Integrating Reflection and Assessment to Capture and Improve Student Learning

Sarah L. Ash, **Patti H. Clayton**, and **Maxine P. Atkinson**
North Carolina State University

Intentionally linking the assessment of student learning outcomes of service-learning with reflection allows each to inform and reinforce the other. This paper traces the evolution of a strategy that uses reflection products as data sources to assess and improve both individual student learning and program-wide approaches to reflection. Two tools were developed to guide the process of reflective writing in two courses. Associated rubrics were used to evaluate the quality of thinking demonstrated in the written reflection. Results suggest that these tools can improve students' higher order reasoning abilities and critical thinking skills relative to academic enhancement, civic engagement, and personal growth, and as a result, can improve the overall quality of their thinking and learning. However, this assessment has also surfaced the need for further improvement, particularly with respect to academic learning outcomes.

Much has been written about the benefits of service-learning to students, ranging from greater self-awareness, to the development of interpersonal and leadership skills, to enhanced preparedness for lives of civic and professional responsibility (Astin, Vogelgesang, Ikeda, & Yee, 2000). Increasingly, however, the focus has been on the need to document the academic and cognitive outcomes produced by the pedagogy (Eyler, 2000; Steinke & Buresh, 2002). Underlying the calls for more research in this particular area are concerns that failure to demonstrate significant intellectual gains will reduce service-learning's legitimacy in the eyes of both administrators—who are deciding how to allocate limited resources across programs—and educators— who are deciding whether the cost of additional time and effort is offset by student learning. As noted by Osborne, Hammerich, and Hensley (1998), "Non-academic benefits are unlikely to motivate many faculty who have not yet tried service-learning," and adoption of the pedagogy will require that "service-learning can be shown to impact the *learning of course content.*" (p. 6).

From our own experience as instructors, the development of critical thinking and meta-cognitive skills is an equally important outcome underlying our commitment to service-learning (SL) pedagogy. Following a call from the Wingspread Conference for research dedicated to identifying and assessing SL-related student learning outcomes (Giles, Honnet, & Migliore, 1991), numerous studies have sought to compare SL and non-SL courses, but the results have been mixed. Some have found improvements in grades (Markus, Howard, & King, 1993) while others have not (Kendrick, 1996; Miller, 1994), and still others have noted gains only in essay, not multiple choice, tests (Kendrick, 1996; Strage, 2000) and in written work (Osborne, Hammerich, & Hensley, 1998), suggesting that "type" of learning may be differentially affected. That is, service-learning may not improve the ability to recall facts over traditional classroom methods, but it may increase the ability to use evidence to support claims or to identify and solve complex problems. Finally, several studies have found that students at least *report* a better understanding of course material as compared to their peers in non-SL classes (Eyler & Giles, 1999; Steinke & Buresh, 2002).

Unfortunately, some of the findings of positive academic and cognitive outcomes can be challenged based on the nature of the research design. For example, students may have self-selected into a service-learning enhanced class, raising the possibility that they are qualitatively different from those in a more traditional classroom to whom they are being compared if random assignment is not part of the protocol. In fact, Sax and Astin (1997) found that service-learning students were more likely to spend more than 20 hours a week studying and doing homework as compared to their traditional classroom peers. The assignments on which grades are based may likewise be qualitatively different. In the study by Markus, Howard, and King (1993), although participants were randomly assigned to either a SL-enhanced or traditional section of the same course, students in the control sec-

tions were required to write a longer research paper to balance time spent in the community by their service-learning counterparts; thus, their lower final grades for the course could have reflected the greater difficulty associated with that requirement. Finally, there are significant limitations associated with the currently common practice of presenting evidence of learning in the form of data from student surveys with Likert-scale statements such as, "The service experience helped me to better understand course material." Students who say they agree with that statement have not said *what* they learned or provided any evidence in support of their claim to better understand the material. As Eyler (2000) suggested, self-report as an assessment strategy often confuses student *satisfaction* with student *learning*. She, therefore, has called for the development of mechanisms that support students in demonstrating concrete learning outcomes—that is, "that they have attained greater understanding, ability to apply their knowledge, problem-solving skills and cognitive development" (p. 11).

The difficulty in demonstrating expected outcomes, as evidenced by the inconsistent research, results at least in part because our approaches, particularly to reflection, do not always tap the full potential of the pedagogy. It is through careful reflection that service-learning—indeed any form of experiential education—generates meaningful learning; as Eyler, Giles, and Schmiede (1996) conclude from their review of approaches to service-learning, "It is critical reflection...that provides the transformative link between the action of *serving* and the ideas and understanding of *learning"* (p. 14). Eyler and Giles (1999) found that the more rigorous the reflection in service-learning, the better the learning, including academic outcomes: deeper understanding and better application of subject matter, increased complexity of problem and solution analysis, openness to new ideas, problem-solving, and critical thinking skills. Despite its centrality, however, quality reflection is perhaps the most challenging component of service-learning, stemming in part from the difficulty of developing and implementing both effective structures to guide it and meaningful strategies to evaluate and deepen its associated learning outcomes (Ash & Clayton, 2004; Rogers, 2001). Hatcher, Bringle, and Muthiah (2004) suggest that research into the relationship between reflection mechanisms and student learning can help put into practice guidelines for the design and implementation of effective reflection; their own guidelines suggest that explicit learning objectives, instructor feedback, and the use of assessment criteria (among other variables) are important influences on student learning outcomes in service-learning.

The authors have developed an integrated approach to reflection and the assessment of student learning that supports students in achieving and demonstrating academic and cognitive outcomes as well as outcomes with respect to personal growth and civic engagement (Ash & Clayton, 2004). A rigorous course-embedded assessment process can help to frame and support reflection, in turn producing stronger learning outcomes. A focus on classroom-based assessment itself as a way to continuously improve learning, rather than simply measure learning after the fact, is increasingly being recommended by those designing assessment strategies: "[F]ormulating assessment procedures for classroom use can spur the teacher to think more specifically about learning goals, thus leading to modification of (both) curriculum and instruction" (National Research Council, 2001, p. 229). In other words, an integrated approach to reflection and assessment can improve our ability to align the practice of service-learning with the theoretical claims of its learning potential. If, as a result of these efforts, we are more convincing in demonstrating significant outcomes, including academic and cognitive outcomes, then we will also have a compelling and effective way to promote, and ultimately expand, the use of this teaching and learning strategy.

We present here the results of a year-long investigation of the effectiveness of this model. This research has confirmed the potential of the model and given us a better understanding of how to refine it further. Of equal importance, however, it has led us to reconsider the fundamental objectives of research into service-learning's student learning outcomes. Perhaps the question ought not to be whether service-learning is "better" than other forms of classroom instruction but whether and how our practice is aligned with the student learning outcomes we believe service-learning is capable of producing. We suggest that as researchers we ought to be asking how we can use the challenge of demonstrating service-learning's effectiveness to facilitate our own careful and deliberate analysis of what it *can* do in theory as compared to what it *does* do in practice. The knowledge gained from such inquiry can then be used to improve the implementation of service-learning such that we more consistently fulfill its promise and actually produce the student learning outcomes we claim for it.

Background

Use of Reflection Products for Course-Embedded Assessment

From its inception, our service-learning program has been guided by a definition of the pedagogy that explicitly identifies the types of student learn-

ing outcomes at stake and the process through which reflection on service helps students to attain them. As we understand it:

> Service-learning is a collaborative teaching and learning strategy designed to promote academic enhancement, personal growth, and civic engagement. Students render meaningful service in community settings that present them with experiences related to academic material. Through guided reflection, students-individually and in groups-examine their experiences critically and articulate specific learning outcomes, thus enhancing the quality of their learning and of their service.

As is increasingly the case across higher education, our administration has for several years focused attention on assessment of student learning. Service-learning is one of several teaching and learning initiatives on The North Carolina State University campus that has emerged during this period of scholarly attention to the relationships among learning objectives, teaching strategies, and learning outcomes. The need to assess student learning outcomes at the program level has been an important driver in the refinement of our service-learning practice.

Encouraged to develop and implement program-wide assessment, we considered a range of possible mechanisms in light of the experience of other programs. We were disinclined to add a survey to end-of-the-semester course evaluations that would provide the data that is more suited to assessing satisfaction than learning, for the reasons described earlier. In addition, limited human and financial resources precluded the use of individual interviews or focus groups (Eyler & Giles, 1999; Gelmon, Holland, Driscoll, Spring, & Kerrigan, 2001) that have the potential to produce more rigorous results than student surveys alone. Our previous participation in the development of course-embedded approaches to assessment in the context of an inquiry-guided learning initiative called our attention to the value of using products generated by students in the course of the semester for assessment. We already had in place a program-wide approach to reflection (Ash & Clayton, 2004) that produced written products called Articulated Learnings (AL). These seemed to be a natural fit for a course-embedded approach to the assessment of learning outcomes, especially given the deliberately close correlation between the types of issues examined in the reflection process and the learning objectives of the program.

The AL is a series of paragraphs written as the culminating step of reflection sessions in each of three categories of learning objectives: academic, civic, and personal. These small group discussions are guided by a reflection framework, which is a series of questions designed to support students in *describing* (stage 1) and then *analyzing* (stage 2) their service experiences in such a way as to generate important learnings in each of the three categories. *Articulating learning* is the third stage in our reflection model; an AL is a vehicle through which students express and continue exploring important learnings that have surfaced through discussions regarding the content of the course (academic), their or others' participation in collective change-oriented processes (civic), and their personal strengths, weaknesses, assumptions, skills, etc. (personal). The AL is structured in accordance with four guiding questions: (a) What did I learn? (b) How, specifically, did I learn it? (c) Why does this learning matter, or why is it significant? and (d) In what ways will I use this learning, or what goals shall I set in accordance with what I have learned in order to improve myself, the quality of my learning, or the quality of my future experiences or service?

Thus, we are able to assess more directly what students actually have learned (because they have to articulate it in their written work), which in turn allows us to determine how closely that learning matches our expectations. This provides us with much more information regarding our program's effectiveness than, for example, the results of a survey from which we might only know what percentage of students think they learned a lot more than they would have in a traditional class.

The AL is designed to be a foundation for learners to carry the results of the reflection process forward beyond the immediate experience, improving the quality of future learning and experience (related to service or to other aspects of their lives). Students refine these reflection products through an iterative process of feedback, on their first draft from the reflection leader—a student trained to guide the reflection session process through which the ALs are produced—and then again after feedback from the instructor on their second drafts, to produce a final product. This process is repeated over the course of several reflection sessions throughout the semester. Instructors integrate the process of articulating learning into the assignment structure of their courses, such that the set of ALs each student produces serves as an alternative to an exam or essay or other assignment that would otherwise gauge particular learning outcomes.

Developing Tools to Assess Reflection Products—Formatively and Summatively

Having identified the ALs as the data source for program-wide student learning outcomes assessment, we began collecting them from SL instruc-

tors at the end of the semester, hoping to be able to qualitatively document learning outcomes across courses in all three categories of service-learning learning objectives. Reviewing these ALs, we quickly realized the disconnect between what the students were learning and what we had hoped they would learn. Taking the academic ALs as an example, we wanted to see deeper understanding of the complexities of course material in light of "real world" application, but many of the ALs demonstrated little more than the students' ability to recognize course concepts when they saw them. Two important realizations followed: a) that we needed a more precise understanding of the learning objectives than "academic enhancement," "civic engagement," and "personal growth" if we were to adequately assess learning in each of these categories, and b) that we needed to better support both students and faculty in achieving important outcomes in each category.

We therefore continued our process of developing program-wide assessment mechanisms by developing a set of learning objectives written specifically for each of the three categories but applicable to any course. This process was informed by our critical evaluation of actual ALs. For example, we read Academic ALs that expressed only statements of fact—as in "I learned that the elderly often do not get the nutrients that they need, even in an institutionalized setting with structured meals"—but that could have expressed deepened understandings of the complexities and subtleties of course concepts (in this case nutrition). For example:

> I learned that attempts to improve the poor nutrient intake that is often seen in the diets of the elderly in an institutionalized setting is complicated by the fact that for many of these residents the immediate pleasure of eating nutrient-poor cakes and cookies far outweighs, at least in their minds, some vague 'long-term risk,' especially since many tell me that they have lived this long eating this way and they aren't about to change. It makes me question whether we should even be telling these people what they should or should not eat.

The difference between the learning expressed in the statement actually produced and the learning that *could* have been expressed echoed the difference between lower and higher order reasoning and thus led us to Bloom's Taxonomy (1956) as a source of structure for the more precise statements of learning objectives we wanted to develop. We realized that the prerequisites for the academic outcomes we had envisioned—the more complex understanding of course material—included *identifying* a course concept when students saw it emerge in their experiences, *applying* the concept in the context of these experiences, *analyzing* the concept as presented in theory through comparing and contrasting its emergence in practice, and ultimately *evaluating* the adequacy of the theory. Clearly many could not reach the higher levels of thinking on their own; the learning objectives therefore were conceived as a way to make transparent and model the very cognitive process in which we were expecting our students to engage. The hierarchical learning objectives thus generated through this process are closely matched to the reflection framework that guides the students as they examine their service experiences from academic, civic, and personal perspectives in the reflection sessions; they are also annotated with a series of prompting questions to help the students use them in guiding their thinking. (See Appendix A for the academic learning objectives.)

Achieving and articulating this more precise understanding of the learning objectives was an important first step. However, we also realized that the process that takes a student from description of an experience to meaningful evaluation of that experience also requires the intellectual discipline of critical thinking. Critical thinking, as outlined by Paul (1993), is based on universal intellectual standards that include accuracy, clarity, relevance, depth, breath, logic, and significance. The oft-cited (Conrad & Hedin, 1990; Stanton, 1990; Strand, 1999) shortcomings of student reflection—reinforced stereotypes, interpretation based on unchallenged assumptions, inappropriate generalizations on the basis of limited data, shallow analysis that yields simplistic solutions to complex issues—appeared frequently in the ALs we reviewed and were clear examples of poorly developed critical thinking abilities. As a result, we recognized the need to provide guidance in this area as a necessary corollary to the hierarchical learning objectives. We therefore produced a second handout with definitions of Paul's standards of critical thinking along with sample AL passages that exemplify the absence of each to introduce students to the elements of critical thinking; this document is intended to support their effort to integrate each element into their thinking process and take their learning from the levels of identification and application, to the levels of analysis and evaluation.

Our assessment approach thus has an explicitly formative component; that is, the learning objectives and critical thinking guide are tools to be used by the students to help focus and deepen their reflective thinking. As noted by the National Research Council (2001), "Students will learn more if instruction and assessment are integrally

related. [P]roviding students with information about particular qualities of their work and what they can do to improve is crucial for maximizing learning" (p. 258). However, these same tools can also be used by instructors or program administrators in a summative fashion. An AL can be evaluated with respect to the highest level of learning objective it achieves, from a "1" indicating thinking at the level of "identify and describe only," to a "4" indicating "thinking encompassing all levels," from identification through evaluation. An AL can be given a "0" on this learning objectives rubric if it fails to follow the guidelines at all (e.g., if a student does not identify a course concept in an Academic AL). In addition, the critical thinking standards can be applied to the ALs in the form of a holistic rubric we developed that organizes the elements into four levels of mastery. (See Appendix B.)

Research Questions

Because the learning objectives and critical thinking guide are grounded in well-established learning theory, we believe that improvements in student scores on these written reflection products represent improvements in student learning. Therefore, as a concluding step in developing our assessment process, we decided to test their effectiveness by asking: (a) do the assessment tools improve the ALs *across drafts within a reflection session*, from first to final version? (b) do the assessment tools improve the *first drafts* of the ALs *over the course of the semester*, from early to later reflection sessions? and (c) are there differences in the degree to which students can achieve mastery among the three categories (Academic, Civic and Personal)?

Answers to the last two questions in particular should help to tease out specific areas of strengths and weaknesses in our ability to support our students' intellectual development and help to further refine course and program approaches to reflection specifically and to the integration of service-learning into our courses more generally. The study under discussion here spanned two semesters and is ongoing; it utilizes student ALs as the data source and generates scores for quantitative analysis through applying learning objectives and standards of critical thinking as rubrics. This study allows us to examine a wide range of learning outcomes associated with service-learning, not only specific outcomes in the Academic, Civic, and Personal categories but—transcending category—cognitive outcomes more generally.

Study Design

Articulated Learnings were collected from a random sample of two classes: a seminar on leadership in the 21st century with 9 students (4 selected) and a course on nutrition across the life cycle with 22 students (10 selected). In the first class, students helped residents at assisted living facilities learn to use computers, and encountered such course concepts as the nature of power in a technologically-intensive society and the various forms "community" takes in contemporary society. In the second class, students worked with community nutrition programs serving children, pregnant women, and older adults, encountering course concepts related to changes in nutrient needs and nutrition-related attitudes and behaviors across the life cycle. At the beginning of the semester, every student received a copy of our *Service-Learning Guidebook, Student Edition*, which includes introductory material on service-learning, the reflection framework itself and discussion of the reflection session process, and the learning objectives and critical thinking standards. Each of the two instructors partnered with a trained Reflection Leader, who met with the students in their service-learning project groups in out-of-class reflection sessions throughout the semester; the guidebook served as a tool to support their work together. The instructor and Reflection Leader both used the learning objectives and the standards of critical thinking to shape the feedback they provided students in the process of AL revision.

The students wrote one AL in each of the three categories (Academic, Civic, and Personal) following each of four reflection sessions. These "raw" ALs were reviewed by the Reflection Leader and returned to the students, who then reworked them to produce "revised" ALs submitted to the instructor for a second round of feedback. The students selected one AL in each category from the first two reflection sessions and, after revising once more following the instructor's feedback, turned in one "final" AL in each category. This selection and submission of three final ALs then repeated during the second half of the semester. Thus for each student there was a *raw*, *revised*, and *final* version of two ALs in each of the three dimensions: Academic, Civic, and Personal. Raw and revised ALs that were not finalized were not included in the analysis. Identifying information was removed from the ALs to maintain student anonymity and then they were randomly sorted. Trained student and faculty scorers independently—and blind to author, draft, and date so as not to prejudice their assessment—rated each AL based on the highest level of learning objective achieved (0-4) and the degree of critical thinking demonstrated (1-4). Scorers then came together at a series of group meetings to dis-

Table 1
Frequency of Learning Objective Scores Across Revisions
(Academic, Civic, and Personal ALs Combined)

	Score	Version
	Raw	Final
0	26% (22)*	15% (12)
1	4% (3)	2% (2)
2	58% (48)	35% (29)
3	11% (9)	43% (36)
4	1% (1)	5% (4)
Total	100% (83)	100% (83)

*Values in parentheses represent the number of ALs.

cuss and resolve discrepancies between their rankings. Through this process, a single consensus-based learning objective score was ultimately assigned to each AL by the group and differences in critical thinking scores was resolved to within one level. This was done to allow us to more carefully evaluate and improve on the learning objectives themselves so as to better communicate to the students how to achieve each level. Critical thinking scores were resolved to within one level as the demarcations between these levels is not as precise, nor are they linked to specific types of mastery. A Test for Observer Agreement identified the scorer whose critical thinking ratings were most likely to represent the majority, and those scores were used in the data being presented. ALs that received a learning objective score of 0 were not given a critical thinking score. There were three essays per student per category (raw, revised, final), in three categories (personal, academic, and civic), written twice during the semester (early and late). Therefore, 249 individual essays were read from the sample of 14 students. (One student did not finalize an Academic AL during the second half of the semester.)

Results

Table 1 presents the overall frequency of learning objective scores for the ALs across revisions in all three categories combined, from first draft ("raw") to third draft ("final"). (Data for second drafts have been omitted from all tables for clarity. We plan on analyzing the relative effects of instructor vs. Reflection Leader feedback in a future study.) There was definite improvement across revisions, as the proportion of students writing at levels 3 or 4 (analysis and evaluation) went from only 12% (11% +1%) on the first, or raw, drafts to 48% (43% + 5%) by the final versions. Table 2 shows the change in learning objective scores across revisions by learning category. In the Academic dimension, while 74% of the raw ALs were written at level 2 (application) and only 7% were written at level 3 (analysis), this improved to 48% at level 2 and 37% at level 3 by the final draft. Improvements were also seen in the Civic and Personal dimensions. In addition, the percentage of Civic ALs that did not meet the minimum criteria, level 0, dropped from 46% to 18%. In general, students appeared to have had a more difficult time making improvements in the Academic dimension. Only 37% of Academic ALs were written at levels 3 or 4, while 47% (36% + 11%) of the Civic and 61% (57% + 4%) of the Personal ALs achieved these scores.

Table 3 shows the frequency of learning objective scores from the first two reflection sessions relative to the final two, comparing the students' first drafts from early in the semester to their first drafts from later in the semester. Here the improvement was less marked as compared to that seen across revisions. Although the percentage of the raw ALs written at level 3 or 4 improved from 2% (2% + 0%) to 22% (20% + 2%), 24% (22% +2%) of the first draft ALs were still being written at levels 0 or 1 by the second half of the semester. Table 4 shows no significant trends when this same data from across the semester is broken down by category. Only a few students were able to write raw ALs at learning objective level 3 or 4 by the second half of the semester. Comparing all three learning categories, a greater percentage of Civic than Personal or Academic ALs still did not meet the minimum requirements, level 0, of the learning objectives, and no raw Academic or Personal ALs achieved level 4.

Table 5 presents the overall frequency of critical

Table 2
Frequency of Learning Objective Scores Across Revisions by Category

Score	Academic		Civic		Personal	
	Raw	Final	Raw	Final	Raw	Final
0	19% (5)*	15% (4)	46% (13)	18% (5)	14% (4)	11% (3)
1	0%	0%	7% (2)	7% (2)	4% (1)	0%
2	74% (20)	48% (13)	36% (10)	28% (8)	64% (18)	28% (8)
3	7% (2)	37% (10)	7% (2)	36% (10)	18% (5)	57% (16)
4	0%	0%	4% (1)	11% (3)	0%	4% (1)
Total	100% (27) †	100% (27)	100% (28)	100% (28)	100% (28)	100% (28)

*Values in parentheses represent number of ALs. † One student did not write an Academic AL in the second half of the semester.

Table 3
Frequency of Learning Objective Scores Across the Semester
(Academic, Civic, and Personal ALs Combined)

Score	Version and Time in the Semester	
	Early, Raw	Late, Raw
0	31% (13)*	22% (9)
1	5%(2)	2%(1)
2	62% (26)	54% (22)
3	2% (1)	20% (8)
4	0%	2% (1)
Total	100% (42)	100% (41)

*Values in parentheses represent the number of ALs.

thinking scores for the ALs across revisions in all three categories combined. A clear trend in improvement was seen, from only 22% of the ALs being written at levels 3 or 4 in the first drafts, to 69% (61% + 8%) being written at those levels by the final version. Table 6 shows that change by category. While only 18% of the Academic ALs were written at levels 3 or 4 in the first draft, 52% (48% + 4%) were written at those levels by the final version; Civic scores improved from 31% at levels 3 or 4 to 74% (65% + 9%); and the percentage of Personal AL scores at a 3 or 4 went from 21% to 80% (68% + 12%). Overall, the improvements were again somewhat greater in the Civic and Personal dimensions as compared to the Academic category.

Unlike the learning objective scores, the data in Table 7 indicate that students did improve on their ability to write a first draft relative to the critical thinking criteria over the course of the semester as the percentage of ALs with a score of 2 went down from 79% after the first half of the course to 52% after the second half, while those written at a level 3 went up from 0 to 42%. Table 8 shows that this trend was similar across categories.

Discussion

Our own informal comparison of ALs produced before and after developing these tools confirms that using learning objectives and critical thinking standards can improve the quality of student reflection and deepen student learning. And the analysis from this initial study also suggests the potential for deepening student thinking further over the course of the semester through their use. The study design is limited by the small number of students and use of only two instructors' classes; however, the classes did represent two very different disciplines (nutrition and leadership development). As we work with other faculty across campus on using this reflection and assessment model, we believe that the general pattern of results could be replicated in other disciplines as well.

Our first research question concerned whether the assessment tools could improve the ALs across drafts within a reflection session, from first to final version. With respect to both the learning objectives and critical thinking standards, we did see improvement in the scores, indicating improvement in the level and quality of thinking, across revisions. The second research question concerned whether the tools could improve the first drafts of the ALs over the course of the semester, that is, whether students could "internalize" them so as to produce higher quality ALs on their own as the semester progressed. The answer with respect to the critical thinking standards was yes—there was a clear shift upward by one level. The results for the learning objective scores, however, were disappointing. The students appeared to remain much more dependent on Reflection Leader and instructor feedback to refine their thinking.

Of course, expecting students largely unfamiliar with service-learning in general and this form of reflection in particular to be able to reason consistently at the highest levels after only 15 weeks is, perhaps, unrealistic, as Perry's (1970) now landmark work on intellectual development has made clear. We were also hampered by having only a few good models of strong ALs to show the students, given the relative newness of the tools themselves. In addition, our own reflection on the process has made clear that the feedback Reflection Leaders and instructors gave the students tended to focus more on critical thinking standards than on learning objectives, perhaps because the former tool can

Table 4
Frequency of Learning Objective Scores Across the Semester by Category

Score	Academic		Civic		Personal	
	Early, Raw	Late, Raw	Early, Raw	Late, Raw	Early, Raw	Late, Raw
0	14% (2)*	23% (3)	57% (8)	36% (5)	21 (3)	7% (1)
1	0%	0%	7% (1)	7% (1)	7% (1)	0%
2	86% (12)	62% (8)	36% (5)	36% (5)	64% (9)	64% (9)
3	0%	15% (2)	0%	14% (2)	7% (1)	29% (4)
4	0%	0%	0%	7% (1)	0%	0%
Total	100% (14)	100% (13) †	100% (14)	100% (14)	100% (14)	100% (14)

*Values in parentheses represent the number of ALs. † One student did not write an Academic AL in the second half of the semester.

Table 5
Frequency of Critical Thinking Scores Across Revisions
(Academic, Civic, and Personal ALs Combined)

Score	Version	
	Raw	Final
1	13% (8)*	1% (1)
2	65% (40)	30% (21)
3	22% (14)	61% (43)
4	0%	8% (6)
Total	100% (62) ‡	100% (71)

*Values in parentheses represent the number of ALs.
‡ ALs given a 0 on the learning objective rubric were not given a critical thinking score.

be more quickly and easily applied—with comments such as, "what do you mean by that," "why might this be true," and "does this follow from what you said earlier"—on a cursory, sentence-by-sentence read of an AL. Effective formative use of the learning objectives as a rubric, however, requires the reviewer to go back and evaluate the AL from a holistic perspective, considering how the learning might best be refocused or strengthened to take it to a higher level; its use seems to come a bit less naturally to Reflection Leaders and instructors. The learning objectives rubric takes more time and is thus apt to be the more easily neglected of the two tools. Improving the use of this rubric will be an important focus for future training efforts.

Our goal is to have students reflecting deeply and articulating quality learning as early in the semester as possible. Students bring fresh experiences to each class discussion and reflection session, and the process for improving their learning and their service is a cumulative one; so the longer it takes to develop solid reflection abilities, the greater risk of "wasting" opportunities. And of course we would like to see a substantial percentage of the ALs demonstrating thinking at the level of evaluation (a Learning Objective score of "4"). Driven by the suggestive but still unsatisfactory nature of our research results, we have developed a four-part tutorial that introduces students to the reflection model, process of articulating learning, and learning objective and critical thinking tools so that they are less dependent on us to help guide them. We plan to gauge the effectiveness of this tutorial in improving the overall quality of the ALs and in producing quality ALs earlier in the semester during the next phase of this research.

Our final research question concerned whether there would be differences in students' ability to achieve mastery among the three dimensions—Academic, Civic, and Personal. We found that, in fact, the Academic dimension posed the most substantial challenge, though it was more problematic for the learning objectives than it was for the critical thinking standards. Upon reflection we have come to realize that while all three categories of learning objectives require students to develop higher-order reasoning and critical thinking skills, the higher level academic learning objectives explicitly require students to critique course material, to look for gaps in the adequacy of a theory. While students seem able to reflect on service-related academic concepts utilizing at least some of the elements of critical thinking, it is clear that bringing specifically analytical and evaluative thought to bear on what they are learning in a course is one of the many ways in which service-learning is "counter-normative" to traditional learning (Howard, 1998), and therefore a reason why higher-order academic learning outcomes related to specific elements of course content may be harder to achieve. Before students can achieve such levels of reasoning they have to be given—and give themselves—permission to judge the work of established authors, and they have to come to believe that their own experience is a legitimate source of knowledge. Our students have helped us understand that to maximize this articulated learning process it is also necessary for them to learn to see writing as a vehicle for ongoing learning, rather than as a representation of learning that has already occurred—which is how they tend to approach academic writing especially. These "shifts in perspective and practice" (Clayton & Ash, 2004) are not easily made, so it follows that demonstration of the associated reasoning levels in the academic arena might lag behind. One of our conclusions from this pattern in the data is the need to provide better sup-

Table 6
Frequency of Critical Thinking Scores Across Revisions by Category

Score	Academic		Civic		Personal	
	Raw	Final	Raw	Final	Raw	Final
1	14% (3)*	0%	0%	4% (1)	21% (5)	0%
2	68% (15)	48% (11)	69% (11)	22% (5)	58% (14)	20% (5)
3	18% (4)	48% (11)	31% (5)	65% (15)	21% (5)	68% (17)
4	0%	4% (1)	0%	9% (2)	0%	12% (3)
Total	100% (22)	100% (23)	100% (16)	100% (23)	100% (24)	100% (25)

*Values in parentheses represent the number of ALs.

Table 7
Frequency of Critical Thinking Scores Across the Semester
(Academic, Civic, and Personal ALs Combined)

Score	Version and Time in the Semester	
	Early, Raw	Late, Raw
1	21% (6)*	6% (2)
2	79% (23)	52% (17)
3	0%	42% (14)
4	0%	0%
Total	100% (29)	100% (33)

*Values in parentheses represent the number of ALs.

port to our students as they learn to learn through service-learning; specifically, we are developing a range of reflection mechanisms designed to help students confront and adjust to the counter-normative nature of service-learning, so they are less hindered by its unfamiliar requirements than they might otherwise be and therefore able to make more progress more quickly, including reflection on academic material.

We began this research seeking to document program-wide student learning outcomes. We have done so, and, moreover, this research process has helped us to refine not only a program assessment strategy but also tools and materials that are serving to encourage if not "enforce" more consistent quality in the service-learning initiative on campus. Thus, the research has played a distinctly formative role in the evolution of our Service-Learning Program as a whole. Through the research process of reaching consensus on AL scores, we have surfaced several ambiguities in the learning objectives and identified specific ways to strengthen them to better support students' thinking in the AL process; the language of the learning objectives has thus been revised, producing the current version represented in Appendix A. The very existence of program-level objectives also gives us a structure around which to build faculty development. It allows us to provide instructors with much-needed guidance as they struggle to give meaningful feedback on their students' service-learning related work, interjects common language into our faculty learning community to support dialogue and scholarship, and helps to focus our understanding of service-learning pedagogy around a set of shared values (e.g., critical thinking). We believe that the tutorial will also make the reflection framework and its associated AL process easier to implement for otherwise reluctant faculty because of the support it provides them in reflection—an area with which they are often unfamiliar and/or uncomfortable.

We realize that there are at least two potential challenges to this process. One could argue that we are "teaching to the test" because we are so explicit about our expectations in the form of learning objectives. However, the "tests" are based on the well-established Bloom's (1956) *Taxonomy of Educational Objectives* that organizes the cognitive domain of learning into a series of levels that build on each other toward the development of intellectual skills, and on equally well-established standards of critical thinking that serve as a guide to the quality of the reasoning being used. A benefit of this transparent process is that it helps to develop both a common language in the classroom and an independent capacity for learning, which students can carry with them into other classes and other areas of their lives.

Second, it can be argued that this reflection model and the associated feedback process is simply too time-intensive to be useful to over-loaded faculty; it may also be suggested that the learning curve requires too many trade-offs of class or homework time and that the model implies inappropriately replacing other assignments with reflection assignments and thus devoting too large a portion of the course to reflection. We would never argue that this process does not take time, on the part of faculty and students alike, or that it does not require us to make changes in our use of time and assignments; it most certainly does, as, of course, does almost any approach to implementing quality service-learning. For us, however, and increasingly for the faculty we work with, time spent on deepening the learning outcomes associated with our teaching is time well-spent. As Howard (1998) suggests, whether we judge time as "away" from the "task at hand" depends on how we define our task. We believe that the power of service-

Table 8
Frequency of Critical Thinking Scores Across the Semester by Category.

Score	Academic		Civic		Personal	
	Early, Raw	Late, Raw	Early, Raw	Late, Raw	Early, Raw	Late, Raw
1	25% (3)*	0%	0%	0%	27% (3)	15% (2)
2	75% (9)	60% (6)	100% (6)	50% (5)	73% (8)	46% (6)
3	0%	40% (4)	0%	50% (5)	0%	39% (5)
4	0%	0%	0%	0%	0%	0%
Total	100% (12)	100% (10)	100% (6)	100% (10)	100% (11)	100% (13)

*Values in parentheses represent the number of ALs.

learning resides primarily in its ability to cultivate capacities for self-directed learning, personal growth, and citizenship, and that this potential can only be tapped through rigorous and high-quality reflection. This conviction helps guide us through the inevitable and admittedly difficult trade-offs that the implementation of this model requires.

Conclusion

We have presented a model for demonstrating the effectiveness of learning in service-learning courses that involves identifying desired student learning outcomes and then crafting an integrated reflection and assessment strategy around them. The following are equally important in this model: (a) the design of reflection mechanisms in accordance with the desired learning outcomes; (b) the recognition that more general cognitive outcomes (improvements in critical thinking and higher order reasoning skills) are tied to the more course- or program-specific learning outcomes; (c) the use of the same tools for formative and summative assessment; and, (d) the use of reflection products as a vehicle for assessment. It is our hope that our approach to integrating reflection and assessment can serve as a model for other faculty and staff as they seek to design service-learning courses and programs. As individual instructors and program administrators, we have found this approach most useful at both levels, in large part because—as any good assessment protocol should—it facilitates continuous improvement in our practice while also giving us data—both qualitative and quantitative—of interest to our institution. In particular, it is helping us to understand better both how our students think and how we can support them in learning to think more deeply and with greater capacity for self-directed learning. The research process helps us maximize the potential for student learning outcomes associated with service-learning.

References

Ash, S.L., & Clayton, P.H. (2004). The articulated learning: An approach to reflection and assessment. *Innovative Higher Education, 29,* 137-154.

Astin, A.A., Vogelgesang, L.J., Ikeda, E.K., & Yee, J.A. (2000). *How service learning affects students*. Los Angeles: University of California, Higher Education Research Institute.

Bloom, B.S. (Ed.). (1956). *Taxonomy of educational objectives, handbook I: Cognitive domain*. New York: David McKay Company.

Clayton, P.H. & Ash, S.L. (2004). Shifts in perspective: Capitalizing on the counter-normative nature of service-learning. *Michigan Journal of Community Service Learning, 11*(1), 59-70.

Conrad, D., & Hedin, D. (1990). Learning from service: Experience is the best teacher—or is it? In J. Kendall & Associates (Eds.), *Combining service and learning. I* (pp. 87-98). Raleigh, NC: National Society for Internships and Experiential Education.

Eyler, J. (2000). What do we most need to know about the impact of service-learning on student learning? *Michigan Journal of Community Service Learning*, [Special issue] *Fall*, 11-17.

Eyler, J., & Giles, D.E. (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*. Nashville, TN: Vanderbilt University.

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

Giles, D., Honnet, E. & Migliore, S. (Eds). (1991). *Research agenda for combining service and learning in the 1990s*. Raleigh, NC: National Society for Internships and Experiential Education.

Hatcher, J.A., Bringle, R.G., & Muthiah, R. (2004). Designing effective reflection: What matters to service-learning? *Michigan Journal of Community Service Learning, 11*(1), 38-46.

Howard, J. (1998). Academic service learning: A counter normative pedagogy. In R. Rhoads & J. Howard (Eds.), *Academic service learning: A pedagogy of action and reflection* (pp. 21-29). San Francisco: Jossey-Bass.

Kendrick, J.R. (1996). Outcomes of service-learning in an introduction to sociology course. *Michigan Journal of Community Service Learning, 3*, 72-81.

Markus, G.B., Howard, J.P., & King, D.C. (1993). Integrating community service and classroom instruction enhances learning: Results from an experiment. *Educational Evaluation and Policy Analysis, 15*(4), 410-419.

Miller, J. (1994). Linking traditional and service-learning courses: Outcome evaluations utilizing two pedagogically distinct models. *Michigan Journal of Community Service Learning, 1*, 29-36.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D.C.: National Academy Press.

Osborne, R.E., Hammerich, S., & Hensley, C. (1998). Student effects of service-learning: Tracking change across the semester. *Michigan Journal of Community Service Learning, 5*, 5-13.

Paul, R. (1993). C*ritical thinking: What every person needs to survive in a rapidly changing world.* Santa Rosa, CA: Foundation for Critical Thinking.

Perry, W.G. (1970). *Forms of intellectual and ethical development in the college years.* New York: Holt, Rinehart and Winston.

Rogers, R. (2001). Reflection in higher education: A concept analysis. *Innovative Higher Education, 26*, 37-57.

Sax, L.J., & Astin, A.A. (1997). The benefits of service: Evidence from undergraduates. *Educational Record, 78*, 25-32.

Stanton, T.K. (1990). Liberal arts, experiential learning and public service: Necessary ingredients for socially responsible undergraduate education. In J. Kendall & Associates (Eds.), *Combining service and learning. I* (pp. 175-189). Raleigh, NC: National Society for Internships and Experiential Education.

Steinke, P., Finch, P. Johnson, C., & Waldstein, F. (2002). An interdisciplinary study of service-learning predictors and outcomes among college students. In S.H. Billig & A.F. Furco (Eds.), *Service-learning through a multidisciplinary lens* (pp. 73-102). Greenwich, CT: Information Age Publishing.

Steinke, P., & Buresh, S. (2002). Cognitive outcomes of service-learning: Reviewing the past and glimpsing the future. *Michigan Journal of Community Service Learning, 8*(2), 5-14.

Strage, A.A. (2000). Service-learning: Enhancing student outcomes in a college-level lecture course. *Michigan Journal of Community Service Learning, 7*, 5-13.

Strand, K.J. (1999). Sociology and service-learning: A critical look. In J. Ostrow, G. Hesser & S. Enos (Eds.), *Cultivating the sociological imagination* (pp. 29-37). New York: American Association for Higher Education.

Authors

SARAH L. ASH received her doctorate in nutrition from Tufts University and is an associate professor in the Departments of Animal Science and Family and Consumer Sciences at North Carolina State University (NCSU). She has been active in the inquiry-guided-learning and general education reform initiatives at NCSU, and she serves as a senior Faculty Fellow with the Service-Learning Program, directing research and scholarship on reflection and assessment. Her special interests include the use of critical thinking standards and the reflective processes of service-learning to improve student learning.

PATTI H. CLAYTON received her interdisciplinary doctorate from the University of North Carolina, Chapel Hill. She is coordinator of the NCSU Service-Learning Program and an adjunct assistant professor in the College of Humanities and Social Sciences. Co-developer of the program's reflection framework, student leadership roles, and faculty development and support processes, she consults with other universities on service-learning capacity building and co-authors materials on reflection and assessment. She is particularly interested in leadership development through reflection and civic engagement.

MAXINE P. ATKINSON received her doctorate in sociology from Washington State University. She is an associate professor in the Department of Sociology and Anthropology at NCSU. She is currently the co-director of the First Year Inquiry Seminar Program. Atkinson's primary teaching interests are training graduate students to teach and teaching first year students using inquiry-guided methods. Her research interests focus on the conceptualization, measurement, and documentation of the scholarship of teaching and learning.

Appendix A: Academic Learning Objectives

Learning Objective (LO) Level	Academic Enhancement Learning Objectives	Associated Guiding Questions
LO 1: Identify and Describe	Identify and describe a specific academic concept that you now understand better as a result of reflection on your service-learning experience.	1.1 Identify an academic concept that relates to your service-learning experience. AND 1.2 Describe the academic concept that relates to your service-learning experience
LO 2: Apply	Apply the academic concept in the context of these experiences.	2.1 How does the academic concept apply to/emerge in your service-learning experience? (e.g., How did you or someone else use the material? When did you see it?)
LO 3: Analyze	Analyze the relationship between the academic material* (and/or your prior understanding of it) and the experience.	3.1 Compare and contrast the academic material and your experience: In what specific ways are the academic material (and/or your prior understanding of it) and the experience the same and in what specific ways are they different? AND 3.2 What are the possible reasons for the difference(s) between the material (and/or your prior understanding of it) and your experience? (e.g., bias/ assumptions/agendas/lack of information on the part of the author/ scientist or on your part.) AND 3.3 In light of this analysis, what complexities (subtleties, nuances, new dimensions) do you now see in the material that were not addressed or that you had not been aware of before?
LO 4: Evaluate	Evaluate the adequacy of the material (and/or your prior understanding of it) and develop a strategy for improved action.	Based on the analysis above: 4.1 How specifically might the material (and/or your prior understanding of it) need to be revised? AND 4.2 If applicable, what additional questions need to be answered and/or evidence gathered in order for you to make a more informed judgment regarding the adequacy/accuracy/appropriateness of the material (and/or your prior understanding of it)? AND 4.3 What should you and/or your service organization do differently in the future (or have done differently in the past) AND what are the associated benefits and risks/challenges?

*Note: "Academic material" includes the concept itself and its presentation (in class, in readings).

Appendix B: Level 4 of 4-Level Holistic Critical Thinking Rubric

Level 4 does most or all of the following:

Integration: Makes clear the connection(s) between the service experience and the learning being articulated

Relevance: Describes learning that is relevant to the AL category and keeps the discussion focused on the learning being articulated

Accuracy: Makes statements that are accurate and well-supported with evidence
(For academic ALs, accurately identifies, describes, and applies appropriate academic material)

Clarity: Consistently provides examples, illustrates points, defines terms, and/or expresses ideas in other ways

Makes very few or no typographical, spelling, and/or grammatical errors

Depth: Thoroughly addresses salient questions that arise from statements being made; avoids over-simplifying when making connections; considers the full complexity of the issue

Breadth: Gives meaningful consideration to alternative points of view and/or interpretations and makes good use of them in shaping the learning being articulated

Logic: Draws conclusions and/or sets goals that consistently follow very well from the line of reasoning presented

Significance: Draws important conclusions and/or sets meaningful goals that address the most significant issue(s) raised by the experience

Modified source: Paul, R. & Elder, L. (2001). *The miniature guide to critical thinking.* Santa Rose, CA: The Foundation for Critical Thinking.