

The Development of an Instrument for Measuring Healing

James Peter Meza, MD, MSA¹

Gail F. Faboome, PhD²

¹Henry Ford Hospital, Department of Family Medicine, Detroit, Michigan

²Wayne State University, College of Education, Detroit, Michigan



ABSTRACT

PURPOSE Our lack of ability to measure healing attributes impairs our ability to research the topic. The specific aim of this project is to describe the psychological and social construct of healing and to create a valid and reliable measurement scale for attributes of healing.

METHODS A content expert conducted a domain analysis examining the existing literature of midrange theories of healing. Theme saturation of content sampling was ensured by brainstorming more than 220 potential items. Selection of items was sequential: pile sorting and data reduction, with factor analysis of a mailed 54-item questionnaire. Criterion validity (convergent and divergent) and temporal reliability were established using a second mailing of the development version of the instrument. Construct validity was judged with structural equation modeling for goodness of fit.

RESULTS Cronbach's α of the original questionnaire was .869 and the final scale was .862. The test-retest reliability was .849. Eigenvalues for the 2 factors were 8 and 4, respectively. Divergent and convergent validity using the Spann-Fischer Codependency Scale and SF-36 mental health and emotional subscales were consistent with predictions. The root mean square error of approximation was 0.066 and Bentler's Comparative Fit Index was 0.871. Root mean square residual was 0.102.

CONCLUSIONS We developed a valid and reliable measurement scale for attributes of healing, which we named the Self-Integration Scale v 2.1. By creating a new variable, new areas of research in humanistic health care are possible.

Ann Fam Med 2008;6:355-360. DOI: 10.1370/afm.869.

INTRODUCTION

The doctor is the identified healer in Western society.¹ Unfortunately, within clinical medicine the focus on healing has not been a major part of recent scientific discussion or the research agenda. The medical lexicon focuses on screening, diagnosing, treating, curing, and managing diseases. Persons, not diseases, can be healed. Complementary and alternative medicine practitioners embrace the term *healing* in their holistic approach to therapy, as illustrated by Dossey:

When I began to explore the world of alternative medicine nearly 3 decades ago...they often used the word healer, which was not part of the lexicon of medical school. In fact, I do not recall the term's ever being used in my medical training...²

Although cross-cultural studies and complementary and alternative medicine help inform practitioners, our project was designed to promote the discussion of healing by building on the heritage of biopsychosocial medicine as described by George Engel³ and on relationship-centered care as described by Beach and Inui.⁴ Different disciplines or individuals label what seem to be the same concept using different words: revealing the true self/the inner child,⁵ self-actualization,⁶ recovery,⁷ reconciliation,⁸ grieving,⁶ becoming conscious,⁹ individuation,¹⁰ and empowerment.¹¹

Conflicts of interest: none reported

CORRESPONDING AUTHOR

James P. Meza, MD, MSA
One Ford Place, 2E
Detroit, MI 48202
jmeza1@hfhs.org

The purpose of our project was to create a nomenclature and methodological tool so that healing can be studied within the ontological framework of Western medicine.¹²⁻¹⁵ Research within this ontological framework will enhance the explanatory narrative cocreated between doctors and patients. The specific aim of this project was to describe the psychological and social construct of healing and to create a valid and reliable measurement scale of the attributes of being healed, thus enhancing our ability to establish a line of inquiry into this important topic.

Domain Analysis

Nunnally and Bernstein¹⁶ stressed the need to have an a priori theory for domain analysis—even if based on common sense or experience—before we develop a measurement scale. We must conceptualize what we try to measure. The importance of domain analysis cannot be overemphasized. Our project is the culmination of more than 6 years immersed in the study of the healing within the context of a lifetime of medical practice by a reflective practitioner. Validity of the domain analysis is ensured by grounding the analysis in selected, well-articulated, comprehensive models.^{5,6,8,17} The explication of the theoretical construct of healing is actually a composite of those models. It is summarized in Table 1.

Situating This Work Within the Healing Research Context

In addition to describing the healing relationship, Miller et al described a so-called “double helix” trial design for studying the impact of healing relationships.¹⁸ In their model, they include outcome measures, anticipating a measurement instrument. They ask a provocative question: “What is the experience of healing and how I am different as a result of healing?” In the text they specify the domains of healing outcomes that are remarkably

similar to the descriptions in the other models presented here and then state, “One could conceivably construct a composite measure of healing using these... dimensions and some combination of these survey and functional measures.”

By creating a valid and reliable psychometric scale, we explicate the very outcome measure referred to by Miller et al in 2003. It measures the trait of being healed in contrast to a state of being healed, which is a short-term temporal experience. The process of healing implies transformation and continues for a lifetime. Because the dominant descriptor associated with healing is wholeness, and one becomes healed by a process of integration of parts into a whole, this scale is called the Self-Integration Scale (SIS). The double helix model for healing research of Miller et al also calls for a measure of the healing relationship. The SIS does not examine the healing relationship.

METHODS

Content Validity

We ensured content validity through the following procedures. Using the above-mentioned theoretical models as the boundaries for the domain of healing, we generated 202 self-attribution statements using brainstorming techniques. We then specifically examined each of the above major theoretical frameworks for omissions, resulting in 20 additional statements. These additions ensured that the composite list comprehensively described the combined theoretical domain and resulted in a complete content sampling.¹⁶

After individually printing the 222 statements onto index cards, we excluded poorly worded or ambiguous statements that could not be anchored into the theoretical framework. Eighty-three statements were discarded using this method, resulting in 139 statements representative of the domain of healing.

We then categorized the 139 cards by pile sorting. Pile sorting is a process by which the individual index cards are placed on a large flat surface, and statements with similar content are placed geographically close to each other. Questions sorted by content clusters relating to the elements described in the domain analysis resulted in clusters or groupings of similar statements. Within each grouping, we selected statements based on clarity of expression and relevance to the domain. Both positively worded and negatively worded statements for each attribute were selected. Eighty-four statements were thus considered redundant after pile sorting, leaving 55 statements eligible for the construction of the initial questionnaire.

We pilot tested these 55 statements using volunteers to screen for practicality. Their instructions were

Table 1. Summary of the Construct Domain

Healing is the human experience of self-discovery and transformation that results in a sense of being whole and connected. It is observable as:

- Antecedent trauma that initiates behaviors resulting in recurring painful experiences
- A beginning of the healing process based in a therapeutic relationship with another “safe” person holding socially constructed power (the healer)
- Using the therapeutic relationship to facilitate discovering and naming emotions, and repairing/improving relationships with oneself, others, and one’s higher power
- Behavior changes that are healthier, more adaptive, and ameliorate the painful experiences described above.
- The ability to articulate one’s purpose/meaning/mission in life, unrestricted in time
- Reaching out to facilitate healing in others

to identify statements lacking sufficient clarity to provide a response. The wording of several statements was adjusted, and one statement was dropped because it was consistently identified as problematic. The remaining 54 statements, hereafter referred to as test items, made up the original questionnaire.

The entire process of generating statements was designed to reflect the theoretical domain of healing as completely and accurately as possible. This domain analysis is critical because "measurement is much more closely related to the sampling of content, as in deciding which test items to include [rather than sampling people]."¹⁶ It is through the use of domain analysis and the procedure for creating and selecting the items that content validity is ensured.

Internal Consistency and Validity of the Questionnaire

For the purposes of clarity, we call the original 54-item instrument the questionnaire, and we refer to the modified 39-item instrument as the Self-Integration Scale v 1.0 (SIS v 1.0) and the 18-item final measurement instrument as the SIS v 2.1. This instrument is available as a Supplemental Appendix, which can be found online at <http://www.annfamned.org/cgi/content/full/6/4/355/DC1>. Items 2, 3, 4, 11, 12, 13, 14, 16, and 17 should be reverse scored.

Institutional review board approval was obtained. We mailed the questionnaire to a randomized stratified sample (N = 1,080) derived from corporate databases of the general population. This stratified sample was modeled so that responses represented an approximation of the American population. The following characteristics were used: age; sex; socioeconomic status determined by geocoding of census tract information; and health care received in urban, federally funded sites vs suburban sites. The initial mailing did not result in an adequate number of responses, so we sent a follow-up mailing to the same sample encouraging them to participate. A separate mailing to a different sample population using the same stratification characteristics resulted in additional responses. These 3 mailings combined resulted in the first sample. Demographic details of the sampling design and response are provided in Table 2.

We tested the responses to the questionnaire for internal consistency and validity using the Statistical Package for the Social Sciences (SPSS v 11.0, SPSS Inc, Chicago, Illinois). Items were reversed coded, where necessary, to reflect a higher score of healing attributes. Missing values were imputed using the linear trend option in SPSS. Cronbach's α was .869, which is very good, given that Nunnally and Bernstein said that only modest reliability (0.70) was necessary at the early stage of instrument development.

We then conducted an exploratory factor analysis for the purpose of data reduction using principal components analysis and varimax rotation, choosing a factor loading of 0.4 as the minimum threshold for the retention of items. The factor analysis resulted in the retention of 39 items.

We used the resulting 39 items to develop a measurement model using the Amos 5 portion of SPSS. The intent of this analysis was to determine the number of factors and to trim the number of items to approximately 20. The process of model building included consideration of 1-, 2-, and 3-factor models. Model building requires cross-validation to minimize the possibility of a sample-specific model. The sample was therefore randomly divided into 2 unequal samples, with the larger sample (n = 300) for the analysis sample and the smaller (n = 144) for the holdout sample. Eighteen of the 39 items were selected based on their representation of the domain content and the magnitude of the pattern coefficients (factor loadings). The data generated similar results with a 2-factor or 3-factor model. We selected the 2-factor model with 9 observed indicators per factor (Figure 1). Our decision was based on both theoretical considerations and model fit. We confirmed this model with the holdout sample.

Construct Validity

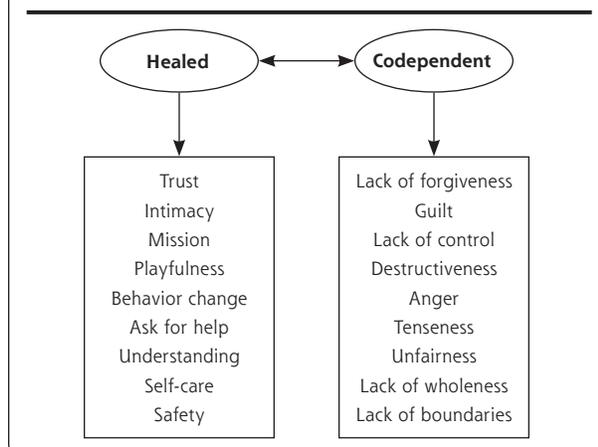
The next step was to test predictive convergent and divergent criterion validity. After the initial measurement model analysis, we mailed the 18-item SIS v 2.0 (embedded in SIS v 1.0) to a randomly selected sample (n = 98), along with the Spann-Fisher Codependency Scale¹⁹ and the SF-36.²⁰ We asked respondents to complete all 3 instruments. We tested predictive con-

Table 2. Demographics of Original Questionnaire Sample (n = 444)

Variable and Strata	Sampling % (Strata)	Data Frequency	Data %
Age, years			
21-35	33	59	13.3
36-55	33	123	27.7
56-75	33	239	53.8
Site			
Suburban	80	356	80.2
Urban	20	65	14.6
Sex			
Male	50	184	41.4
Female	50	237	53.4
Income			
<\$40,000	40	238	53.6
≥\$40,000	60	183	41.2

Note: 23 responses were missing a demographic code.

Figure 1. Two-factor trimmed self-integration model.



vergent criterion validity against the emotional and mental health subscales of the SF-36, for which we expected a weak positive correlation. Predictive divergent criterion validity was tested against the Spann-Fischer Codependency Scale, for which we predicted a strong negative correlation.

Test-Retest Reliability

We mailed SIS v 2.0 (embedded in SIS v 1.0) to a randomly selected sample ($n = 136$) with instructions that they would receive the same questionnaire in approximately 1 month. We monitored dates of mailing and postmarked dates of responses carefully and calculated estimated mail delivery to an approximately 1-month interval. This group of respondents was used to calculate the test-retest reliability coefficient.

As mentioned above, we tested the SIS v 2.0 for model fit against a holdout sample from the questionnaire. We also confirmed the measurement model using the SIS v 2.0 embedded in all initial SIS v 1.0 responses received for the test-retest correlation coefficient grouped with the SIS v 1.0 responses for the convergent and divergent validity mailing. In each of these combinations, we obtained consistent model fit indices, supporting the appropriateness of the model.

RESULTS

Of the 1,080 questionnaires in the first sample, 20 persons declined to participate, and 80 questionnaires were returned because of inaccurate addresses, leaving 980 potential responders. For the second sample ($N = 1,080$), 23 persons declined to participate, and 61 questionnaires were returned because of inaccurate addresses, leaving 996 potential responders. Of the combined samples, 444 were returned for overall

return rate of 22%. The demographics of the resultant sample are displayed in Table 2.

SIS v 2.0 Is Psychometrically Reliable and Valid

Cronbach's α of the original questionnaire, which is a measure of internal consistency, was .869 (very good). Cronbach's α for the 18-item scale from the original sample ($n = 444$) was .862, and for the combined second sample ($n = 323$), it was .857. These values are highly consistent over the 2 samples. The test-retest reliability coefficient, which is a measure of stability over time, was 0.849 (excellent).

Validity

Face validity was confirmed when the items were instantaneously grouped into the 2 factors by the content expert. The content expert also named the 18 items within minutes, further supporting the face validity of the SIS v 2.0. The eigenvalue for factor 1 (healed) was 8 and the eigenvalue for factor 2 (unhealed) was 4, both well above the accepted standard of 1.

Divergent validity, as measured against the Spann-Fischer Codependency Scale, had a predicted strong negative correlation. The correlation coefficient was -0.648 , which provides good support for criterion validity. Divergent validity was further supported in that the correlations between healed and codependent in the 3 structural equation models were not large (analysis sample = 0.645, holdout sample = 0.424, and second sample = 0.557). Convergent validity, as measured against the emotional and mental health subscales of the SF-36, had statistically significant correlations of 0.38 and 0.6, respectively, supporting good criterion validity.²¹

Construct validity was evaluated by the use of confirmatory factor analysis using Amos (v 7.0). The parameter estimates were examined for significance, expected direction, and magnitude. The pattern coefficients in each of the three 2-factor models were significant. The standardized coefficients were between 0.433 and 0.701 for healed and between 0.368 and 0.692 for codependence for all 3 samples.

The goodness of fit for the 2-factor model was estimated by the following fit indices: χ^2 divided by degrees of freedom (CMIN/DF), Bentler's Comparative Fit Index (CFI), root mean square residual (RMR), root mean square error of approximation (RMSEA), and the confidence interval for RMSEA. The values for the selected fit indices for the analysis sample, the holdout sample, and the combined second sample are displayed in Table 3. The fit indices across the 3 samples are very consistent, with the fit of the holdout sample slightly worse than the analysis sample and the combined second sample, which would be expected given the

Table 3. Selected Goodness-of-Fit-Indices for the 3, 2-Factor Models: Analysis Sample, Holdout Sample, and Combined Second Sample

Fit Index	Initial Analysis Sample (n = 300)	Initial Holdout Sample (n = 144)	New Combined Sample (n = 323)
CMIN/DF	2.289	1.976	2.413
CFI	0.890	0.777	0.871
RMR	0.110	0.138	0.102
RMSEA	0.066	0.083	0.066
RMSEA CI	0.056-0.075	0.068-0.097	0.057-0.075

CFI = comparative fit index; CI = confidence interval; CMIN/DF = χ^2 divided by degrees of freedom; RMR = root mean square residual; RMSEA = root mean square error of approximation.

smaller sample size of the holdout sample. The fit of the analysis sample and the second sample are almost identical. Moreover, all fit indices are close to or within the generally accepted criteria for each index. CMIN/DF values for the 3 samples are all less than 3.0 and closer to 2.0, indicating adequate fit. RMR values of less than 0.10 are generally considered favorable. The values for the samples are very close, with RMR = 0.102 for the second sample. RMSEA is a parsimony-adjusted index that does not assume perfect fit. All 3 values are less than the recommended 0.10, and the analysis sample and second sample are between 0.05 and 0.08, suggesting close approximate fit. RMSEA provides a 90% confidence interval. Ideally the lower limit would be below 0.05 and the upper limit would be lower than 0.10. In both the analysis sample and second sample, the SIS v 2.0 limits are between 0.05 and 0.08, supporting reasonable error of approximation. CFI is an incremental fit index. CFI of less than 0.90 (roughly) is indicative of good fit. The SIS v 2.0 values for both the analysis sample and the second sample are greater than 0.87, suggesting reasonably good fit. The consistency of the fit indices over the 3 samples provides further support for the 2-factor measurement model.

DISCUSSION

Healing exists within the doctor-patient relationship regardless of whether it is acknowledged. Michael Balint²² referred to the drug "doctor" that has effects and side effects. He vehemently called for research in this area of clinical practice. Similarly, Arthur Kleinman, MD, PhD,²³ describes "medical psychotherapy" and states that it happens with or without its being acknowledged. He also called for research in the "medical social science and humanities":

As long as we lack such knowledge, the development of new paradigms of practice and effective treatment strategies will be delayed, and the research enterprise

will remain enormously unbalanced toward disease questions.²³

The most important feature of this work is that it explicates the construct of healing as it is relevant to Western medical practice; furthermore, this attempt is the first to quantify attributes of being healed in this setting. Instead of using such terms as reconciliation, self-actualization, and recovery, it is now possible

to refer to healing as measured with the SIS v 2.1 with specificity and clarity of meaning, facilitating improved scientific communication. The SIS v 2.1 also gives researchers a new tool to explore relationships between variables in health and health care. As Stange et al²⁴ stated, new knowledge is required to create new models of health care. This work offers a reference point for future research in the topic area and a counterbalance to the biomedical model of care.

Although we made attempts to create a sample group representative of the general US population, our actual sample has the inherent limitations of any research project. Although the SIS v 2.1 promises to be a useful research tool, it would be inappropriate to use it for individuals or in clinical practice. Normative data for subgroups of the population do not exist.

To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/cgi/content/full/6/4/355>.

Key words: Healing; research design; measurement; psychometrics

Submitted September 24, 2007; submitted revised March 20, 2008; accepted April 3, 2008.

References

1. Frank JD, Frank JB. *Persuasion and Healing—A Comparative Study of Psychotherapy*. 3rd ed. Baltimore, MD: The Johns Hopkins University Press; 1991.
2. Dossey L. *Healing Beyond the Body*. 1st ed. Boston, MA: Shambhala Publications, Inc; 2001.
3. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196(4286):129-136.
4. Beach MC, Inui T. Relationship-centered care—a constructive reframing. *J Gen Intern Med*. 2006;21(Suppl 1):S3-S8.
5. Whitfield CL. *Healing the Child Within*. Deerfield Beach, FL: Helath Communications, Inc; 1989.
6. Maslow AH. *Toward a Psychology of Being*. 3rd ed. New York, NY: John Wiley & Sons, Inc; 1999.
7. Alcoholics Anonymous World Services I. *Alcoholics Anonymous*. 2004 2001.

8. Egnew TR. The meaning of healing: transcending suffering. *Ann Fam Med*. 2005;3(3):255-263.
9. Frattaroli E. *Healing the Soul in the Age of the Brain*. New York, NY: Viking Press; 2001.
10. Stevens A, Storr A. *Jung, Freud & Jung—A Dual Introduction*. New York, NY: Barnes and Noble, Inc; 1994.
11. Herman JL. *Trauma and Recovery*. New York, NY: Basic Books; 1992.
12. Sulmasy DP. Is medicine a spiritual practice? *Acad Med*. 1999;74(9):1002-1005.
13. Loxterkamp D. A vow of connectedness: views from the road to Beaver's Farm. *Fam Med*. 2001;33(4):244-247.
14. Katon W, Sullivan M. Of healing and human suffering. *J Am Board Fam Pract*. 1997;10(1):62-65.
15. Borrell-Carrio F, Suchman AL, Epstein R. The Biopsychosocial Model 25 years later: principles, practice, and scientific inquiry. *Ann Fam Med*. 2004;2(6):576-582.
16. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw-Hill, Inc; 1994.
17. Waters DB, Rudolph A. *A Curriculum in Healing*. Paper presented at: STFM 34th Annual Spring Conference, 2001; Denver, CO.
18. Miller WL, Crabtree BF, Duffy B, Epstein R, Stange K. Research guidelines for assessing the impact of healing relationships in clinical medicine. *Altern Ther Health Med*. 2003;9(3):A80-A95.
19. Fischer JL, Spann L, Crawford D. Measuring codependency. *Alcohol Treat Q*. 1991;8(1):87-100.
20. Ware J, Snow K, Kosinski M, Gandek B. *SF-36 Health Survey*. Boston, MA: The Health Institute, New England Medical Center; 1993.
21. Kline RB. *Principles and Practice of Structural Equation Modeling*. New York, NY: The Guilford Press; 2005.
22. Balint M. *The Doctor, His Patient and the Illness*. Edinburgh: Churchill, Livingstone; 1957 [2000].
23. Kleinman A. *The Illness Narratives—Suffering, Healing, and the Human Condition*. New York, NY: Basic Books; 1988.
24. Stange KC, Miller WL, McWhinney I. Developing the knowledge base of family practice. *Fam Med*. 2001;33(4):286-297.