God.
2 I don't know whether there is a God and I don't believe there is any way to find out.
3 I don't believe in a personal God, but I do believe in a higher power of some kind.
4 I find myself believing in God some of the time, but not at other times.
5 While I have doubts, I feel that I do believe in God.
6 I know God really exists and I have no doubts about it.
7 None of the above represents what I believe. What I believe about God is _____

Q2 Which of the following statements comes closest to expressing what you believe about Jesus? (circle one number)

1 Frankly, I'm not entirely sure there ever was such a person as Jesus.
2 I think that Jesus was only a man although an extraordinary one.
3 I believe that Jesus was a great man and very holy, but I don't see Him as the Son of God any more than all of us are children of God.
4 While I have some doubts, I basically believe that Jesus is divine.
5 Jesus is the Divine Son of God and I have no doubts about it.
6 None of the above represents what I believe. What I believe about Jesus is _____

Q3 Do you claim to be a Christian? (circle one number)

1 NO
2 YES, I respect and attempt to follow the moral and ethical teachings of Christ.
3 YES, I have received Jesus Christ into my life as my personal savior and Lord.
4 YES, I have received Jesus Christ as my personal Savior and Lord and I seek to follow the moral and ethical teachings of Christ.

Q4 If you answered YES to the above question (Q3), how many years have you been a Christian?

_____ YEARS

Q5 Which of the following statements comes closest to expressing what you believe about the Bible as the basis for your religious faith and belief? (circle one number)

1 Every person has the ability to determine what is true and I don't need the Bible for this.
2 The Bible is God's word and is the ultimate source of truth for me.
3 In addition to the Bible, religious experiences (e.g., speaking in tongues) are just as important.
4 In addition to the Bible, decisions by the church hierarchy (such as the Pope) are another source.
5 In addition to the Bible, writings or sayings by others are equally valid.
6 I'm not sure how to answer this.
7 None of the above state what I believe. What I believe about the Bible is _____

Q6 How often do you participate in a religious activity of any type? (circle one number)

1 LESS THAN ONCE A YEAR
2 ONCE OR TWICE A YEAR
3 3 TO 11 TIMES A YEAR
4 1 TO 3 TIMES A MONTH
5 WEEKLY
6 MORE THAN ONCE A WEEK

Appendix I

Research Announcement (Study One)

MEASURE SPIRITUAL GROWTH: CAN IT BE DONE?

Dan Brinkman, a doctoral student at Western Seminary, is currently studying methods of measuring spiritual growth and needs volunteers to help him. If you would like to participate in this study which would involve filling out some questionnaires now and again a year from now please indicate by signing your name below and turning this form in. You will receive the results back from these questionnaires next year to compare them with the ones you filled out earlier as a way of looking at your own growth. All information will be mailed to you and will be kept in strict confidence.

Yes, I would like to participate:

Name:
Address:
City:
State, Zip:

Appendix J

Instructions (Study One)

April, 1988

Greetings!

Thank you for your willingness to participate in this study of spiritual growth. As you will recall, this is a two part study which will involve filling out and returning the enclosed forms now and again a year from now. The surveys being used are all in various stages of research and development and any comments or suggestions you wish to make would be appreciated and can be written on the forms or on a separate sheet.

When filling out the surveys remember that there are no right or wrong answers. Each survey will be used only in comparison with the one that is filled out next year by you, not in comparison with anyone else. Please answer each question so that it reflects you as you are now, not as you would like to be. The number on the front of the forms is a number which has been assigned to you and will be used to match surveys. When collecting and analyzing the data your name will not be on any of the paperwork. This is done to help maintain confidentiality.

After the surveys from next year are turned in they will be returned back to you so that you can see the difference a year has made. Please fill out the survey now and mail it back in the enclosed self-addressed, stamped envelope today. Thank you again for your cooperation and help.

Sincerely,

Dan Brinkman

Appendix K

Research Announcements (Study Two)

Research Proposal

Research Description

There are many tests today which are supposed to measure different aspects of a person's faith or religion. However, the majority of these tests were not developed with regard to accepted standards for test construction. One of the areas that is often neglected is called reliability, or how consistently a test measures something. One of the ways to investigate this property is to have people take a test and then have the same people take the same test at a later time and compare both scores. The present research, which is part of my doctoral dissertation at Western Conservative Baptist Seminary, is examining this. The tests being used ask about one's beliefs and attitudes about God and life.

Research Proposal

A large sample size is needed for this study. What is proposed is that the members of the Sunday school classes at your church be asked to voluntarily participate in this study. The extent of their involvement would be to fill out 3 short tests and a background questionnaire during the class and then to complete the same tests again six weeks later. It takes about 20 minutes to finish. If some people missed class the second time, they would have an opportunity to complete the tests at a later date. All of the tests would be kept in strictest confidence and individual names would not be associated with the tests when scored. Names would be associated with the tests initially in order to match up the first and second sessions but removed after they're finished.

Benefits for the Church

The individual who completes the tests will be given an opportunity to receive feedback on what the tests are measuring and what their scores mean.

The pastoral staff will be given a summary of the group results and an opportunity to discuss with me what they mean.

What is needed from the Church

In order to accomplish this task, the permission and support of the pastoral staff and Sunday school teachers are needed as well as some class time during the Sunday school hour to complete the tests. More specific details will be worked out with you. I would like to administer the first packet as soon as it can be arranged, with the second administration taking place six weeks later.

Examples of the tests are included with this proposal. If there are any questions I can be contacted at the Seminary through Box 158 or at home at (206) 892-0530.

Sincerely,

Dan Brinkman

RESEARCH OPPORTUNITY

Yes, this is a genuine opportunity to accomplish several things at once. First, to expand our scientific knowledge base; second, learn something more about yourself; third, help out a needy graduate student; and fourth, earn extra credit for your psychology course. "How can this be done ?" you might ask. Well, Dan Brinkman, a doctoral student is conducting research for his dissertation. What this involves is having individuals like yourself complete some tests now and again about six weeks from now. It would take approximately 20 minutes of your time to complete.

There will be four opportunities to come by and participate in the project. It will be on campus in room 2550. The times are:

Wednesday, October 19, 1988 at 12:00 noon Room 2550
Thursday, October 20, 1988 at 12:00 noon Room 2550
Wednesday, October 26, 1988 at 12:00 noon Room 2550
Thursday, October 27, 1988 at 12:00 noon Room 2550

After completing the second session, your names will be submitted to your professor for extra credit. You will also have an opportunity to receive feedback on the results of the study.

Sincerely,

Dan Brinkman

RESEARCH PARTICIPATION REMINDER

For those of you who have already participated in the first phase of the research project during one of the lunch hours on October 19, 20, 26, or 27th this is a reminder about the second session. (The project involved the completion of some surveys concerning your attitudes and beliefs).

There will be three opportunities to come by and complete your part in the project. It will be in the same room as before. The times are:

Wednesday, November 30, 1988 at 12:00 noon Room 2550
Thursday, December 1, 1988 at 12:00 noon Room 2550
Wednesday, December 7, 1988 at 12:00 noon Room 2550

After this second session, your names will be submitted to your professors for extra credit. You will also receive a handout that explains the purpose of the study when you are done. In addition, there will be an opportunity to sign up to discuss the results of your testing either in person or through a phone conversation. In person feedback times will be on the Wednesday afternoons listed above starting at 1:00.

Sincerely,

Dan Brinkman

Appendix L

Instructions (Study Two)

Greetings,

Allow me to introduce myself and explain why I am here. My name is Dan Brinkman. I am a student at WCBS--I've been there for 5 years.

Anyway, one of the things we do as part of our studies at the seminary is to look at religious belief and behavior. We do this to try and understand it better in the hopes that this will help us do the work that we have been called to do in a more effective way.

The leadership here has graciously agreed to let us ask you to help us out in one our studies, as many other churches have already done.

The particular study that I am coordinating consists of two parts. The first part will be done right now and involves completing some forms that ask about the beliefs and attitudes you have about God. The second part will be done here about six weeks from now and will involve completing some other similar forms at that time. All information that you give will be kept strictly confidential. The church leaders here will be given a summary of the results of the whole group but will not know how individuals answered. It is hoped that the information you as a church supply will enable your leaders to serve you better. When I come back next time I will tell you more about what we are studying and will also give each of you a chance to learn about how you did.

Are there any questions?

PASS OUT PACKETS

Inside each packet is a four page handout and an index card. Please print your name on the index card. The only reason for doing this is to be able to match up these forms with the ones you will hopefully complete in a few weeks.

Now complete all four pages. The last page asks for information that helps us understand the variety in responses. Remember there are no right or wrong answers. Please answer each question as it describes you now, not as you would like to be or think you should be.

When you are finished, put the forms back in the packet along with the index card and hand it back in. Please don't seal the envelope.

Note: if a question is not understood, leave it blank.

Appendix M

Handout (Study Two)

INFORMATION CONCERNING THE RESEARCH PROJECT

Thank you for participating in this research project. Your willingness to help was much needed and appreciated. Let me tell you a little more about the project itself. What you have completed is called a test-retest reliability study. Reliability is one way in which tests and scales are evaluated and is concerned with how consistent responses are on a test. The specific way consistency was examined in this study was by having you complete some scales once and then completing the exact same scales a few weeks later. The responses that were given each time are compared to each other to see how similar they are.

The three scales that were used in the study are all in the research and development stage and are not available for general use. The first one you completed is called the Spiritual Well-Being Scale (SWB). It is designed to measure two areas of an individual's well-being: in relation to God, and in relation to life purpose and life satisfaction. The SWB Scale yields three scores: 1) a religious well-being (RWB) score; 2) an existential well-being (EWB) score; and a fullscale score (SWB) which is the sum of the first two scores. For this study there were two different ways of measuring your responses. Half of you received a 6 point Likert scale, the other half a 0-100 continuous scale. On the Likert scale RWB and EWB scores ranged from 10-60, and SWB from 20-120 with higher scores indicating higher well-being. On the continuous scale scores ranged from 0-100 and 0-200, respectively. For more information about this instrument see "Spiritual Well-Being: Conceptualization and Measurement" by C. W. Ellison in The Journal of Psychology and Theology, Volume 11, No. 3, pp. 330-340, 1983.

The second scale is called the Concept of God Scale (COG). It consists of 75 adjectives that the individual is asked to rate whether each one is like God or not like God. The adjectives are then combined into 11 different subscales, e.g. Traditional Christian, Kindliness, Wrathfulness, and Evaluation. For more information about this scale see "The Conceptualization of God as Seen in Adjective Ratings" by R. L. Gorsuch in the Journal For The Scientific Study of Religion, Volume 7 pp. 56-64, 1968.

The third scale is called the Spiritual Distress Scale (SDS). This scale attempts to measure distress of the human spirit by looking at one's relationship to God and how that influences feelings of forgiveness, love, hope, trust, and meaning in life. It yields one score. The higher the score, the more distress an individual is reporting. Scores can range from 22-132. For more information about this scale see "Development of a Measure to Assess Spiritual Distress in the Responsive Adult" by R. Flesner. This is an unpublished master's thesis from Marquette University, 1981.

A more complete description of this study and the results should be forthcoming in my dissertation which is scheduled to be finished in early 1989. It is titled "An Evaluation of the Spiritual Well-Being Scale: Reliability and Response Measurement" by Daniel D. Brinkman, Western Conservative Baptist Seminary, Portland, OR. It should be accessible through the Seminary library or through University Microfilms.

If you have any questions or comments please feel free to contact me through the seminary in Portland at (503) 233-8561 ext. 395, or at home in Vancouver at (206) 892-0530. Thank you again.

-Dan Brinkman

Appendix N

Raw Data

Data Interpretation Key for Study One

Column

A Identification Number
B Religious Well-Being subscale (Test Version)
C Existential Well-Being subscale (Test Version)
D Spiritual Well-Being fullscale (Test Version)
E Religious Well-Being subscale (Original Version)
F Existential Well-Being subscale (Original Version)
G Spiritual Well-Being fullscale (Original Version)
H Spiritual Maturity Index
I Religious Orientation Scale - Extrinsic
J Religious Orientation Scale - Intrinsic
K COG scale - Traditional Christian subscale
L COG scale - Benevolent Deity subscale
M COG scale - Companionable subscale
N COG scale - Kindliness subscale
O COG scale - Wrathfulness subscale
P COG scale - Deisticness subscale
Q COG scale - Omni-ness subscale
R COG scale - Evaluation subscale
S COG scale - Irrelevancy subscale
T COG scale - Eternality subscale
U COG scale - Potently Passive subscale
V Importance of Religious Beliefs
W Extent of Religious Knowledge
X Life Satisfaction
Y Estimation of Spiritual Maturity
Z Hours per Week in Ministry

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
101	634	746	1380	46	47	93	128	19	15	280	63	35	64	30	10	24	29	4	24	7	5	6	5	4	3
102	865	780	1645	53	47	100	140	26	14	275	62	34	59	54	11	24	29	4	24	9	6	5	5	5	10
103	900	940	1840	56	55	111	145	12	12	291	62	37	69	45	16	24	26	4	24	13	6	5	6	5	2
104	930	940	1870	58	60	118	145	14	12	302	69	40	71	59	8	24	30	4	24	6	7	6	6	6	2.5
105	998	932	1930	60	56	116	171	19	11	302	72	42	72	50	5	24	27	4	24	17	7	7	6	6	
106	915	930	1845	55	55	110	138	19	12	292	71	36	69	36	6	24	28	4	24	10	6	5	5.5	5	3
107	915	940	1855	60	57	117	166	17	9	306	72	42	72	43	5	24	30	4	24	13	7	5	6	4	15
108	1000	960	1960	60	57	117	162	14	9	304	72	42	72	44	5	24	30	4	24	18	7	5	5	5	10
109	905	930	1835	55	56	111	137	19	11	295	72	40	72	34	8	24	30	4	24	11	5	5	6	5	
110	930	920	1850	57	52	109	131	21	15												6	4	5	5	
111	895	840	1535	54	38	92	141	19	19	302	61	42	72	48	16	24	30	4	24	9	6	4	4	4	1
112	700	710	1410	40	41	81	104	24	23	288	71	40	70	46	6	14	30	7	24	11	5	2	2	2	1
113	970	930	1900	56	58	114	142	16	13	300	72	40	72	38	5	24	30	4	24	11	7	5	6	5	5
114	930	925	1855	47	49	96	141	15	12	291	68	39	68	56	9	24	21	4	24	10	7	4	4	5	2
115	1000	940	1940	60	51	111	149	12	14	294	72	40	72	29	5	22	28	4	24	6	7	4	4	4	
116	982	970	1952	60	56	116	166	18	10	294	72	41	72	23	5	24	29	4	24	5	7	6	6	5	20
117	904	980	1884	59	59	118	146	20	14	301	67	37	67	69	10	24	30	4	24	8	6	2	5	3	2
118	847	914	1761	50	58	108	138	24	16	286	62	35	62	57	10	24	27	4	24	10	6	7	5	5	13
119	880	960	1840	51	55	106	126	31	12										24		7	5	6	5	
120	990	985	1975	60	60	120	166	23	10	304	72	42	72	26	5	23	30	4	24	8	7	4	6	5	2
121	848	745	1593	49	44	93	136	28	12	279	60	35	61	44	15	24	28	4	24	10	6	6	4	6	12
122	1000	700	1700	60	47	107	155	19	11	302	72	42	72	54	5	24	30	4	24	10	6	5	5	5	70
123	977	955	1932	60	59	119	133	25	13	298	69	41	72	47	7	24	27	4	24	13	6	5	5	5	4.5
124	1000	960	1960	60	58	118	157	17	15	306	72	42	72	33	5	24	30	4	24	13	7	4	7	4	12
125	977	741	1718	60	54	114	137	12	21	301	72	42	72	43	5	24	30	4	24	13	7	7	5	6	27.5
126	990	850	1840	60	53	113	171	23	9	304	72	42	72	39	5	24	30	4	24	13	7	5	6	5	10
127	910	750	1660	46	39	85	117	24	18	303	62	42	72	32	15	24	29	4	24	10	5	4	4	4	1
128	790	790	1580	43	40	83	136	26	17												5	3	3	4	8
129	960	890	1850	60	54	114	149	24	13	306	70	42	72	71	7	24	30	4	24	13	7	4	4	4	6.5
140	875	860	1735	49	55	104	163	20	15	296	72	40	70	59	5	24	26	4	24	13	7	5	5	5	
145	960	940	1900	60	58	118	157	23	10	306	72	42	72	70	5	24	30	4	24	9	7	5	4	3	2
146	975	920	1895	60	55	115	157	29	15	294	66	39	67	62	8	24	29	4	24	12	7	7	6	6	4
149	890	960	1850	60	54	114	159	18	13	293	71	37	72	47	6	22	24	5	24	9	7	5	5	5	12
201	960	955	1915	58	51	103	139	19	13	302	72	42	72	46	5	24	30	4	24	11	7	5	6	5	2.5
203	536	442	978	40	37	77	100	35	16	273	56	30	59	71	16	24	20	4	24	16	4	5	2	3	.5
204	920	760	1700	58	50	108	150	17	10	298	71	41	71	61	6	24	29	4	24	13	7	5	5	5	2.5
205	895	955	1850	56	47	103	143	15	11	301	72	42	72	23	5	24	30	4	24	8	7	6	5	6	5
206	930	720	1650	55	47	102	149	17	11	304	72	42	72	47	5	24	28	4	24	13	7	4	4	4	3
207	970	915	1885	60	57	117	148	23	13	295	72	42	69	38	5	24	30	4	24	11	7	5	6	5	9
208	880	820	1700	58	52	110	146	23	10	298	71	41	71	61	6	24	29	4	24	12	7	5	5	5	4.5
211	925	760	1685	56	45	101	149	16	10	302	72	42	72	31	5	24	30	4	24	12	7	7	5	5	10
212	943	922	1865	55	50	105	142	19	15	300	72	41	71	50	5	24	27	4	24	8	7	4	5	5	3
213	745	750	1495	50	47	97	114	26	16	275	63	32	67	38	11	24	25	5	24	13	5	6	3	5	1.5
214	900	820	1720	47	42	89	139	17	12	297	72	42	68	45	5	24	30	4	24	8	6	5	5	5	6.5
215	870	690	1560	60	54	114	164	15	10	302	72	42	72	62	5	24	30	4	24	10	7	5	7	5	50
216	970	800	1770	56	45	101	157	14	12	300	72	42	72	18	5	24	30	4	24	3	7	7	6	6	40
217	950	870	1820	59	58	111	141	26	18	296	64	40	71	58	17	24	28	4	22	8	5	4	4	4	0
218	895	760	1655	58	49	101	134	13	12	306	72	42	72	53	5	24	30	4	24	18	7	5	6	5	2
219	730	830	1560	48	53	101	115	22	25	278	62	37	70	45	10	24	28	4	24	12	6	3	4	3	.5
220	947	952	1899	60	60	120	152	16	13	301	72	42	72	43	5	24	30	4	24	18	7	5	5	5	2
228	1000	990	1990	60	59	119	164	12	10	296	72	42	72	38	5	24	30	4	24	13	7	6	6	5	25
223	910	845	1755	58	50	108	157	27	10	301	65	40	72	59	11	24	29	4	24	13	7	7	6	5	2
224	893	600	1493	54	36	90	145	19	13	300	72	40	70	60	5	24	30	4	24	11	6	4	3	4	4
225	900	1000	1900	60	60	120	170	23	9	301	72	42	72	26	10	24	30	4	24	13	7	7	6	5	3
227	950	925	1875	59	55	114	135	24	13	305	71	42	72	51	6	24	30	4	24	12	7	4	5	4	8
228	940	830	1770	55	60	115	156	16	12	301	72	42	72	31	5	24	30	4	24	13	6	6	5	6	6
230	759	730	1489	48	46	94	122	20	15	297	69	39	71	61	8	24	30	4	24	10	5	4	5	3	2
231	1000	920	1920	60	56	116	159	21	11	300	72	42	70	38	5	24	30	4	24	8	7	2	3	4	2
232	815	880	1695	53	57	112	146	18	12	296	72	42	72	18	5	24	30	4	24	8	7	5	6	5	1.5
234	845	840	1685	48	45	93	132	20	12	298	66	41	70	60	14	24	27	4	24	14	7	6	5	5	8
299	602	574	1176	42	39	81	104	33	15	243	48	24	40	72	17	24	22	4	24	14	5	5	3	4	1
301	870	800	1670	51	50	101	152	23	14	258	64	33	63	31	11	24	24	11	22	10	6	5	6	5	8
302	919	610	1529	48	48	96		29	15																
304	867	945	1812	53	60	113	125	22	12	281	65	38	63	35	8	20	26	5	23	13	6	3	6	5	5
305	605	780	1385	43	50	93	125	27	14	218	53	27	49	30	14	14	23	6	22	12	6	3	5	5	5.5
306	774	734	1508	48	46	94	137	18	14	278	66	39	70	39	10	24	25	7	24	10	6	5	5	4	5
307	675	750	1425	43	44	87	116	26	22	300	72	42	72	48	5	24	30	4	24	12	6	3	5	3	3.5
308	955	810	1765	60	57	117	158	26	10	297	72	41	71	33	5	24	29	4	24	11	7	7	7	6	60
310	825	800	1625	53	52	105	126	21	15	298	71	41	68	41	8	24	30	4	24	10	6	5	6	5	3
311	855	485	1340	54	38	92	130	18	15	287	70	42	72	50	8	22	20	4	24	15	6	5	3	4	20
312	710	725	1435	44	45	91	137	19	16	250	63	35	67	27	11	20	21	6	22	9	6	4	5	5	9
313	935	760	1695	58	51	109	135	24	12	285	69	40	70	51	10	24	28	4	24	14	7	7	5	6	9

Raw Data Interpretation Key for Study Two
Original Version Spiritual Well-Being Scale

Column

A Identification Number
  100's = community college
  200's = Baptist high school class
  300's = Baptist college class
  400's = Baptist ladies class
  500's = Baptist young adults class
  600's = Baptist middle age adults class
B Religious Well-Being subscale (First Session)
C Existential Well-Being subscale (First Session)
D Spiritual Well-Being fullscale (First Session)
E Religious Well-Being subscale (Second Session)
F Existential Well-Being subscale (Second Session)
G Spiritual Well-Being fullscale (Second Session)
H Spiritual Distress Scale
I COG scale - Traditional Christian subscale
J COG scale - Kindliness subscale
K COG scale - Companionable subscale
L COG scale - Wrathfulness subscale
M COG scale - Deisticness subscale
N COG scale - Omni-ness subscale
O COG scale - Evaluation subscale
P COG scale - Irrelevancy subscale
Q COG scale - Eternality subscale
R COG scale - Potently Passive subscale
S COG scale - Benevolent Deity subscale
T Age
U Gender 1 = Female 2 = Male
V Marital Status
  1 = Single (Never Married) 2 = First Marriage
  3 = Separated or Divorced 4 = Remarried
  5 = Living Together 6 = Other
W Income
  1 = Less than $10,000 2 = $10-20,000
  3 = $20-30,000 4 = $30-40,000
  5 = $40-50,000 6 = Over $50,000

Raw Data Interpretation Key for Study Two
Original Version Spiritual Well-Being Scale (continued)

X Education
1 = Did not finish high school
2 = Completed high school
3 = Attended or completed trade school
4 = Some college 5 = Completed college
6 = Some graduate work 7 = Graduate degree

Y Ethnic Identification
1 = Black 2 = Chicano 3 = Native American
4 = Oriental 5 = White 6 = Other

Z Religious Identification
1 = Catholic 2 = Jewish 3 = Protestant
4 = Other 5 = No identification

1 Estimation of Spiritual Maturity

2 Estimation of Spiritual Well-Being

3 Belief in God
1 = Don't believe 2 = No way to know
3 = Higher Power 4 = Believe sometimes
5 = Basically believe 6 = No doubts
7 = None of the above

4 Belief in Jesus
1 = Don't believe 2 = Only a man
3 = Not Son of God 4 = Basically believe
5 = No doubts 6 = None of the above

5 Profession of Christianity
1 = No
2 = Moral ethical
3 = Received Christ as Savior
4 = Received Christ and follow Him

6 Years a Christian

7 Belief in Bible
1 = Not needed
2 = Bible ultimate source
3 = Bible plus religious experiences
4 = Bible plus church hierarchy
5 = Bible plus sayings of others
6 = Not sure 7 = None of the above

8 Religious Participation
1 = Less than once/year 2 = 1-2/year
3 = 3-11/year 4 = 1-3/month
5 = Weekly 6 = More than 1/week

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	UVWXYZ123456 78
101	53	56	109	50	55	108	37	264	30	53	26	10	24	24	4	24	9	62	42	131455656619911
103	49	46	98	46	44	90	57	306	42	72	76	30	24	30	24	24	18	52	22	21345257654 411
104	59	52	111				75	304	42	72	73	23	24	28	7	24	16	60	24	11945477
106	59	56	115	56	59	117	40	296	42	72	25	8	22	29	4	24	9	72	18	11245345653 633
107	60	53	113				35	298	42	72	25	5	19	28	4	24	5	72	43	14245456
110	38	48	80	39	49	88	53												36	121751554429953
111	47	49	96	48	44	92	56	269	41	71	29	16		30	4	24	9	64	31	12145535419972
112	60	59	119	60	56	116	41	299	42	72	21	9	24	29	4	24	11	68	37	13145477654 926
113	10	55	65	28	37	65	65	31	7	12	13	5	4	5	9	4	3	32	40	1444555551119971
114	41	49	90	42	58	94	44	229	32	57	19	17	17	16	7	22	7	52	33	14535344569962
117	52	57	109				45	299	40	72	27	6	24	29	4	24	6	72	19	11525450
122	56	49	107	56	58	110	37	296	41	72	29	7	23	29	4	24	11	71	19	21625356454 726
125	48	51	93	56	57	113	44	279	40	71	24	13	22	24	4	24	8	64	19	1534535564441965
128	59	51	110	59	47	106	58	293	42	72	40	10	17	30	4	24	15	67	19	11345455454 826
130	60	56	116	60	51	111	41												18	11245356654 1126
132	32	52	84				51	260	34	60	44	20	19	26	6	24	16	56	18	11425424
142	43	58	98				48			67					9	24			19	11345445
143	44	51	95	44	53	97	44	260	38	69	37	10	17	25	10	24	8	66	19	21545356654 1962
144	59	57	116				40	298	42	72	34	5	24	28	4	24	15	72	22	11145377
145	58	43	98	58	40	98	64	268	36	64	32	17	24	27	4	24	10	59	18	21325335554 1176
148	40	36	76				71	221	30	54	49	16	15	22	7	21	10	56	20	11345445
151	51	52	103	54	55	109	43	287	40	70	38	8	24	28	4	24	10	69	32	12545555458 353
152	44	46	90	45	55	100	53	267	40	64	23	16		21	4	24	11	56	35	1254545654 23561
156	35	50	85	35	53	88	56	284	42	70	35	19	13	29	4	24	8	61	19	11145454419931
157	14	39	53				89	187	27	42	46	27	14	15	9	19	13	37	19	21625544
158	46	44	90	38	46	84	63	246	32	65	24	11	12	21	7	23	3	68	22	11125544419962
159	60	56	116	57	59	116	33	301	42	72	28	5	24	30	4	24	17	72	23	12445425654 5265
160	35	50	85	48	52	100	74	290	42	71	31	20	17	30	7	24	16	57	21	11245433653 724
161	30	50	80				63	183	24	44	44	19	16	15	11	15	12	41	18	21525544
162	45	51	96	56	59	115	49	277	33	64	37	13	23	24	4	24	8	66	18	11425556654 1373
164	44	60	104				46	298	42	72	47	5	24	30	4	24	3	72	28	23125415
200	55	48	100	52	46	98	57	304	40	70	71	8	24	30	4	24	16	69	24	21345311664 876
201	56	50	106	54	53	111	52	304	42	72	46	5	24	28	4	24	11	72	17	11315355654 926
204	56	33	89				66	288	42	70	64	21	16	30	6	22	14	56	14	11995444
203	48	48	84	46	48	88	77	295	40	72	56	15	21	25	4	24	11	62	14	21615411654 926
206	57	47	104	58	58	113	57	304	42	72	64	5	24	30	4	24	8	72	16	21915365654 1126
207	59	43	102	58	51	110	49	298	40	72	43	11	24	23	4	24	13	66	15	21115456654 1026
210	48	46	94	43	48	88	67												15	21985456499944
214	55	52	107	57	51	108													99	21695355554 926
215	40	40	80				62	274	40	70	48	15	22	25	4	22	13	60	14	11995456
216	46	49	97	45	49	94	68	298	42	72	49	8	24	23	4	24	13	70	14	11415444553 1026
217	60	50	110	60	54	114	30	301	42	72	65	5	24	30	4	24	13	72	41	13845344654 1325
218	51	51	102				54	276	34	63	38	12	17	28	4	24	9	67	14	11415455
219	60	53	113	60	53	113	44			72				30		24			14	11415999654 126
226	32	39	71				60	179	25	42	46	18	14	18	14	14	11	42	15	11495962
230	60	59	119	60	55	115	49												15	11113455654 926
231	36	33	69				85												17	11415343542 724
232	38	25	61				02	256	35	56	60	13	13	25	4	24	16	57	17	11115471649996
233	36	38	68				69						4						16	2999999999999
234	36	49	105				58												14	19999945654 823
300	57	51	108				59	273	39	69	31	14	17	25	6	24	9	60	19	21385455
301	55	51	106				51	300	41	70	56	5	24	30	4	24	7	71	24	11245455
308	47	38	85	60	43	103	55	301	42	72	54	12	23	28	8	24	16	66	19	11945356654 1326

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	UVWXYZ123456 78
304	58	57	115	59	56	115	36	304	40	70	67	5	24	30	4	24	8	72	31	2225534565422 6
306	52	44	96				69	290	38	68	35	8	22	26	4	24	4	70	16	1152294465A 965
307	50	44	94				63	282	34	65	37	7	24	28	4	24	7	70	26	1125533565A1426
403	60	59	119	60	55	115	32		42	72		5	24	28	4	24	3	72	30	1294534565A1226
405	60	60	120	60	60	120	29	294	40	72		5	24	28	4	24	11	72	48	12945366465A 926
407	60	60	120				31	298	42	72	27	5	24	30	4	24	7	72	37	1234536664543026
408	48	42	90				52	288	38	70	43	5	24	24	4	24	13	72	43	1254534565A3826
504	50	39	89				79												28	12445434
506	59	53	112	60	57	117	46	303	40	69	50	5	24	30	4	24	6	72	30	1295533565A1426
510	58	46	104	57	40	97	55	297	40	72	44	6	24	30	4	24	12	72	34	2447535565A3025
514	54	47	101	55	60	115	57	301	42	72	66	5	24	30	4	24	16	72	31	1258545565A9926
515	59	59	118	57	59	116	39	300	40	71	53	5	24	28	4	24	9	72	31	2255534465A1726
519	58	57	115	58	60	118	34	306	42	72	48	5	24	30	4	24	15	72	29	1236535565A2126
520	60	60	120	60	60	120	35	305	42	72	64	5	24	30	4	24	16	72	99	2267535565A2826
521	56	52	108	60	52	112	37	306	42	72	58	5	24	30	4	24	13	72	34	2244535665A2526
523	60	55	115	60	57	117	37	301	40	72	44	5	24	30	4	24	5	72	35	2244535765A2026
524	50	37	87				71	306	42	72	71	9	24	30	4	24	12	68	30	2344543
525	59	58	117				57	306	42	72	63	10	24	30	4	24	11	67	33	12445344
526	60	58	118				46	304	42	72	67	8	24	28	5	24	8	69	35	22445358
527	42	43	85				78	204	28	44	52	20	16	20	16	16	12	44	27	22275481
528	55	46	101				41	286	37	72	56	10	24	27	4	24	13	67	29	12235443
529	58	48	106				53	299	42	72	58	5	24	30	4	24	13	72	38	2258535665A1626
530	60	44	104				56	301	42	72	61	5	24	30	4	24	12	72	30	2287534465A2526
531	60	50	110				44	278	36	67	46	12	23	27	4	24	11	65	29	2256535665A1026
532	60	58	118				34	296	42	72	28	5	24	28	4	24	12	72	29	12545349
601	48	47	95				35								4		11		42	12435344
603	49	50	99	53	53	106	47	270	40	69	30	18		27	7	23	7	65	38	2224534565425 5
606	55	53	108	59	52	111	50	294	40	67	43	14	24	29	4	24	13	61	33	1234534465A1926
609	58	49	107	54	49	103	39	299	40	72	59	5	24	30	4	24	6	72	40	2237535665A2226
610	35	43	78				82	272	32	58	61	13	24	25	4	24	11	57	30	12545358
611	59	58	117	60	58	118	39	297	41	72	25	5	22	30	4	24	5	72	38	22245444653 826
612	54	46	100				62	290	35	66	53	7	24	27	4	24	13	68	30	22545359
614	48	43	91	54	47	101		298	36	65	50	7	24	30	6	24	13	70	47	1498535399999999
616	59	55	114				41	291	37	70	59	5	24	27	4	24	11	69	30	12335356
617	60	54	114	59	50	109	51	300	42	72	50	7	24	27	4	24	10	69	35	1244535665A2826
621	60	49	109				58	298	39	70	42	6	24	30	4	24	13	70	41	22355366
623	59	56	115	58	59	114	30	304	42	72	63	5	24	28	4	24	6	72	58	2267534665A4476
624	33	31	64				01	306	42	72	49	5	24	30	4	24	18	72	46	22345344
626	31	43	74				76	196	16	28	69	29	24	11	11	24	15	34	46	24455564
630	58	57	115				30	304	40	72	66	8	24	30	4	24	13	69	32	22545346
631	53	52	105				38	291	39	70	44	9	24	27	4	24	9	68	29	22345345
632	56	54	110	60	56	116	39	299	40	70	48	5	24	30	4	24	18	72	41	1434534665A1225
634	44	47	91	45	43	88	64	281	35	67	53	19	24	24	4	24	9	58	36	2237535565A1726
635	50	48	98				66	301	42	72		5	24	30	4	24	7	72		6543226
636	54	41	95				69								4	24			99	92331349
637	55	39	94				56	296	42	72		8	22	28	4	24	13	70	30	12341444
638	42	38	80				42	287	40	68	46	10	22	27	4	24	10	66	56	1432534465A2525
639	60	46	106				46	293	42	72	21	5	24	28	4	24	11	72	54	1244536665A4126
640	55	58	113				31	303	42	72	37	5	24	29	4	24	11	72	37	2255524565A9925
641	58	55	113				44	304	42	72	50	5	22	30	4	24	8	72	37	1268536665A2525
642	52	48	100				48	299	42	72	26	5	24	30	4	24	13	72	68	2225535565A3626
643	60	59	119					306	42	72	38	5	24	30	4	24	13	72	66	1224535565A3726
644	59	55	114				31	296	37	67	46	9	24	30	4	24	11	72	38	2227535665A1426
645	53	48	98				62	301	42	72	68	5	19	30	4	24	11	72	50	23215343

Raw Data Interpretation Key for Study Two
Test Version Spiritual Well-Being Scale

Column

A Identification Number
  100's = community college
  200's = Baptist high school class
  300's = Baptist college class
  400's = Baptist ladies class
  500's = Baptist young adults class
  600's = Baptist middle age adults class
B Religious Well-Being subscale (First Session)
C Existential Well-Being subscale (First Session)
D Spiritual Well-Being fullscale (First Session)
E Religious Well-Being subscale (Second Session)
F Existential Well-Being subscale (Second Session)
G Spiritual Well-Being fullscale (Second Session)
H Spiritual Distress Scale
I COG scale - Traditional Christian subscale
J COG scale - Kindliness subscale
K COG scale - Companionable subscale
L COG scale - Wrathfulness subscale
M COG scale - Deisticness subscale
N COG scale - Omni-ness subscale
O COG scale - Evaluation subscale
P COG scale - Irrelevancy subscale
Q COG scale - Eternality subscale
R COG scale - Potently Passive subscale
S COG scale - Benevolent Deity subscale
T Age
U Gender 1 = Female 2 = Male
V Marital Status
  1 = Single (Never Married) 2 = First Marriage
  3 = Separated or Divorced 4 = Remarried
  5 = Living Together 6 = Other
W Income
  1 = Less than $10,000 2 = $10-20,000
  3 = $20-30,000 4 = $30-40,000
  5 = $40-50,000 6 = Over $50,000

Raw Data Interpretation Key for Study Two
Test Version Spiritual Well-Being Scale (continued)

X Education
1 = Did not finish high school
2 = Completed high school
3 = Attended or completed trade school
4 = Some college 5 = Completed college
6 = Some graduate work 7 = Graduate degree

Y Ethnic Identification
1 = Black 2 = Chicano 3 = Native American
4 = Oriental 5 = White 6 = Other

Z Religious Identification
1 = Catholic 2 = Jewish 3 = Protestant
4 = Other 5 = No identification

1 Estimation of Spiritual Maturity

2 Estimation of Spiritual Well-Being

3 Belief in God
1 = Don't believe 2 = No way to know
3 = Higher Power 4 = Believe sometimes
5 = Basically believe 6 = No doubts
7 = None of the above

4 Belief in Jesus
1 = Don't believe 2 = Only a man
3 = Not Son of God 4 = Basically believe
5 = No doubts 6 = None of the above

5 Profession of Christianity
1 = No
2 = Moral ethical
3 = Received Christ as Savior
4 = Received Christ and follow Him

6 Years a Christian

7 Belief in Bible
1 = Not needed
2 = Bible ultimate source
3 = Bible plus religious experiences
4 = Bible plus church hierarchy
5 = Bible plus sayings of others
6 = Not sure 7 = None of the above

8 Religious Participation
1 = Less than once/year 2 = 1-2/year
3 = 3-11/year 4 = 1-3/month
5 = Weekly 6 = More than 1/week

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	UVWXYZ123456 78
100	960	690	1650				31		40	67	21	7		85	4	19	12	62	99	14345477
102	720	820	1540	910	930	1840	48	308	42	72	46	13	84	30	4	24	11	69	18	211253456S3 666
103	915	825	1740				50	287	36	69	31	6	24	89	4	24	11	70	19	11345345
104	545	760	1305				53	286	34	64	37	16	19	26	4	23	12	61	36	22445255
109	1000	1000	2000	1000	1000	2000	32												18	211154876S4 126
115	1000	1000	2000				32	306	48	78	33	5	84	30	4	24	13	78	18	11645477
116	60	910	970	80	970	1050		77	10	15	13	11		14	10	4	3	32	18	21425564711991I
118	490	450	940				68	261	35	59	48	25	17	29	10	24	11	50	21	15126544
119	850	580	1430	840	810	1650	73	298	40	70	46	9	84	84	5	24	11	71	19	11585433654108S
120	134	810	944	98	940	1038	57	152	23	40	50	23	16	9	5	16	12	27	18	21349555121991I
121	715	845	1560	905	905	1810	48	284	32	66	25	10	24	29	4	24	11	63	19	21643255654 226
123	820	710	1530	770	850	1620	44	276	42	72	26	8		28	4	24	11	70	34	1284511565462562
124	860	780	1640	880	799	1659	39	246	21	53	33	11	20	84	4	24	6	63	35	12545457853I546
126	854	741	1595	580	777	1357	58	284	36	66	60	22	20	26	4	24	13	60	19	21643645531484
127	995	922	1917	990	900	1890	39	295	42	71	44	10	24	30	4	24	13	67	18	11525467854I436
129	670	310	980	510	400	910	61	189	25	52	48	13	16	18	15	12	14	53	21	21625365454I113
131	610	845	1455				57	226	34	55	37	15	14	23	8	16	8	52	18	11345488
133	515	700	1215				72	237	31	60	36	14	19	27	7	21	9	58	18	11585156
134	145	750	895	50	800	850	82	206	25	47	44	29	15	16	15	21	10	36	20	15145582311997I
135	530	815	1345	480	745	1225	54	284	29	70	26	15	20	27	4	24	10	64	25	15445524541996I
136	130	520	650				87	142	23	49	34	20	16	14	10	22	9	47	23	22245163
137	984	720	1704				47	299	42	72	41	10	24	25	9	24	4	67	22	11385348
138	975	755	1730	948	897	1639	50	298	42	72	38	7	22	29	4	24	16	70	36	9845345565541685
139	770	838	1608	710	820	1530	41	213	36	59	32	22	18	27	9	12	12	40	33	12245564761997I
140	986	840	1826	994	947	1941	50	285	37	66	39	9	24	26	4	24	9	72	18	11485356454I126
141	485	485	850	470	600	1070	65												18	11445163548 3x3
146	950	840	1790	1000	820	1620	42	301	42	72	28	5	24	30	4	24	13	72	32	14285345634997I
147	500	515	1015				86	276	38	72	40	20	19	25	9	23	13	59	18	21585144
149	800	840	1640	630	960	1490	41	245	32	54	48	12	16	22	8	22	7	55	23	15245455543 264
150	940	840	1780	895	760	1655	51	279	39	70	30	6		27	7	24	6	71	19	11445045534I166
153	718	738	1456				68	269	36	63	38	13	22	24	4	24	9	64	19	21345443
154	400	970	1370				60	234	31	60	53	21	16	23	13	20	12	50	31	12445343
155	600	875	1475	760	1000	1760	41	226	31	54	41	16	12	22	6	18	8	56	23	12545877654I333
163	948	852	1800				48	290	38	69	43	6	24	23	9	24	6	71	18	11425366
165	443	730	1813				66	286	33	57	71	26	24	30	5	24	18	51	18	21125464
202	989	930	1919	808	804	1609		289	38	65	59	13	24	25	7	24	13	68	15	21195444754 626
203	750	640	1390	848	680	1525	57		42				24		4	24	10	67	14	11315345454 926
208	765	565	1330	780	780	1560	72	257	37	65	61	16	12	26	8	21	14	57	15	21615433683 963
209	630	710	1340	850	630	1280	78	230	33	55	60	20	18	22	18	18	13	50	14	21615454644 446
211	975	825	1800	580	500	1080		297	38	68	66	13	19	30	5	24	9	64	15	21415366999999
212	660	770	1430	710	790	1500	73	278	40	72	43	11	13	24	4	24	16	69	15	21615444453 636
213	559	440	999	447	442	889	76			58					8				15	21995344854I063
220	990	760	1750	840	480	1320	56	301	42	72	31	5	24	25	4	24	13	72	17	21415444454 126
221	870	830	1700	883	535	1388	66	294	42	72	36	3		28	4	24	6	70	17	11915533654I426
222	995	890	1885	998	908	1900	41	303	42	69	66	18	24	30	7	24	13	65	16	21215444454I426
223	620	480	1100	740	410	1150	64	293	42	72	52	8	24	30	4	24	11	70	17	11115444454 726
224	580	800	1380				63	239	36	59	34	13	14	22	7	21	11	54	14	11995464
225	690	730	1620	930	770	1720	58												14	11991456553 323
227	885	760	1585	570	840	1810	66	294	32	62	63	18	24	29	5	24	13	60	17	11415354754 936
228	639	784	1623	940	748	1688	43	296	40	70	56	5	24	27	4	24	13	72	31	12234356454I426
229	988	638	1777	971	886	1857	41	289	42	71	36	5	24	24	8	24	6	71	14	21615354454 626
303	770	720	1490				67	257	35	66	40	19	18	23	13	22	11	61	19	11145454
305	1000	800	1800	950	650	1600	55	294	41	69	58	6	23	27	4	24	10	71	20	11145364454I586

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	UVWXYZ123456 78
400	880	780	1660	950	660	1610	49	299	41	70	87	5	24	30	.4	24	12	72	39	1295539996542826
401	1000	960	1960	990	940	1930	33	296	42	72	18	5	24	30	4	24	8	72	40	12945366654 826
402	980	970	1950	985	970	1955	39	306	42	72	44	5	24	30	4	24	8	72	44	1256835645543826
404	605	585	1190	1000	820	1680	50	301	42	72	50	13	24	30	9	22	15	65	47	133353446654 225
406	980	980	1960				58	296	42	72	83	5	22	30	4	24	3	72	35	12946377
500	985	880	1605	985	845	1770	57	295	40	71	34	6	22	30	4	24	14	72	31	123583556541626
501	770	535	1305	575	496	1071	56	274	31	60	44	16	24	27	4	24	16	56	39	2445834415482586
502	835	775	1610	810	780	1590	46	300	41	70	70	10	24	30	4	22	16	66	33	1447534455420826
503	995	875	1870	999	974	1973	48	306	42	72	68	5	24	30	4	24	16	72	38	1225534565419826
505	785	797	1582	795	730	1525	49	294	40	68	61	12	24	30	4	24	9	66	32	2234535555541926
507	840	700	1540				61	297	39	71	44	6	24	28	8	24	4	71	32	22965335
508	810	835	1645				61	303	42	72	59	5	24	28	4	24	12	72	31	12445335
509	908	892	1800	885	840	1725	42	301	42	72	48	5	24	30	4	24	8	72	36	2245535556542026
511	990	800	1790	950	810	1760	53	300	41	71	61	10	24	28	4	24	17	69	28	1254536765422826
512	935	810	1745				48	301	39	70	56	5	24	30	4	24	11	72	37	12445356
513	970	880	1850				49	308	42	72	68	9	24	30	4	24	13	68	29	224535365
516	1000	980	1980	990	960	1950	38	301	42	72	43	5	24	30	4	24	13	72	40	2267535556542826
517	1000	1000	2000	1000	1000	2000	30	298	42	72	58	5	24	30	4	24	18	72	35	1244536665420826
518	880	870	1550	900	810	1710	46	270	33	56	46	13	24	24	4	24	10	60	35	2245534465499826
522	970	940	1910	980	980	1960	40	306	42	72	68	10	24	30	4	24	18	67	31	1255536645426826
600	840	710	1550	990	840	1830	56	292	42	72	47	5		28	4	24	14	72	32	1243535765436283
602	650	820	1470	780	770	1490	65	298	39	72	58	6	24	28	4	24	10	66	44	2245535565436826
604	835	790	1625	960	900	1860	54	298	42	72	52	6	24	30	4	24	11	71	33	2244535556541826
605	1000	780	1780	1000	900	1900	41	275	40	70	35	5		20	4	24	6	72	39	1244535666548725
607	940	840	1580	880	740	1620	48	301	42	72	26	5	24	30	4	24	8	72	38	1234535565436026
608	995	933	1928	1000	934	1938	36	299	42	72	43	5	24	27	4	24	8	72	36	1223535565426826
613	840	830	1670	780	800	1580	73	299	41	72	59	6	24	29	4	24	15	71	40	1244534465426826
615	922	827	1749	956	828	1784	47	301	41	71	62	13	24	30	4	24	13	63	39	2245535655426126
616	987	825	1752	960	910	1870	53	303	42	72	56	5	24	29	5	24	11	72	38	2244534465436026
619	790	745	1535	780	690	1470	58	306	42	72	44	5	24	30	4	24	7	72	42	1244534465426825
620	760	850	1610	760	870	1630	42	293	41	72	29	5	24	30	4	24	3	72	39	2244534365417125
622	750	570	1320	740	640	1380	69	298	42	72	46	5	21	30	4	24	18	72	48	2222534565326825
625	710	590	1300	765	870	1435	54	271	37	62	37	9	20	26	5	24	9	67	38	1445535565326226
627	980	860	1840	880	880	1760	34	299	40	72	43	5	24	28	4	24	11	72	38	1225536665436285
628	880	705	1585	890	690	1580	57	300	42	71	44	5	24	30	4	24	8	72	36	2287535565426926
629	1000	890	1890				36	306	42	72	30	5	24	30	4	24	3	72	56	16345356
633	865	750	1615	930	810	1740	43	300	40	70	58	5	24	30	4	24	8	72	39	2257535565411226

Appendix O

Vita

DANIEL D. BRINKMAN

6919 N.E. 182nd Avenue
Vancouver, Washington 98662
(206) 892-0530

EDUCATION

TRINITY WESTERN UNIVERSITY
Associate of Arts - General Studies, 1979 - honors
Bachelor of Arts - Psychology, 1981 - high honors

WESTERN CONSERVATIVE BAPTIST SEMINARY
Master of Arts - Clinical/Counseling Psychology, 1985
(Doctoral studies, Clinical Psychology in progress)

INTERNSHIP EXPERIENCE

CENTER FOR BEHAVIORAL MEDICINE
PORTLAND ADVENTIST MEDICAL CENTER 1987-89

Half-time internship on psychiatric inpatient locked, voluntary, adolescent, and eating disorder units and outpatient clinics with a multi-disciplinary team. Primary responsibilities were conducting clinical interviews, individual therapy, testing, and co-leading group therapy for patients and for their families. Other activities included accompanying psychiatrists on rounds and developing new programs. Diagnoses based on multiaxial DSM-III-R criteria. Worked with schizophrenia, affective disorders, substance abuse, psychosis, eating disorders, adolescents, neuropsychiatric cases, personality disorders, and multiple personalities.

Supervisors: Robert C. Wolgamott, M.D.; Rodger K. Bufford, Ph.D.
Peggy Loveless, M.S.W., R.C.S.W.

PRACTICUM EXPERIENCE

CATHOLIC FAMILY SERVICES 1986-87

Delivery of counseling services, intake interviews, and testing for outpatients in individual, couple, and family contexts. Provided treatment for anxiety, depression, thought disorders, abuse, marital and family problems, and geriatric issues.

On Site Supervisor: Rikki Schoenthal, R.C.S.W.

HEALTH HELP 1985

Prepared and administered all phases of individual therapy to walk-in clients seeking treatment. Administered, scored, and interpreted many mental health tests, including the MMPI and TAT. Dealt with a variety of mental problems, personality disorders, and histories including depression, anxiety, abuse, marital issues, manic-depression, and anti-social personalities.

On Site Supervisor: Jan Zeedyke, Ph.D.

PORTLAND ADVENTIST CONVALESCENT CENTER 1984-85

Assigned to treat adjustment disorders of geriatric inpatients by developing rapport and addressing crucial and sensitive issues in a low key, non-threatening manner. Also administered mental status exams and became familiar with the social care aspects and concerns of nursing homes. Treated patients adjusting to handicaps, feelings of bitterness, abandonment, loss, denial, and various states of delirium and dementia.

On Site Supervisor: Sharon Nordloff, M.S.W.

RELEVANT EXPERIENCE

CLACKAMAS COUNTY MENTAL HEALTH

Intake Counselor 1987

Contract position involving initial assessment of potential clients requesting mental health services. Assessments included information gathering, evaluating severity of condition, and determining appropriateness of agency to meet their needs. Developed crisis intervention skills and knowledge of community referrals and resources.

Supervisor: Don Sichel, Ph.D.

CATHOLIC FAMILY SERVICES

Social Worker - Friends To The Elderly 1986-87

Part-time position with inter-disciplinary, inter-agency team created to coordinate volunteers to render social support for elderly clients. Main tasks were assessing clients and volunteers, matching participants, training and supporting volunteers.

Supervisor: Judy Alleman, R.N., M.S.

Mental Health Consultant 1986

Contract position providing consultation to service providers of geriatric inpatients in dealing with behavioral and other problems effecting staff and residents. Duties included problem diagnosis, treatment plan formulation, and communication of strategies to service personnel.

Supervisor: Judy Alleman, R.N., M.S.

WESTERN CONSERVATIVE BAPTIST SEMINARY

Group Leader 1986-87

Group leader for several seminary students as part of their course requirements. Group met weekly for two semesters dealing with topics raised by members and addressing interpersonal and group process issues.

Supervisors: Norm Thiesen, Ph.D.; J. Grant Howard, Th.D.

Graduate Fellow 1987-1989

Graduate Fellow for Neal McBride, associate vice president for academic affairs and professor of psychological research. Main responsibilities were assisting students with the statistical aspects of their research, teaching the SPSS/PC+ statistical program, and overseeing the care and proper use of the department computer equipment. Also involved with various class and academic projects.

Supervisor: Neal McBride, Ed.D, Ph.D.

TRINITY WESTERN UNIVERSITY

Resident Assistant 1978-80

Live-in leader and counselor for 15-20 peers in college dorm setting. Assessed individual and group needs then developed and implemented goals and strategies to meet those needs. Commended many times for ability to lead, shape group identity, creatively formulate and organize activities, and resolve problems. Counseling issues confronted included vocational guidance, finances, life goals, depression, relationships, and academics.

Supervisor: Arvid Olson, M.A.

TEST PROFICIENCY

Ammons and Ammons Quick Test
Beery Visual Motor Integration Test
Bender Visual Motor Gestalt Test
Benton Visual Retention Test-Revised
House-Tree-Person
Interpersonal Behavior Survey
Luria-Nebraska Neuropsychological Battery
Luria-Nebraska Screening Test
Minnesota Multiphasic Personality Inventory
Rorschach
Rotter Incomplete Sentences Blank
Stanford-Binet Intelligence Scale: Fourth edition
Thematic Apperception Test
Wechsler Intelligence Scale for Children
Wechsler Adult Intelligence Scale
Wechsler Memory Scale-Revised
Wide Range Achievement Test

PROFESSIONAL INTERESTS

Personality Disorders, Measurement Instruments, Religious Issues

References available upon request.