

Assessing Service-Learning

Pamela Steinke
North Carolina State University

Peggy Fitch
Central College

Abstract

In an effort to increase the quality and quantity of service-learning assessment, this article provides a brief rationale for engaging in service-learning assessment and reviews a selection of available tools for doing so. The paper includes a matrix that lists cognitive outcomes, including critical thinking and problem solving, and the instruments that measure them. The conclusion emphasizes the role of service-learning assessment in transforming current assessment debates.

Assessing Service-Learning

Over two decades of research demonstrate that high quality service-learning enhances student learning (Eyler & Giles, 1999, 2001). As calls for greater accountability and evidence-based practice in higher education increase (e.g., U.S. Department of Education, 2006), service-learning researchers and practitioners are poised to provide leadership for a “new model of excellence” in higher education (Zlotkowski, 1998). In order for service-learning to gain recognition in this role, however, it must be fully integrated into departmental and general education assessment processes at colleges and universities. In this article we explain why quality assessment of service-learning is important to higher education. We then describe several tools that measure cognitive outcomes associated with service-learning.

Virtually all definitions of service-learning refer to an organized educational experience that both meets needs of the community and fulfills learning objectives. However, for the purposes of this paper, service-learning also incorporates credit-bearing courses that include reflection activities that connect the student’s experience with course content and the wider discipline (Bringle & Hatcher, 1995). Consistent with this focus on academic service-learning, assessments discussed here will include only those designed to measure student learning outcomes specific to gains in cognitive skills. Therefore, other data-gathering activities often associated with service-learning, such as logging the number of hours students complete, are not included.

Given the ever-growing role of assessment in shaping curricula, assessments must demonstrate that local service-learning efforts enhance locally specified student learning outcomes. Regardless of favorable published research, without this evidence, service-learning will never gain the influence it needs to transform the curriculum. Systematic assessments of service-learning provide opportunities to demonstrate the powerful impact that this pedagogy can have on student learning in a way that speaks directly to those individual faculty members and administrators who design the curriculum at the local level.

Better assessment also provides a way of interjecting service-learning into the national dialogue about the quality of undergraduate education in the U.S. (e.g., Association of American Colleges and Universities [AAC&U], 2002; Bok, 2006; Hersh & Merrow, 2005). Critics argue that higher education fails to focus on the skills that are most important for succeeding outside academia including critical thinking and real world problem solving. In addition, little effort is being made to ensure that students can transfer what they learn in one course to other courses or to the outside world. Indeed, grades and other assessments are often based on how students perform on short answer or multiple-choice questions that require little higher-order thinking or novel application (Bok, 2006).

Although service-learning may be able to assuage some of these indictments levied against higher education, it is often left out of the debate. For example, Bok (2006) discusses volunteer community service in his chapter on citizenship and civic engagement, but does not discuss academic service-learning in a chapter that includes critical thinking and problem solving. As a pedagogy, service-learning inherently teaches the kind of thinking skills and knowledge application necessary for success outside academia. Student products of service-learning, including comprehensive projects and analytic journals, require demonstration of critical thinking and problem solving skills in multiple contexts. In addition, service-learning often provides the right balance of challenge and support to foster intellectual growth and development (Ey-

ler & Giles, 1999). Assessing knowledge application, critical thinking/problem solving, and intellectual development outcomes of service-learning provides a platform for shifting the national conversation away from a focus on weaknesses in the U.S. education system toward a dialogue about how to build upon its strengths.

Given that the main goal of assessment is to improve learning, increasing formative assessment of service-learning will also provide feedback necessary to improve the quality of service-learning practice and thereby enhance student learning. Many faculty who include service-learning in their curricula already collect data that could be used for assessment (e.g., journals, projects, surveys, observations) but they do not systematically document their findings or collaborate with colleagues to assess outcomes across courses or curricula. More systematic examination of the data and increased discussion with colleagues about the results will inform improvements that can be shared with other service-learning practitioners and researchers.

Finally, increasing the involvement of service-learning practitioners in assessment will have the additional benefit of getting faculty members from a variety of institutions more professionally engaged in the scholarly service-learning assessment and research literature. As classroom, program, and college-wide assessments of service-learning are further developed, faculty can present their work at conferences related to assessment, service-learning, scholarship of teaching and learning, or disciplinary based teaching, thereby stimulating more service-learning researchers and research tools from the grass-roots level.

Tools to Assess Service-Learning

Assessment begins with the goals and objectives of the specific program being assessed (Maki, 2004; Walvoord, 2004). From these broader statements come more specific outcomes; finally, measures that can best address these outcomes are identified. Although indirect measures, such as attitudinal surveys, can supplement and provide context to assessment, direct measures of student learning must be included in any substantive assessment report. Measures must be meaningful to those who make decisions about the program so the results are actually used for improvement.

For many service-learning outcomes, faculty create their own assessments including locally developed rubrics. For other outcomes, however, tools developed for wider use as well as those developed specifically for service-learning research can be adapted for purposes of assessment. The remainder of this paper includes a review of tools that can be used to measure cognitive outcomes of service-learning. While service-learning clearly affects many other important outcomes such as civic engagement and ethical development, the focus of this review is on tools that assess cognitive outcomes as these are the main focus of student learning assessment. In addition, tools related to assessing program quality or the institutionalization of assessment, while important, also are not the focus of this article. This selection is not meant to be exhaustive but rather to provide the reader with a place to start. The tools discussed below are organized by format (research scales, written essays/protocols, interviews/qualitative approaches), and for assessment purposes each one can be aligned with one or more cognitive outcomes such as knowledge application, critical thinking and problem solving, and intellectual development (see Table 1). Inclusion of a tool in the table is not meant to be an endorsement of the tool. Readers will need to review relevant information on the reliability and validity of each instrument and to evaluate whether the resultant measures will address their outcomes.

Research Scales

One example of an instrument developed and tested specifically to assess cognitive outcomes of service-learning is the Cognitive Learning scale developed by the first author (Steinke & Fitch, 2003; Steinke, Fitch, Johnson, & Waldstein, 2002). This eight-item scale includes both a pretest version that asks about typical course requirements and a post-test version that asks about requirements of that course. Identical items ask about the degree to which requirements beyond participation in class and assigned readings addressed specific outcomes such as ability to spontaneously generate examples and application of course material to real world problems. While it is an indirect measure of learning because it relies on students' judgments of whether they have learned, the comparison between the pretest and post-test scores provides data beyond that from a typical indirect measure.

Bringle, Phillips, and Hudson (2004) have compiled information on research scales that were not originally created to assess service-learning but are relevant for common service-learning outcomes. The Problem-Solving Inventory assesses an individual's perceived problem solving skills; as such, it is an indirect measure. Bringle et al. include three direct measures specific to critical thinking: (a) the Watson-

Table 1
Cognitive Outcomes of Selected Assessment Tools

Tools	Cognitive outcomes		
	Knowledge application	Critical thinking and problem solving	Intellectual development
Direct Measures			
Articulated Learning (AL)	X	X	
California Critical Thinking Skills Test (CCTST)		X	
Cornell Critical Thinking Test (CCTT)		X	
Cognitive Level and Quality Writing Assessment Instrument (CLAQWA)	X	X	
Critical Thinking Rubric Direct		X	
Problem-Solving Analysis Protocol (P-SAP)	X	X	X
Problem-Solving Interview Protocol	X	X	X
Steps for Better Thinking		X	X
Watson-Glaser (WGCTA)		X	
Mixed Measures			
Learning Environment Preferences (LEP)			X
Measure of Intellectual Development (MID)			X
Measure of Epistemological Reflection (MER)			X
Perry Interview Mixed			X
Scale of Intellectual Development (SID)			X
Schommer Epistemological Questionnaire			X
Indirect Measures			
Cognitive Learning Scale	X		
Problem-Solving Inventory		X	

Glaser Critical Thinking Appraisal (WGCTA) provides a measure of critical thinking based on five subscales (i.e., Inference, Recognition of Assumptions, Deduction, Interpretation and Evaluation of Arguments; <http://harcourtassessment.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8191-013>); (b) the California Critical Thinking Skills Test (CCTST) provides a measure of five cognitive skill dimensions of critical thinking (i.e., Analysis, Evaluation, Inference, Deductive Reasoning, and Inductive Reasoning; <http://www.insightassessment.com/test-cctst.html>); and (c) the Cornell Critical Thinking Test (CCTT) provides a measure of six aspects of critical thinking (i.e., Induction, Deduction, Observation, Credibility, Assumptions and Meaning; http://www.criticalthinking.com/series/055/index_c.html). Finally, to measure intellectual development Bringle et al. include the Scale of Intellectual Development (SID) developed by Erwin (1983) to measure three of the four stages in Perry's (1968/1999) scheme of intellectual development (i.e., Dualism, Relativism and Commitment) plus an Empathy factor. The authors provide the source and description for each scale in their book, as well as information on reliability and validity.

Test (CCTT) (pp. 210-213) provides a measure of six aspects of critical thinking (i.e., Induction, Deduction, Observation, Credibility, Assumptions and Meaning); (http://www.criticalthinking.com/series/055/index_c.html). Finally, to measure intellectual development Bringle et al. include the Scale of Intellectual Development (SID) (pp. 205-207) developed by Erwin (1983) to measure three of the four stages in Perry's (1968/1999) scheme of intellectual development (i.e., Dualism, Relativism and Commitment) plus an Empathy factor. The authors provide the source and description for each scale in their book, as well as information on reliability and validity.

The intellectual development measures such as the SID introduce a new dimension to the direct versus indirect way of conceptualizing measures. The SID has students rate statements that represent how they think so, in this sense, it assesses students' beliefs and attitudes and could be classified as an indirect measure. However, the responses are not taken at face value. They are keyed to a larger theoretical framework of intellectual development and students never self-report on where they are in this framework, so it could also be classified as a direct measure. Another way to think about this is that it is a direct measure of

how students perform on a self-report task. Several of the measures discussed here fall into this category, which we will refer to as mixed measures.

Another checklist-type measure of Perry's intellectual development scheme that can be classified as a mixed measure is the Learning Environment Preferences (LEP) developed by Moore (1989) (<http://www.perrynetwork.org/recognitionformat.html>). This scale yields scores for Dualism, Early Multiplicity, Late Multiplicity, and Contextual Relativism, as well as a total Cognitive Complexity Index.

A mixed measure scale related to intellectual development is Schommer's (1990, 1998; Schommer, Crouse, & Rhodes, 1992) Epistemological Questionnaire. This multidimensional instrument assesses college students' beliefs about knowledge and learning on four dimensions: stability of knowledge, structure of knowledge, speed of learning, and ability to learn.

Written Essays and Protocols

The Problem-Solving Analysis Protocol (P-SAP) is a direct measure that uses open-ended problems to assess critical analysis and was developed specifically for service-learning (Steinke & Fitch, 2003). The rubric for writing was adapted from the original coding scheme for scoring interviews developed by Eyler and Giles (1999) based on the work of King and Kitchener (1994). The assessment begins with a prompt; an issue specific to a course is presented to students with a set of follow-up questions. The protocols are scored using two sets of rubrics, one for causal and solution complexity and one for causal and solution locus. The rubrics measure the use of critical thinking for problem analysis (<http://www.ncsu.edu/assessment/resources/p-sap.htm>).

Two essay-type mixed measure instruments are available to measure Perry's intellectual development scheme: the Measure of Intellectual Development (MID) (<http://www.perrynetwork.org/essayformat.html>) and the Measure of Epistemological Reflection (MER) (http://unixgen.muohio.edu/~magoldpm/faculty_faculty_intro_mbm_docs.html). Both present a series of questions and probes about students' conceptions of knowledge and learning and must be scored by trained raters. The MID offers pre- and posttest versions as well as an alternate essay about career planning.

The Cognitive Level and Quality Writing Assessment Instrument (CLAQWA) is a direct measure that was developed at the University of South Florida to help instructors assess the writing skills and cognitive skills of students (Flateby & Metzger, 1999; 2001). It includes two rubrics for scoring including a cognitive level skills scale based on Bloom's Taxonomy (Bloom, 1956). The rubrics can be applied to an essay assignment that is part of normal coursework so it can be easily applied to service-learning assignments (<http://usfweb2.usf.edu/eval/claqwa/>).

Steps for Better Thinking provides a model for teaching and assessing higher-order thinking skills including critical thinking and open-ended problem solving (<http://www.wolcottlynch.com>). It includes a simplified rubric designed primarily for feedback to students in classroom settings and a more complex and reliable rubric designed as a direct measure for research and formal assessments. This work began over ten years ago (Lynch, 1996; Lynch & Wolcott, 2001) and is grounded in King and Kitchener's (1994) reflective judgment model and Fischer's dynamic skill theory (Fischer, 1980). Dynamic skill theory describes the psychological underpinnings of the developmental sequence.

The Critical Thinking Rubric was developed at Washington State University to foster students' higher-order thinking skills and reform faculty practice. The rubric was designed to be integrated into courses and to provide a direct measure of critical thinking when scoring student essays so it can be easily applied to service-learning assignments. It continues to be developed and expanded (<http://wsuctproject.wsu.edu/ctr.htm>).

Interviews and Other Qualitative Tools

In addition to purely quantitative assessments or using well-developed rubrics to define or quantify various written works, other service-learning researchers have developed qualitative approaches to exploring and assessing cognitive outcomes. Eyler and Giles (1999, 2002) developed the Problem-Solving Interview Protocol for their research on outcomes of service-learning. This protocol questions students about the causes, solutions and strategies for action in response to a specific social problem both before and after students have encountered it in their service-learning experiences. It is based on the work of King and Kitchener (1994). As with the P-SAP discussed above, responses are scored for a direct measure of both locus and complexity.

Ash and Clayton (2004) have developed a reflection model to engage service-learning students in a deeper level of analysis based in part on Bloom's Taxonomy (Bloom, 1956). The product of the reflection process is an "articulated learning" (AL) in which students not only explain what they learned, but also how they learned it, why it was important and the ways in which it will be used for improvement. The AL is structured around specific learning objectives and results in written works useful for assessment. Ash and Clayton have adapted Paul's (1993) standards for a holistic scoring of the AL for purposes of direct assessment.

Perry's scheme of intellectual development can also be measured using a semi-structured interview (<http://www.perrynetwork.org/interviews.html>). Besides the standard set of questions, alternate protocols are available, such as clarifying convictions about competing opinions, as well as looking backward over the college experience and looking forward to future goals. This open-ended, mixed measure format provides rich data because it is flexible and allows for follow up on students' responses as well as adding questions to assess other outcomes related to intellectual development such as disciplinary perspectives.

Final Thoughts on Assessing Service-Learning

In this paper we have presented arguments for assessing service-learning and have provided some tools for faculty as they work on integrating service-learning into the assessment efforts on their campus. The focus of this paper has been on how to increase service-learning assessment with anticipated benefits ranging from improving service-learning outcomes to impacting the national dialogue on the quality of undergraduate education.

We conclude this paper by considering a more subtle benefit of assessing service-learning. Because of the goal-based, real world nature of this pedagogy, enhancing the quality of service-learning assessment can also provide a fresh perspective on the increasingly complex and often contentious assessment debates at colleges and universities across the country. The nature of service-learning often demands authentic assessments as faculty struggle to capture the real world transfer skills they believe are developing in their students. An increase in service-learning assessment may lead to a greater emphasis overall on assessments that better measure those skills and abilities needed for success outside academia. This possibility specifically addresses current critiques about the quality of higher education and its lack of relevance to real world demands (Association of American Colleges and Universities [AAC&U], 2002; Bok, 2006; Hersh & Merrow, 2005). Neither the push for standardized measures to produce results that are comparable across institutions nor the move for local outcomes-based, faculty-driven assessments has addressed the issue of assessing real world skills, including knowledge transfer, adequately. Perhaps an increase in service-learning assessment will bring this important issue to the foreground of the debate.

References

- Association of American Colleges and Universities (2002). *Greater expectations: A new vision for learning as a nation goes to college* (National Panel Report). Washington, DC: Author.
- Ash, S. L., & Clayton, P. H. (2004). The articulated learning: An approach to guided reflection and assessment. *Innovative Higher Education*, 29(2), 137-154.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives, handbook I: Cognitive domain*. New York: David McKay Company.
- Bok, D. (2006). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more*. Princeton, NJ: Princeton University Press.
- 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., Phillips, M. A., & Hudson, M. (2004). *The measure of service learning: Research scales to assess student experiences*. Washington, DC: American Psychological Association.
- Erwin, T. D. (1983). The scale of intellectual development: Measuring Perry's scheme. *Journal of College Student Personnel*, 24(1), 6-12.
- Eyler, J., & Giles, D. E. (1999). *Where's the learning in service-learning?* San Francisco: Jossey-Bass.
- Eyler, J., & Giles, D. E. (2001). *At a glance: What we know about the effects of service-learning on college students, faculty, institutions, and communities, 1993-2000*. Retrieved 5/17/07 from Campus Compact Web site: <http://www.compact.org/resources/downloads/aag.pdf>

- Eyler, J., & Giles, D. E. (2002). Using the problem solving interview to assess the impact of service-learning on understanding and critical thinking. In A. Furco & S. H. Billig (Eds.), *Advances in service-learning research: Vol. Service-learning: The essence of the pedagogy* (pp. 147-160). Greenwich, CT: Information Age Publishing.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477-531.
- Flateby T. L., & Metzger, E. (1999). Writing assessment instrument for higher order thinking skills. *Assessment Update*, 11(2), 6-7.
- Flateby, T. L., & Metzger, E. (2001). Instructional implications of the cognitive level and quality of writing assessment. *Assessment Update*, 13(1), 4-5.
- Hersh, R. H., & Merrow, J. (Eds.) (2005). *Declining by degree: Higher education at risk*. New York: Palgrave Macmillan.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment*. San Francisco: Jossey-Bass.
- Lynch, C. L. (1996). Facilitating and assessing unstructured problem solving. *Journal of College Reading and Learning*, 27, 16-27.
- Lynch, C. L., & Wolcott, S. K. (2001). *Helping your students develop critical thinking skills* (IDEA Paper #37). Manhattan, KS: The IDEA Center. Available at http://www.idea.ksu.edu/papers/Idea_Paper_37.pdf
- Maki, P. L. (2004). *Assessing for learning: Building a sustainable commitment across the institution*. Sterling, VA: Stylus.
- Moore, W. S. (1989, November). The *Learning Environment Preferences*: Exploring the construct validity of an objective measure of the Perry scheme of intellectual and ethical development. *Journal of College Student Development*, 30, 504-514.
- Paul, R. (1993). *Critical thinking: What every person needs to survive in a rapidly changing world*. Santa Rosa, CA: Foundation for Critical Thinking.
- Perry, W. G. (1999). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston. (Original work published in 1968)
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504.
- Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology*, 68(4), 551-562.
- Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, 84(4), 435-443.
- Steinke, P., & Fitch, P. (2003). Using written protocols to measure service-learning outcomes. *Advances in service-learning research: Vol 3. Research exploring context, participation, and impacts* (pp. 171-194). Greenwich, CT: Information Age Publishing.
- Steinke, P., Fitch, P., Johnson, C., & Waldstein, F. (2002). An interdisciplinary study of service-learning predictors and outcomes among college students. In A. Furco & S.H. Billig (Eds.), *Advances in service-learning research: Vol 2. Service-learning research through a multidisciplinary lens* (pp. 73-102). Greenwich, CT: Information Age Publishing.
- U.S. Department of Education (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, D.C.: Author.
- Walvoord, B.E. (2004). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. San Francisco: Jossey-Bass
- Zlotkowski, E. (1998). A new model of excellence. In E. Zlotkowski (Ed.), *Successful service-learning programs: New models of excellence in higher education*. Boston, MA: Anker Publishing.