The Measure of
Service Learning

Research Scales
to Assess Student Experiences

Robert G. Bringle
Mindy A. Phillips
Michael Hudson

American Psychological Association, Washington, DC

CONTENTS

Page numbers refer to *the original.*

Chapter 1: Understanding Service Learning (p. 3 of original)

Along with *foundational* knowledge (content and cross-disciplinary knowledge) and *professional* knowledge (practitioner skills and content), *socially responsive* knowledge should be an integral part of the undergraduate curriculum. How can the challenge of educating future generations include socially responsive knowledge in a manner that is pedagogically sound?

John Dewey provided the theoretical underpinnings for understanding good instruction. He specified four conditions that maximize the potential for inquiry-based learning. It must:

(a) generate interest in the learner;
(b) be intrinsically worthwhile to the learner;
(c) present problems that awaken new curiosity and create a demand for information; and
(d) cover a considerable time span and be capable of fostering development over time.

Dewey wrote (1933) that

> Thinking begins in what may fairly enough be called a forked-road situation, a situation that is ambiguous, that presents a dilemma, that proposes alternatives... Demand for the solution of a perplexity is the steadying and guiding factor in the entire process of reflection.

During the past 70 years, research has identified a number of elements that enhance depth of understanding in the learning process. These include

(a) active learning;
(b) frequent feedback from experts, students, or others (e.g., community practitioners) that is provided in nonthreatening ways;
(c) collaboration;
(d) cognitive apprenticeship (i.e., a mentor with whom students can discuss and learn generalization of principles, transfer of knowledge between theory and practice, and how to analyze perplexing circumstances); and
(e) practical application in which students are involved in tasks that have real consequences but have a safety net for high-stakes mistakes (Marchese, 1997).

These elements are compatible with Dewey's analysis and with other approaches to good pedagogy. From multiple perspectives, then, there has been a convergence of opinion on the qualities that are integral to designing effective learning environments.

Service Learning as a Pedagogy (p. 5 in original)

Service learning classes engage students in service activities that simultaneously pursue two goals: (a) benefit to community stakeholders (e.g., agency, clients, neighborhood

residents) and (b) academic learning outcomes. The service experience provides a rich text from which academic lessons are learned through the interplay between theory and practice. The educational outcomes are derived from community service through reflection activities (e.g., journals, small group discussions, directed writing; see Eyler, Giles, & Schmiede, 1996). The presumption is that community service does not necessarily, in itself, produce learning.

There are three goals for socially responsive knowledge:

(1) to educate students in the problems of society;
(2) to let have them experience and understand first-hand social issues in their community; and, most important,
(3) give students the experiences and skills to act on social problems. (Altman, 1996, pp. 375-376).

There is increasing evidence that service learning is effective in helping students develop socially responsive knowledge as well as facilitating learning in the more traditional domains of content and skills, such as the capacity to view phenomena from multiple perspectives and to apply knowledge developed in one setting to other settings (Eyler & Giles, 1999; Eyler, Giles, Stenson, & Gray; 2001).

Service learning represents a paradigm shift in higher education because it heightens the role that students can assume as constructors of knowledge. Furthermore, service learning shifts the role of the instructor from the center of instruction to the facilitator of learning that occurs outside the classroom. Part of the strength of service learning is illustrated in its compatibility with other pedagogical trends in education, such as collaborative learning, problem-based learning, and diversity education. Service learning is also compatible with other changes taking place in higher education that has observed: (a) a shift from a focus on teaching to student learning; (b) moving from an emphasis on autonomous, individualistic work to collaborative, interdisciplinary work; and (c) a change from the isolated character of higher education to a more public and democratic approach to academic work. (Rice, 1996)

Well-executed service learning involves a coordinated partnership between the campus and the community, with the instructor tailoring the service experience to the educational agenda and community representatives ensuring that the students' community service is consistent with their goals (Zlotkowski, 1999). In addition, well-designed reflection activities should

(a) intentionally link the service experience to course-based learning objectives,
(b) be structured,
(c) occur regularly,
(d) allow feedback and assessment, and
(e) include the clarification of values (Bringle & Hatcher, 1999; Hatcher & Bringle, 1997).

<u>Research on Impact of Service Learning</u> (p. 7 of original)

 This volume (Bringle, Phiilips, and Hudson, 2004) is devoted to **scales that can be used in research on students in service learning class**es. Chapter 2 provides an overview of the research process to discuss the interplay between theory and research. Because focusing on scales risks suggesting that a scale's selection s the major task in conducting research, chapter 2 discusses the deductive and inductive processes that surround the measurement of constructs. Thus, the characteristics of good measurement (reliability, construct validity) are put within the context of other qualities (qualities of good theories, external validity) that must be considered when designing and conducting good research.

 Chapter 3 discusses in more detail the nature of multiple-item scales and the role these scales can play in developing research to answer different types of questions. A Glossary compiles the definitions of key terms.

<u>Summary</u> (p. 8 of original)

 Research needs to establish that service learning contains the critical elements that promote good academic learning: active engagement, frequent feedback, collaboration, cognitive apprenticeship, and practical application (Marchese, 1997).

REFERENCES

Altman, I. (1996). Higher education and psychology in the millennium. American Psychologist, 51, 371-378.
Brewer, C. L. (1997). Undergraduate education in psychology: Will the mermaids sing? American Psychologist, 52, 434-441.
Bringle, R. G. (2002, October). Enhancing theory-based research on service learning. Keynote address presented at the Second International Conference on Service Learning Research, Nashville, TN.
Bringle, R. G., & Hatcher, J. A. (1995). A service-learning curriculum for faculty. Michigan Journal of Community Service Learning, 2, 112-122.
Bringle, R. G., & Hatcher, J. A. (1999). Reflection in service learning: Making meaning of experience. Educational Horizons, 77, 179-185.
Bringle, R. G., & Hatcher, J. A. (2000, Fall). Assessment: Meaningful measurement of theory-based service-learning outcomes. Michigan Journal of Community Service Learning, 68-75.
Dewey, J. (1916). Democracy and education. New York: Macmillan.
Dewey, J. (1933). How we think: A restatement of the relation of reflective thinking to the educative process. Boston: D.C. Heath.
Enos, S. L., & Troppe, M. L. (1996). Service-learning in the curriculum. I n B. Jacoby & Associates (Eds.), Service-learning in higher education (pp. 156-181). San Francisco: Jossey-Bass.

Eyler, J., & Giles, D. E., Jr. (1999). Where's the learning in service-learning? San Francisco: Jossey-Bass.

Eyler, J., Giles, D. E., & Schmiede, A. (1996). A practitioner's guide to reflection in service-learning: Student voices and reflections. Nashville, TN: Vanderbilt University.

Eyler, J., Giles, D. E., Jr., Stenson, C. M., & Gray, C. J. (2001). At a glance: What we know about the effects of service-learning on college students, faculty, institutions and communities, 1993-2000 (3rd ed.). Nashville, TN: Vanderbilt University.

Furco, A. (1996). Service-learning: A balanced approach to experiential education. In Corporation for National Service, Expanding boundaries: Serving and learning (pp. 2-6). Columbia, MD: Cooperative Education Association.

Hatcher, J. A. (1997). The moral dimensions of John Dewey's philosophy: Implications for undergraduate education. Michigan Journal of Community Service Learn- ing, 4, 22-29.

Hatcher, J. A., & Bringle, R. G. (1997). Reflections: Bridging the gap between service and learning. Journal of College Teaching, 45, 153-158.

Howard, J. (1993). Community service in the curriculum. In J. Howard (Ed.), PRAXIS I: A faculty casebook on community service learning (pp. 3-12). Ann Arbor, MI: OCSL Press.

Giles, D. E., & Eyler, J. (1994). The theoretical roots of service-learning in John Dewey: Towards a theory of service-learning. Michigan Journal of Community Service Learning, 1, 77-85.

Giles, D. E., Jr., & Eyler, J. (1998). A service-learning research agenda for the next 5 years. In R. Rhoads & J. Howard (Eds.), Academic service learning: A pedagogy of action and reflection (pp. 65-72). San Francisco: Jossey-Bass.

Marchese, T. J. (1997). The new conversations about learning: Insights from neuro- science and anthropology, cognitive studies and work-place studies. In Assessing impact: Evidence and action (pp. 79-95). Washington, DC: American Association for Higher Education.

Mathews, D. (1995). The politics of diversity and the politics of difference: Are academics and the public out of sync? Higher Education Exchange, 66-71.

McGovern, T. V., & Reich, J. N. (1996). A comment on the Quality Principles. American Psychologist, 51, 252-255.

Rice, R. E. (1996, January). Making a place for the new American scholar. Paper presented at the American Association for Higher Education Conference on Faculty Roles and Rewards, Atlanta, GA.

Zlotkowski, E. (1999). Pedagogy and engagement. In R. Bringle, R. Games, & E. Malloy (Eds.), Colleges and universities as citizens (pp. 96-120). Boston: Allyn & Bacon.

Chapter 2: Overview of Scientific Research and Measurement (p. 11 of original)

Collecting information about outcomes associated with teaching a course can be done in a variety of ways, including structured interviews, casual conversations with students after class, classroom assessment techniques, student portfolios, peer observation, student evaluations, course examinations, and reading student journals. In addition, practitioners can conduct formal research that uses standardized scales to measure constructs that are related to or part of intended educational outcomes. Although any of these methods may be useful, some provide more meaningful information than others do for specific purposes.

This chapter provides an overview of the deductive process of translating concepts into researchable questions, gathering meaningful information, and generalizing those findings to other circumstances and to theories. More-detailed attention is given to the rationale for multiple-item scales and the advantages that they offer to research on service learning.

The Nature of Inquiry (p. 12 of original)

The effectiveness and meaningfulness of research is tied to its theoretical context. Theories clarify, simplify, and provide context within which diverse observations, information, inferences, and connections can be interpreted and understood. They are explanations for events that provide a rich set of heuristics through which to explore auxiliary phenomena, boundary conditions (i.e., when the theory is not applicable), and alternative conceptual frameworks. Theory and research are equally important to the process of accumulating knowledge through the scientific method (Bringle, 2002).

Deduction (p.13)

Deduction is the process of going from general, abstract constructs to specific, concrete manifestations of a construct. *Constructs* are abstract or hypothetical entities that make sense of a diverse set of phenomena. For example, consider critical thinking. The verbal, written, or behavioral manifestations of critical thinking may be apparent to an observer; they can help teachers identify differential growth among students; and they provide a means for ranking persons on attributes associated with critical thinking. However, the construct of critical thinking is not directly accessible; it is inferred from a coherent set of behavioral and verbal manifestations that are presumed to indicate and represent the construct. Theorists can map the conceptual domain and identify attributes that are presumed to be indicative of good or poor critical thinking, but the construct itself does not exist in a tangible way.

Theories are composed of statements about the nature of constructs, their manifestations (cognitive, affective, behavioral), and the relationship between constructs. Theories ask and answer "why?" questions. For example, a theory would explain why a particular course quality (e.g., written reflection on alternative problem-solving strategies for a community issue) produces a particular outcome (e.g., enhanced self-efficacy on community issues).

Variables are the "causes" and "effects" of a construct. They are (a) quantitative—they vary in intensity or degree (e.g., attitude, satisfaction), or (b) qualitative—they differ in kind (e.g., gender, types of instruction). Many variables that can be associated with a particular construct.

Operationalization is the way a variable is measured (e.g., observation of behavior, scale) or manipulated in research (e.g., offering a service section and a traditional section of a course, varying the type of reflection activities that students in the two sections undertake). Typically, there are multiple ways to operationalize a variable, and a researcher must select among them. For example, civic skills can be measured by observing a student at a service site, analyzing journal entries, and presenting civic problems to a student and asking experts to code solutions proposed by a student. The adequacy of each operationalization depends on the degree to which it corresponds to the underlying construct that is embedded within a theoretical context. The progression from construct to variable to operationalization goes from abstract to more concrete, from broad to more specific, and from theoretical to empirical.

Induction: Generalizability (p. 14 of the original)

Induction is the generalization to broad conclusions from specific facts or findings. For example, generalizing from observations of a single service learning class to a general principle about learning is an inductive process. The quality of induction relies on evaluating the degree to which the specific information forms an appropriate basis for a general conclusion. Because such a conclusion exceeds the information on which it is based, the evaluation of the meaningfulness of inductions will be a matter of degree (i.e., of likelihood or probability).

External validity is the degree to which research findings may be generalized to additional cases, situations, instances, and so forth. Several aspects of the evidence determine the *level of confidence* with which generalizations can be made, i.e., the extent to which we can feel assured that the outcomes owes to the independent (causal) variable) and not to random chance). The more representative a research sample is of the larger population, the stronger the association will be.

Measurement (p. 16)

Interviews

The simplest approach to evaluating what people do, believe, feel, etc. is to ask them. However, researchers understand the difficulties and biases inherent in this simple prescription. One risk inherent in face-to-face interviews, for example, is biases that are introduced by the interviewer (e.g., his or her gender, age, status, race, socioeconomic status, formal education, etc.). Inaccuracies can be generated as well by the manner in which interviewers carry out an interview (e.g., not following directions, asking questions out of order, tone of voice, improvisation).

Standardized Scales (p. 17)

A *scale* is essentially a structured interview on paper. The questions and prompts can be open-ended, with no specified format or categories for responding, or the responses can be structured (e.g., from "strongly agree" to "strongly disagree"). One advantage of a paper-and-pencil scale is that all sources of variability attributable to characteristics of the interviewer are eliminated.

Characteristics of Good Standardized Scales (p. 17)

Standardized scales are a product of a measurement strategy that has its roots in the deductive translation of a construct into a set of measurement procedures (Nunnally, 1967). A paper-and-pencil standardized scale is intended to measure a single construct (e.g., attitude toward elderly persons, self-efficacy). As such, a scale should display qualities consistent with the assumptions of being unidimensional and measuring only one construct. Although most *questionnaires* are composed of several scales, each individual *scale* is a multiple-item measure of only one construct. *Measurement* is the process through which numbers are assigned to interpersonal variations in a construct or attribute. A good scale, then, should produce numbers that correspond to individual differences in the attribute.

Validity (p. 20)

[from Wikipedia]: "*Validity* has no single agreed definition but generally refers to the extent to which a concept, conclusion or measurement is well-founded and corresponds accurately to the real world. … The validity of a measurement tool (for example, a test in education) is considered to be the degree to which the tool measures what it claims to measure."

Validity is the ability of a procedure to measure the construct it is supposed to measure. It is the degree to which a particular measure of a construct produces scores that confirm the expectations contained in the theory. A construct thus takes its meaning from the theory that specifies the construct's nature, manifestations, and confirmation of hypothesized relationships of the measure to other constructs (Cronbach & Meehl, 1955).

There is no single index of construct validity. Assessing the validity of a measure of a construct must occur within the theoretical context of the construct. Validity is context-specific, is situation-specific, and is established for a particular use of a measure. For many constructs (e.g., racial tolerance, critical thinking, civic responsibility), there is no single criterion against which all scores can be compared.

Example: the construct of "social justice" can have different meanings depending on whether it is embedded in a theoretical perspective that is moral, psychological, or economic. The validity of a scale that measures one meaning of social justice is connected to its theory, and its construct validity is dependent on how well the scale measures that particular meaning and conforms to the predictions of that theory. A different scale that comes from a different theoretical perspective must similarly be evaluated in terms of its theoretical context and predictions.

Reliability (p. 18)

[from Wikipedia:] "*Reliability* is the consistency of a set of measurements or of a measuring instrument. Reliability is inversely related to random error."

Reliability does not imply validity. A reliable measure is one that measures a construct consistently. But you may not be measuring what you want to be measuring. For example, while there are many reliable tests of specific abilities, not all of them would be valid for predicting, say, job performance. In terms of accuracy and precision, reliability is analogous to precision, while validity is analogous to accuracy.

An example often used to illustrate the difference between reliability and validity in the experimental sciences involves a common bathroom scale. If someone who is 200 pounds steps on a scale 10 times and gets readings of 15, 250, 95, 140, etc., the scale is not reliable. If the scale consistently reads "150," it is reliable, but not valid. If it reads "200" each time, then the measurement is both reliable and valid. This is what is meant by the statement, "Reliability is *necessary but not sufficient* for validity."

Reliability may be estimated through a variety of methods that fall into two types: multiple-administration and single-administration.

Multiple-administration methods require the administration of two assessments. In the "test – retest"method, reliability is estimated as the Pearson product-moment correlation coefficient between two administrations of the same measure: In the "alternate forms" method, reliability is estimated by the Pearson product-moment correlation coefficient of two *different* forms of a measure, usually administered together.

Single-administration methods include *split-half* and *internal consistency*.   The "split-half" method treats the two halves of a measure as alternate forms. This "halves reliability" estimate is then stepped up to the full test length using the Spearman-Brown prediction formula.

[from Wikipedia:] "Internal consistency is typically a measure based on the correlations between different items on the same test (or the same subscale on a larger test). It measures whether several items that propose to measure the same general construct produce similar scores. For example, if a respondent expressed agreement with the statements "I like to ride bicycles" and "I've enjoyed riding bicycles in the past", and disagreement with the statement "I hate bicycles", this would be indicative of good internal consistency of the test. The most common internal consistency measure is Cronbach's alpha, which is usually interpreted as the mean of all possible split-half coefficients."

Scales and Other Measurement Procedures (p. 22)

Standardized scales are valuable measuring tools because they are easily transportable; they are efficient and effective ways of obtaining information; and they are readily coded and analyzed. They are not the only way to conduct research on service learning. As you design your research, you should consider alternative forms of measurement, such as observation, interviews, qualitative data, physiological measures, focus groups, and archival information. Researchers who employ multiple sources of data and multiple methods for collecting data

are in a better position to establish converging results (see Gelman, Holland, Driscoll, Spring, & Kerrigan, 2001).

Single- and multiple-indicator indices can be taken from observations of behaviors, reports by key persons (e.g., site supervisors), samples of journal entries, physiological measures of empathic reactions, ratings of job performance, and archival measures (e.g., transcripts). In all cases, these indices are improved by having multiple indicators.

Summary (p. 23)

Service learning practitioners can draw on multiple sources of information to gauge the development of their work and to convey its value to other audiences. Different types of inquiry (reflection, experimental, survey, qualitative, correlational) have both strengths and weaknesses, and the meaningfulness of any one source or type of information varies depending on the appropriateness of its application. Scientific research represents a particular type of systematic inquiry incorporating inductive and deductive inferential processes between theories and research procedures.

Theories contain statements about the nature of constructs and the relationships among constructs. Research on constructs must employ measurement procedures that are accurate (i.e., error free), consistent across time (i.e., reliable), and coherent (i.e., internally consistent).

REFERENCES (p. 23)

Angelo, T. A., & Cross, K. P. (1993). Classroom assessment techniques: A handbook for college teachers (2nd ed.). San Francisco: Jossey-Bass.

Bringle, R. G. (2002, October). Enhancing theory-based research ori service learning. Paper presented at the Second International Conference on Service Learning Research, Nashville, TN.

Bringle, R. G., & Hatcher, J. A. (2000, Fall). Assessment: Meaningful measurement of theory-based service-learning outcomes. Michigan journal of Community Ser- vice Learning, 68-75.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psycho-metrika,l6, 297-334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-301.

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

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. Journal of Personality and Social Psychology, 37, 1097-1126.

Gelmon, S. B., Holland, B. A., Driscoll, A., Spring, A., & Kerrigan, S. (2001). Assessing service-learning and civic engagement: Principles and techniques. Provi-dence, RI: Campus Compact.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw Hill.

Walvoord, B. E., & Anderson, V. J. (1998). Effective grading : A tool for learning and assessment. San Francisco: Jossey-Bass.

CH. 3: THE USE OF SCALES IN SERVICE LEARNING RESEARCH (p. 25)

Program evaluation and research have a long history in education (Sansone, Morf, & Panter, 2003). To some extent, the existing literature on measurement and design issues can be applied to service learning with good results. However, service learning has special characteristics such as unique aspects of the pedagogy (e.g., reflection) that warrant additional resources for good research. This is particularly the case for unique outcomes expected from service learning classes (changes in values, moral develop- ment, civic outcomes; Shumer, 2000)

Scales: Existing versus Modifying versus Developing

Using existing scales has many advantages over developing original scales. Existing scales take less time to incorporate into research, are usually prepared by researchers who have professional expertise, may have norms available against which a particular sample can be compared, and have a known record of psychometric qualities (although these may vary from sample to sample).

    Existing scales may also be modified to suit a particular research context or question associated with a service learning class. Adapting a scale has the advantage that most of the work has been completed and the resulting scale may be more appropriate than the original scale. Modifying a scale, however, runs the risk of changing a known quantity in unknown ways.

    Developing an original scale allows a researcher to design a scale for a specific need. However, developing a new scale takes time and resources (e.g., literature review, pilot testing), requires knowledge of scale development procedures, and runs the risk that the result may not be a good scale.

Inclusion and Exclusion Criteria (p. 26)

Two primary criteria were used to select scales for this volume from the empirical literature that was examined: (a) relevance to service learning, and (b) the quality of supporting evidence.

Relevance to Service Learning

The scales were selected because they include constructs that were examined in past research on service learning or they are present in or related to conceptual discussions of service learning (e.g., Eyler & Giles, 1999; Giles & Eyler, 1998).   For selected scales that have not been used in past research on service learning, we judged the relevance of a scale's content on the basis of we thought it might prove useful.

    Some scales were selected not because they are appropriate as an outcome measure but because they might prove interesting as a moderating or mediating variable. A mediating variable describes an intervening variable that is assumed to explain the relationship between

two variables (Baron & Kenny, 1986). Mediating variables are useful for exploring theoretical explanations for why a relationship exists between two variables.

A moderating variable, which describes an "it depends" relationship, specifies the conditions for which the relationship holds and other conditions for which a different relationship holds (Baron & Kenny, 1986). As such, moderator variables are useful for exploring the boundary conditions for a phenomenon

This book ***does not include*** *scales to measure the* **academic** *performance of students* in service learning classes. Achievement scales of this type are much too course-specific to be readily transportable. The book also ***does not contain*** *scales that measure* **student satisfaction** for a service learning class. Furthermore, the focus of this book is exclusively on the processes and outcomes of *students*. **No** attempt has been made to compile scales to study other aspects of service learning such as **community impact** (Cruz & Giles, 2000), **institutional attributes** (Holland, 2000), and **faculty** *(e.g., motives, professional development, outcomes*; Driscoll, 2000).

Practitioners must determine the appropriateness of a construct and scale. This determination should consider the design of the course (e.g., educational goals), the implementation of the course (e.g., selection of course activities and reflection assignments), and expected outcomes. A sound rationale that relates goals, course activities, and outcomes must exist for an expected outcome prior to the selection of a scale to measure that outcome. Research and program evaluation directed at service learning will benefit immensely when more practitioners develop multiple-item measures of constructs that target specific outcomes of service learning (e.g., civic skills, civic responsibility, citizenship), for which few scales could be found.

Some constructs are conceptualized as being stable across time and situations (e.g., self-esteem, self-efficacy). These constructs are trait-like in nature, and the measure considers the construct in a global sense. In contrast, there are other constructs that are assumed to be more malleable and sensitive to environmental influences (e.g., attitude, mood). These constructs are state-like in that the measure focuses on a particular time and setting.

Scales measuring trait-like constructs may be of interest to researchers as precursors that predispose individuals to behave, think, or feel in particular ways during a service learning experience. In this case, their role may be as a moderator variable clarifies a set of conditions under which certain effects take place. In contrast to this use, a measure of a trait-like construct would be inappropriate as a pre-, post test dependent variable if there is little reason to assume that a brief educational experience (i.e., a course) would change a trait (e.g., self-esteem). Malleable, -dependent constructs would be more appropriate for dependent variables in this case (e.g., attitudes).

The nature of some constructs may vary in the degree to which they are construed as being trait-like or state-like in nature. For example, self-esteem is assumed to be a relatively stable attribute that is a function of a lifetime of experiences. However, one could conceive a construct that is the person's self-esteem as a student. In this case, the construct might more reasonably be influenced by experiences during a course or semester. Furthermore, the construct could be the person's self-esteem as a student in a particular class; that is, how favorably or unfavorably does the person feel as a student during this service learning class? In this case, the focus of the construct is more situation specific and the expectations for the

construct being influenced by the experience are more reasonable than with the global, trait-like construal. Thus, scales that are designed to measure trait-like constructs can be customized to focus the attention on state-oriented aspects of the attribute.

Quality of Supporting Evidence (P. 28)

Scales were selected not only because of their relevance to research on service learning but also because they are multiple-item measures of constructs that possess minimally acceptable psychometric characteristics (i.e., reliability, validity).   Numerous scales were located in research on service learning that are multiple-item measures, but few were accompanied by supporting evidence that the collection of items was intended to measure a specific construct.

<u>Scale Entries</u> (p. 29)

Researchers must select scales based on the appropriateness of the scale to the broader research questions that are being asked. To assist researchers, **the entries in chapters 4-9 provide a description of each scale that identifies the construct that is measured, a description of the scale, and some of the evidence about the psychometric properties of the scale** that could be identified in existing published and unpublished (e.g., dissertations) research. In most cases, the descriptions have been shared with authors and the entries have been edited and updated according to their suggestions. The purpose of these entries is to provide sufficient information to guide researchers in identifying potentially useful scales. However, researchers who are selecting scales are advised to consult the primary sources to learn more about the theory and research that is critical to determining if a particular scales is appropriate. Potential users of these scales must also be aware that additional research may be available that would influence their decision to select a scale.

   <u>Some unpublished scales are unpublished</u>. In these cases, the authors' contact information at the time of publication is given.

<u>Summary</u> (p. 30)

Although the scientific approach that emphasizes quantitative measurement may seem too structured and too narrow (see Shumer, 2000), accumulating this type of research is important (a) to practitioners as a means of improving knowledge and confidence, and (b) to external audiences to increase their appreciation for how service learning can help the academy do its most important work: educating students to be better informed, more fully functioning, and active individuals in their communities. *The scales contained in this volume are but one tool to be used by practitioners to engage in scientific research*. Researchers must use good judgment and creativity to determine which tools are appropriate (e.g., scales, observations, archival measures, interviews, qualitative methods), which procedures are best suited to investigate a particular question, what inferences should be drawn from their research, and what are the best ways to communicate the results to multiple audiences.
.

REFERENCES (p. 30)

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considera- tions. Journal of Personality and Social Psychology, 51, 1173-1182.

Cruz, N. I., & Giles, D. E., Jr. (2000, Fall). Where's the community in service-learning research? Michigan Journal of Community Service Learning, 28-34.

Driscoll, A. (2000, Fall). Study faculty and service-learning: Directions for inquiry and development. Michigan Journal of Community Service Learning, 35-41.

Eyler, J., & Giles, D. E., Jr. (1999). Where's the learning in service-learning? San Francisco: Jossey-Bass.

Fink, A., & Kosecoff, J. (1998). How to conduct surveys: A step-by-step guide (2nd ed.). Thousand Oaks, CA: Sage.

Giles, D. E., Jr., & Eyler, J. (1998). A service-learning research agenda for the next 5 years. In R. Rhoads & J. Howard (Eds.), Academic service learning: A pedagogy of action and reflection (pp. 65-72). San Francisco: Jossey-Bass.

Holland, B. A. (2000, Fall). Institutional impacts and organizational issues related to service-learning. Michigan Journal of Community Service Learning, 52-60.

Markus, G. B., Howard,). P. F., &King, D. C. (1993). Integrating community service and classroom instruction enhances learning: Results from an experiment. Educational Evaluation and Policy Analysis, 15, 410-419.

Sansone, C., Morf, C. C., & Panter, A. T. (2003). The sage handbook of methods in social psychology. Thousand Oaks, CA: Sage.

Shumer, R. (2000, Fall). Science or storytelling: How should we conduct and report service-learning research? [Special issue]. Michigan]ournal of Community Service Learning, 76-83.

Thorton, J. (1999). A field guide to the soul: A down-to-earth handbook of spiritual practice. New York: Bell Tower.