Assessing Academic/Intellectual Skills in Keene State College’s Integrative Studies Program

Abstract

Key to the design of Keene State College’s Integrative Studies Program (ISP) was the intention to create a program based on a conceptual framework. The framework came from AAC&U’s Greater Expectations Report (2002) in which the essential learning outcomes for a liberal education were identified. In 2007 the program was implemented. The program has three sets of outcomes (intellectual skills, perspectives and interdisciplinary outcomes and integrative outcomes). The focus of this paper will be on the development of our process and on the assessment of intellectual skills.

Introduction

Keene State College is a founding member of the Council of Public Liberal Arts Colleges (COPLAC), and its mission is to ensure the goals of a liberal education are realized. Beginning in 2003, Keene State College embarked on another journey to revise its general education curriculum. In 2006, the new Integrative Studies Program (ISP) was approved. Key to the design of the program was its conceptual framework based on AAC&U’s Greater Expectations Report (2002) in which the essential learning outcomes for a liberal education were identified. Since then AAC&U has recognized the program as a LEAP (Liberal Education and America’s Promise) exemplar program. In 2007 the program was implemented. The program has three sets of outcomes (intellectual skills, perspectives and interdisciplinary outcomes and integrative outcomes). The focus of this paper is on the assessment of intellectual skills.

Over the last four years of the ISP writing, critical thinking, quantitative literacy and information literacy skills have been assessed using artifacts from the two foundation courses (Thinking and Writing and Quantitative Literacy), and writing, critical thinking, and quantitative reasoning skills using artifacts from our perspectives and interdisciplinary courses. Results are shared with faculty cohorts who are provided opportunities to discuss and address recommendations. The literature regarding curriculum development, pedagogical methodologies and assessment is replete with evidence that tells us we must clearly articulate expectations; design experiences and use methods that will help us meet those expectations; and use the evidence gathered to discuss whether or not what we are doing is working. The focus of this paper is on the assessment process developed to assess ISP outcomes and on initial findings of the assessment of intellectual skills (writing, critical thinking, information literacy, quantitative literacy and reasoning).
Review of Literature

A Liberal Education

AAC&U asserts that “all students need the knowledge, skills and capacities developed by a liberal education” (2006. p.6) and that “helping students master the arts of inquiry, analysis, and communication is the signature strength of a liberal arts education” (2007 p.18).

Most of us in higher education likely agree that an educated person has an understanding of disciplinary and interdisciplinary ways of knowing. There also seems to be growing agreement that “educated people should also possess a number of core proficiencies, in areas such as writing, quantitative reasoning, logical analysis, the use of computers [technology], and the ability to search out, evaluate, and integrate knowledge from many sources and contexts” (AAC&U 2001 p. ix).

Students determine what “counts” or “matters’ in higher education based on the messages they receive from us. What messages are students receiving about the broad skills and capacities they are expected to demonstrate in all the courses they are taking? How are students interpreting what a liberal education means? According to Humphreys (2006) “too often students associate the term ‘liberal education’ with the general education component of a liberal education rather than with a set of essential capacities developed across both general education and the major” (p.4). If certain skill proficiencies are the hallmark of a baccalaureate degree, then we must challenge students to develop those skills/proficiencies at high levels. “The skills must be developed for all college students in more than just the course they take to satisfy their ‘general education’ requirements. They must be further deepened and developed in a student’s major” (AAC&U, 2007 p.9).

In our third year of implementing outcomes based programs at Keene State College, we continue to struggle with our ability to ensure that essential skills and capacities are being developed in the ISP and in major courses. Early ISP assessment results suggest that much work needs to be done in helping students develop skills, and more discussion needs to occur related to the consistency between ISP skill expectations and the expectations in the majors. Both the majors and the ISP have a responsibility to ensure that the skills outcomes are transparent, transferable, and integrated.

Intellectual Skills

Many faculty in higher education have indicated for many years that they feel students are coming to college underprepared. Whatever knowledge and skills students are entering college with, the primary task of college faculty is to “move students’ skills in analysis and application to a much higher level” (AAC&U 2007 p. 31). It is our responsibility to identify our expectations and to find strategies that will move students to higher skill levels irrespective of the level at which they enter. College and university
learning in the United States is about higher learning, not just longer learning, and this applies to intellectual skill development as much as it does to gaining higher levels of knowledge and understanding.

There seems to be some resistance among those in higher education in using the word “skill” to describe the demonstrated ability of students to most fully engage with course content. Some have stated that skill development is not the responsibility of college faculty, others see skill development as remedial, and still others believe that students should enter college possessing the skills they need to do college level work. It is not difficult to understand these perspectives given that most faculty teaching in higher education have had little if any formal ‘training’ or “education” in teaching, let alone purposefully and intentionally helping college students develop higher order intellectual skills. As Bok (2006), indicates, most college faculty have not been formally ‘trained/educated’ to do one third of their job. However, “safely insulated from reliable evidence of how their students are progressing, most faculty members have happily succumbed to the Lake Wobegone effect. As surveys have confirmed, close to 90 percent of college professors consider their teaching above average” (Bok, 2006, p.316).

There is little argument that college faculty are experts in their fields and that they have been formally prepared to help students gain knowledge and develop understanding about specific disciplinary areas. However, “preoccupied with research and suspicious of anything ‘vocational’, Arts and Sciences departments have never made a serious effort to prepare Ph.D. candidates as teachers, even though most of their graduate students over the course of their careers will be primarily interested in their teaching and will spend more time at it than they devote to scholarship” (Bok, 2006 p.314).

Bok (2006) goes on to state that, “…colleges and universities, for all the benefits they bring, accomplish far less for their students than they should. Many seniors graduate without being able to write well enough to satisfy their employers. Many cannot reason clearly or perform competently in analyzing complex, non-technical problems, even though faculties rank critical thinking as the primary goal of a college education” (p.8).

AAC&U (2007) suggests that students will not be prepared “for the real-world demands of work, citizenship, and life in a complex and fast-changing society” (p.4) without having further developed their ability to read, write and speak effectively, think critically and creatively, reason quantitatively, and access information and use technology appropriately. The Leap National Leadership Council recommends an education that fosters “high-level intellectual and practical skills (AAC&U, 2007 p.4). However, the evidence that students are able to demonstrate higher level skills is sobering. Findings from Academic Profile, Educational Testing Service in 2003-2004 indicated that 5% of first year students and 8% of seniors were proficient at level 3 math; 11% of seniors were proficient at level 3 writing; and 6% of seniors were proficient in critical thinking” (AAC&U, 2007 p.8). Pascarella and Terenzini “found little consistent evidence that one’s major has more than a trivial net impact on one’s general level of intellectual or cognitive outcomes. Most of the progress in critical thinking – the skill so rightly prized by the
faculty- seems to take place during the first two years of college, before many students even start their concentrations (majors)” (p.140).

Justifiably, those in higher education have resisted and continue to resist external control in setting outcome expectations. This puts the burden on each institution and each academic program needing, then, “to set standards and guidelines for the expected level of student accomplishment” (AAC&U, 2007 p. 27). “In some instances, the judgments of professors on what skills are fit for teaching in a university will be influenced less by an enlightened sense of what their undergraduates need than by their own professional interests and priorities masquerading as sound educational principles” (Bok 2006, p 37).

What should college graduates know and be able to do? To date “there has been a near-total public and policy silence about what contemporary college graduates need to know and be able to do’ (AAC&U, 2007 p.1). But employers have not been silent. They have been telling us for some time that our graduates are not performing at a high skill level. “Seventy-three per cent of employers want colleges to place more emphasis on critical thinking and analytical reasoning and writing and oral communication. Seventy percent want more emphasis to be placed on information literacy and creativity and innovation; sixty-four percent on complex problem solving and sixty per cent on quantitative reasoning” (AAC&U, 2008 p.11). Today’s employers are ready to take responsibility for specific job training; however, they express concerns about the broader general skills of their new employees including “the lack of problem-solving skills and verbal and written communication skills” (AAC&U 2006 p.8). The 2008 National Survey findings (study conducted by AAC&U and Peter A. Hart Associates in 2007) note the following: “Employers reported the following percentages of recent college graduates being well prepared in various skill areas: writing – 26%; critical thinking – 22%, oral communication – 30%, quantitative reasoning – 32%.” (Kuh,2008. p.5).

The research being conducted makes a strong case for the need to become more focused on helping students develop college level intellectual skills. Huba and Freed (2000) state that “orchestrating stages in the skill development of students is also part of curriculum development” (p. 14). They pose a set of questions central to an assessment process that help frame a strategy for improving teaching and learning and which are: Where in the curriculum will students learn and practice skills like writing, speaking, teamwork, and problem solving? What teaching strategies will faculty use to help students develop these skills, and how will professors give feedback to students on their progress? Will all professors be responsible for these skills? Will the skills only be addressed in the general education component of the curriculum? Will some courses throughout the course of study be targeted as ‘intensives’ (p.14)?

**Assessment**

On most college and university campuses “faculty are not systematically engaged in asking powerful questions and getting answers to questions that impact and improve student learning. The fundamental assessment question, which should be built into the fabric of the curriculum and be the province of the faculty, is whether the faculty are
asking the questions and collecting the evidence that fosters constant improvement of teaching and the curriculum" (Lee, 1992 p.76). Committing to effective assessment requires faculty to “become researchers, posing central questions to better inform their sense of students' learning, their approach to teaching strategies, and the development of their own reflective habits” (Zessoules and Gardner, 1991, p.65).

One goal of assessment is to help students see the value in developing intellectual skills and in recognizing how they contribute to both academic and career success. As part of an effective education, students need to be able to “demonstrate what they have learned and what they can do as a result of their college experience” (AAC&U 2006 p.8).

"If assessment is to move from a 'minimal effort-minimal impact' mode, faculty must develop the understanding that assessment is not a foreign, hostile intrusion into academic life, but is rather a quintessentially scholarly activity. Assessment calls upon faculty to bring all of their professional skills and creativity to bear on how best to understand the impact of their work on students and how to increase the positive impact on students' intellectual and personal development" (Walvoord, 1998, p. 83).

Those faculty who are hesitant about the rigor of assessment procedures might consider Bok’s (2006) assertion that “the proper test for universities to apply is not whether their assessments meet the most rigorous scholarly standards but whether they can provide more reliable information than the hunches, random experiences, and personal opinions that currently guide most faculty decisions about educational issues” (p. 320).

“The overriding purpose of assessment is to understand how educational programs are working and to determine whether they are contributing to student growth and development. Hence the ultimate emphasis of assessment is on programs rather than on individual students” (Palomba and Banta 1999, p. 5). Faculty must move beyond student evaluations and individual student success or failure to determine whether students have meet programmatic expectations. We need more “systematic information about how much and how well students are actually learning” (Jacobi, Astin, Ayala, 1987, p. 17), and how well the curriculum is structured to advance faculty expectations for student learning.

Conducting programmatic assessment involves guidelines that, among others, Maki (2004) and Walvoord (2004) have identified in their work. As KSC embarked on identifying an assessment strategy for the ISP, we followed those guidelines: Identify outcomes, criteria and standards; identify in which courses outcomes are being addressed; determine methods to be used; determine who will be assessed and what will be assessed; establish a schedule; determine who will assess and interpret results; determine how and with whom information will be shared, and how changes have impacted future results.

In establishing an assessment timeline, advice from Allen (2006) was helpful. We realized that we did not have to “assess every general education outcome in every student every year, and that attempting to do so might lead to the trivialization or abandonment of efforts. A well-designed plan systematically addresses all outcomes in a multi-year cycle.
The plan specifies how a relevant, representative sample of students will be selected for each study, how data will be collected and analyzed, and how the campus will close the loop by identifying and integrating implications for change [action]” (p. 18).

Faculty, administrators and professional staff all play a role in creating effective assessment processes. The Cal State U experience with assessment in the 1980s "highlights the importance of faculty participation as a necessary condition for success. Findings also reinforce that assessment must be led. A key is keeping faculty informed and keeping assessment moving. Visible administrative commitment represents a crucial, and large nonsubstitutable, precondition for success” (Ewell, 1992, p. ii). It is unlikely the progress we have made at Keene State would be the same without administrative commitment and leadership, and without ongoing collaboration between faculty and administrators.

Keene State College’s Integrative Studies Program (ISP) Assessment Process

The Integrative Studies Program (ISP) has approved three sets of program outcomes (intellectual skills, perspectives and interdisciplinary outcomes and integrative outcomes). In the initial stages of programmatic assessment we have focused on assessing the intellectual skills outcomes. The Integrative Studies Program Council (ISPC) members, faculty facilitators and faculty teaching ISP courses have in the last three years identified and continue to identify and revise criteria that are used to assess all program outcomes. These constituencies make recommendations regarding the revision of outcomes and the adjustments that need to be made to the assessment process.

We have created a process that identifies the work which will be assessed, and the participants who will conduct assessments, write reports, and disseminate information. We have provided opportunities for faculty to identify student work that can be used to assess multiple outcomes (content and skills) and to discuss how changes will be made if changes are warranted based on assessment results. Our goal is to create an effective process that is as least labor intensive as possible while giving us useful information about what students are able to demonstrate about their learning.

ISP Assessment Committee Responsibility

All students in the program are told by faculty that they must submit their work (artifacts) to blackboard (Bb) sites for the purposes of programmatic assessment. The chair of the ISP Assessment Committee notifies faculty when and which blackboard (Bb) sites are available.

After submissions, student assignments are randomly selected for the purposes of program assessment. Faculty teams (three reviewers in each team) participate in a “norming” session facilitated by the chair to assure there is consistency in their assessments. After that session, each member of the team assesses twenty assignments (artifacts) using the program rubrics. Reviewers are paid a stipend for their work. One member of the team submits a report to the ISP Assessment Committee. That report is
shared with all faculty and results and recommendations are discussed among the faculty cohorts teaching in the program, facilitated by the area coordinator.

The submission of work is not voluntary but required. If work were received only from students who wished to submit their work, our findings would be skewed. The chair of the ISP Assessment Committee coordinates blackboard submissions and contacts all faculty each semester regarding the process used. Faculty are notified regarding who submits work so that they can address issues with students who do not submit work.

**Faculty Responsibility**

All course proposals have identified ISP outcomes addressed in the course. These outcomes need to be named on course syllabi and faculty should be clear with students about outcome expectations.

It is expected that faculty teaching perspectives and interdisciplinary courses will have one or more assignments that will address these learning outcomes (that will allow students the opportunity to demonstrate them). Faculty are encouraged to share program expectations with students and to incorporate as many outcomes in one assignment as is feasible. For example, one faculty has combined the writing and critical thinking outcomes students are expected to demonstrate in a single essay: “Outcomes being addressed in this second essay are: **writing** – using grammar and organization to effectively communicate my ideas; **critical thinking** - using credible evidence from our readings and discussion to support or refute ideas about poverty, and synthesizing ideas and information from our text, readings and class discussions to create a new understanding for me of what poverty is, who experiences it and what is needed to help alleviate it.”

Ultimately faculty need to identify **one assignment** (part of the course grade) to be submitted per outcome or per outcomes if there are two or more being addressed. Students should not submit multiple assignments for one outcome. Faculty are asked to have students include a cover page for the assignment with this information included: student name, student ID, title, and outcomes being addressed.

**Student Responsibility**

Students are required to submit, toward the end of the semester, course assignments that require them to demonstrate the intellectual skills being assessed. If one assignment is used to demonstrate two or more skills, students submit the same assignment to two or more sites. Students receive an e-mail from the chair of the ISP Assessment Committee with instructions for submitting their assignments.

**Intellectual Skills**

As indicated earlier, KSC has begun our assessment of ISP program outcomes by focusing on the assessment of intellectual skills. In higher education intellectual skills have not always been purposefully and intentionally developed, and when they have been
addressed with a sense of purposefulness and intentionality, coherence across a program and programs has not always been achieved. The Integrative Studies Program has identified eight skill sets (critical reading, writing, quantitative reasoning, critical thinking, creative thinking, critical dialogue, information literacy, and media and technological fluency) that need to be further developed across the eleven required courses in the program. The development of skill is different than gaining knowledge and understanding. Without skill development it is unlikely that students will be able to effectively convey what they know and understand, and without skill proficiency they are handicapped in their ability to most fully engage content. Skill development begins with a clear understanding of where students are and what needs to be done to help them develop skills at a higher level. Faculty must be clear about their expectations and must understand their role in helping students further develop intellectual skills. First steps involve identifying criteria that students are expected to meet for each of the identified program outcomes. We need to ask what our expectations are for performance in upper level courses and then we need to commit to building capacity in lower level courses.

The intent of initial assessments was to determine if the outcomes had been clearly articulated and were assessable and that there was appropriate inter-rater reliability. First round assessments typically resulted in revising outcomes and/or criteria. Norming sessions have been successful, resulting in high levels of inter-rater reliability.

Work in the two foundation courses, Integrative Thinking and Writing (ITW) and Integrative Quantitative Literacy (IQL), is assessed annually with work from both semesters being submitted to assure a representative sample. Work from IQL courses is assessed at the end of the fall semesters, and work in the ITW courses at the end of the spring semesters. Intellectual skills using work from the perspectives and interdisciplinary courses are assessed on a rotating basis at the end of each semester.

Assessing Foundation Courses in Keene State College’s Integrative Studies Program

The ISP facilitates an integrative teaching and learning process in which connections can be and are made. Integrative teaching and learning are more effective when purposeful, intentional and transparent. The program provides sustained opportunities to develop skills while engaging content, and to make connections across disciplines and beyond the classroom (applying learning, connecting theory and practice). To that end, the Thinking and Writing (ITW) and Quantitative Literacy (IQL) courses establish the foundation for reading, writing, information literacy, quantitative literacy, and critical thinking across the program. Outcomes for the ITW and IQL courses were created by faculty responsible for those courses.

Outcomes for the ITW course reflect students’ ability to: demonstrate skills and ways of thinking that are essential for all students as they move through the academic curriculum; and to write about an issue of special interest to them by focusing on a creative and complex question, investigating the question with critical analysis of readings, research and data, and using appropriate research techniques in documentation.
Outcomes for the quantitative literacy course reflect students’ ability to: apply the basic methods of descriptive statistics, including both pictorial representations and numerical summary measures, analyze data; use appropriate software to create spreadsheets, tables, graphs and charts; read and interpret visually represented data; distinguish among various types of growth models (e.g., linear, exponential) and the types of situations for which the models are appropriate; critically read and interpret a quantitative problem; and apply quantitative skills and concepts to describe, analyze and interpret real life data.

Students enrolled in the ITW courses are responsible for producing a research paper they have developed throughout the course. This paper will reflect and be used to assess the program outcomes for the course. Students enrolled in the IQL courses are responsible for producing a project that will reflect and be used to assess the program outcomes for the course.

Assessment (2007-2008)

In 2007, we began assessing the writing, critical thinking, quantitative literacy, and information literacy outcomes using artifacts from the two foundation courses - ITW and IQL.

Assessments occur at the end of the fall and spring semesters. Students submit their work to blackboard (Bb) sites. Random samples of student work are drawn. Students taking IQL courses also take a pre-test during the second week of classes and a post-test during the last week of classes. Faculty reviewers assess student work using rubrics created by faculty. The results are interpreted by the faculty reviewers and reported by them.

Norming

Prior to each round of artifact review, a norming session is conducted with reviewers in order to develop some consensus regarding the traits and the different levels of accomplishment defined by the rubrics used in assessment. During norming sessions, several randomly selected artifacts are reviewed and an anchor paper is identified and used as a benchmark in actual reviewing sessions.

Assessment of Writing and Critical Thinking Programmatic Outcomes Using Artifacts from the Integrative Thinking and Writing (ITW) Courses

Fall 2007

Procedures

A common practice in writing assessment is to calculate interrater reliability to measure consistency among raters. However, this practice has shortcomings. As Brookhart (2004) indicated, rater accuracy is a concern of all subjectively scored work. Having more than one rater to assess an artifact does not necessarily reduce subjectivity; even if they have
high agreement in their ratings. Further, intercorrelations between raters can show only consistency among the rank orders of the rated artifacts, but cannot tell the severity or leniency differences among raters (Bond & Fox, 2007). Moreover, in addition to consensus agreement among raters, consistency of a single rater is also a source of variations in raters’ judgment (Jonsson & Svingby, 2007; Moskal & Leydens, 2000).

To overcome the above mentioned shortcomings, since 2009, we have experimented with the Many Facets Rasch Model (MFRM). Many Facets Rasch Measurement (MFRM), also known as Facets Model, expands the original Rasch model by including other facets (e.g., rater severity) than person ability and item difficulty (Linacre & Wright, 2004). There are several advantages of using MFRM in conducting writing assessments. Assessing writing is labor intensive and costly if we have to rely on more than one rater to reach consensus with their ratings. Instead of having three raters review the same set of student writing artifacts, we developed a judging plan to make sure that any two of those three raters review only a few of the same artifacts. MFRM does not require every student be rated by every rater on every item as long as there is some network that links them together (Linacre & Wright, 2004). A certain level of objectivity about student writing performance can be modeled through MFRM, independent of raters. Last and most important, raters’ ways of interaction with the rating scale, in our case, rubrics, can also be modeled, which allows us to find ways to improve the rubric design and raters’ rating behavior.

An initial set of randomly selected artifacts from ITW sections were used to assist the raters in calibrating the rubrics. Then a random sample of 74 artifacts was selected for programmatic assessment from the total sample (N=512). 512 students out of the 681 (75%) enrolled in ITW courses submitted their work. Student ID numbers were the only identifying data included on the artifact. The writing skills outcome was operationalized as two interlocking components: writing using research (Table 1) and incorporating research into the written product (Table 2). Since the critical thinking rubrics tested during the pilot fell below acceptable levels of precision, two new rubrics were devised by this cohort of raters (see Tables 3 and 4).

Assessments were conducted using the rubrics as shown in Tables 1 -4. The rationale for using such a large sample was predicated upon the findings of the pilot study. The inter-rater reliability data for the pilot fell within acceptable limits but since the sample size was small, these estimates may have been unduly influenced by uncontrolled errors. In order to ascertain a deeper understanding of the efficacy of this rubric, a larger sample size was required for this iteration of program review.
Table 1. Skills Outcomes: writing using research

<table>
<thead>
<tr>
<th>Skill</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research</td>
<td>absent</td>
<td>1-2 in-text references from nonacademic, popular sites are present</td>
<td>3-4 in-text references, but no scholarly sources, are present</td>
<td>5 or more references no scholarly sources</td>
<td>Multiple in-text references; at least 1-2 are from scholarly sources</td>
<td>Multiple in-text references from 3 or more scholarly sources</td>
</tr>
</tbody>
</table>

Table 2. Skills outcome: incorporating research in writing

<table>
<thead>
<tr>
<th>Skill</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporate appropriately</td>
<td>absent</td>
<td>Sources present but cut-and-paste</td>
<td>Sources present with attempt to connect to topic</td>
<td>Sources present and mostly connected to topic</td>
<td>Integrated because most sources are controlled by writer</td>
<td>Highly integrated with all sources controlled by writer</td>
</tr>
</tbody>
</table>

Table 3. Skills Outcomes: Critical thinking (context)

<table>
<thead>
<tr>
<th>Critical Thinking Outcomes</th>
<th>Needs Improvement (1)</th>
<th>Met Expectation (2)</th>
<th>Exceeded Expectation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can the student examine a topic within a broader context (for example, where does this issue sit within a larger social, political, philosophical or historical framework?)</td>
<td>No attempt to examine the topic within a broader context.</td>
<td>Attempt to examine the topic within a broader context.</td>
<td>Demonstrates knowledge of the historical, cultural, or social forces that shape and are shaped by this issue</td>
</tr>
</tbody>
</table>

Table 4. Skills Outcomes: Critical thinking (perspective)

<table>
<thead>
<tr>
<th>Critical Thinking Outcomes</th>
<th>Needs Improvement (1)</th>
<th>Met Expectation (2)</th>
<th>Exceeded Expectation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can the student examine an issue, topic, or idea from more than one perspective (for example, reading not just those authors who support the writer’s position or viewpoint)</td>
<td>Examines the issue from one perspective only.</td>
<td>Attempt to examine an issue from more than one perspective.</td>
<td>Demonstrates broad knowledge of various positions on the issue and integrates them effectively</td>
</tr>
</tbody>
</table>
Results

Approximately 61% of the artifacts met or exceed the critical thinking-perspective expectation. Approximately 34% of the artifacts met or exceed the critical thinking-context perspective. However, these estimates reflect that the raters did not agree on the underlying nature of the latent variable. These data mimic the data derived during the pilot test. Articulating the latent variable(s) underlying critical thinking and accurately measuring them continues to be quite challenging.

Not all papers appeared to meet the course outcomes. A few of the papers contained no research and consisted mostly of personal stories. Several of the papers seemed to indicate that the students were paraphrasing without explicitly attributing a source. Despite the faculty development opportunities ITW faculty have been given, and despite a tight cohort, it was noted that there were still many discrepancies between papers. It was clear that while many faculty were assigning papers consistent with the course outcomes, others are not. Many of the papers did not have a thesis or claim and appeared to be “reports”—compilations of material not necessarily tied together in any meaningful way. Although the course outcomes clearly emphasized writing with purpose and stating and developing complex claims, many of the claims that students made were overly simple, indicating that ITW workshops may need to emphasize what it means to write about a topic in depth. Reviewers questioned the incentive for faculty who have fulfilled their “official” faculty development obligations to continue attending meetings and working on meeting course outcomes.

Recommendations

Reviewers reported there was a great deal of work to do before the writing and critical thinking outcomes could be effectively assessed. They indicated clearer outcomes were needed. Many skills outcomes were vague and hard to translate into clear and usable rubrics. For example, “write with an organizational schema” could be so open as to include any kind of organization. “Write with purpose” is another example—all writing has some purpose (instructions for taking medication, diary entries, and research papers) though not all purposes are appropriate to assess. The outcomes need to be meaningful to all who teach in the program before they can be accurately and adequately assessed.

The reviewers identified the need for clearer, more usable rubrics. While current writing rubrics were described as being effective and relatively easy to use, this was not true of the critical thinking rubric. It was determined that the critical thinking rubric needed to be revised. For example, instead of using the term “attempt” to describe what it means to meet “acceptable” levels of critical thinking competence, greater specificity was needed about what an “attempt” looks like.

Without a works cited or reference pages attached to all drafts of the artifacts, reviewers could not accurately assess students’ use of source materials. The ITW coordinator issued a reminder to all ITW faculty to this effect.
Training sessions were identified as necessary to improve inter-rater reliability.

Spring 2008

Procedures

There were 559 students enrolled in ITW courses during the Spring 2008 semester. Of those 471 submitted papers to the Blackboard (Bb) site (84.3%). A random sample of 60 artifacts was drawn and distributed to three reviewers. Although the original plan was to have all reviewers read 20 different papers from that set, reviewers conducted a norming session reviewing twenty papers for the purposes of increasing inter-rater reliability.

In order to ensure subject anonymity, student IDs were the only identifying data included on the artifact. The reference sections were included with the papers. In this assessment session, the writing skills outcome was operationalized as two components: writing using research and incorporating research into the written product (Table 5). Since the critical thinking rubrics tested during the assessment conducted in fall 2007 fell below acceptable levels of precision, the rubrics were refined and revised (Table 6) by this cohort of reviewers.

Table 5 Writing Outcomes: Use of Research

<table>
<thead>
<tr>
<th>Skill</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research References</td>
<td>absent</td>
<td>1-2 in-text references from nonacademic, popular sites are present</td>
<td>3-4 in-text references, but no scholarly sources, are present</td>
<td>5 or more references, no scholarly sources</td>
<td>Multiple in-text references; at least 1-2 are from scholarly sources</td>
<td>Multiple in-text references from 3 or more scholarly sources</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skill</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporate appropriately</td>
<td>absent</td>
<td>Sources present but cut-and-paste; no clear connection to argument</td>
<td>Sources present with some connection to argument</td>
<td>Sources present and mostly support argument</td>
<td>Integrated because most sources are controlled by writer and support argument</td>
<td>Highly integrated with all sources controlled by writer; sources clearly support argument</td>
</tr>
</tbody>
</table>
Table 6  Critical Thinking Outcomes

<table>
<thead>
<tr>
<th>Critical Thinking Outcomes</th>
<th>Needs Improvement (1)</th>
<th>Met Expectation (2)</th>
<th>Exceeded Expectation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can the student examine a topic within a broader context (for example, where does this issue sit within a larger social, political, philosophical or historical framework?)</td>
<td>No attempt to examine the topic within a broader context.</td>
<td>Examined the issue within a broader context, but may be done briefly, addressing one specific context (historical or sociopolitical, for example)</td>
<td>Demonstrates knowledge of numerous historical, cultural, or social forces that shape and are shaped by this issue.</td>
</tr>
<tr>
<td>Can the student examine an issue, topic, or idea from more than one perspective (for example, reading not just those authors who support the writer’s position or viewpoint)</td>
<td>Examines the issue from one perspective only.</td>
<td>Examines another side of or perspective on the issue; this is confined to one paragraph or section and may not be fully integrated into the argument as a whole.</td>
<td>Demonstrates broad knowledge of competing or conflicting positions on the issue and integrates them effectively into the argument.</td>
</tr>
</tbody>
</table>

While the original plan was to have each reader assess 20 artifacts from the sample of 60, it was decided that readers should review the same 20 to ensure inter-rater reliability, as the Fall 2007 assessment had inter-rater reliability data that fell below acceptable limits for three out of four outcomes.
Results

This study represented a replication of the previously conducted assessment of the two research rubrics. Findings for each outcome assessed are presented in Table 7.

Table 7 Summary of Findings

<table>
<thead>
<tr>
<th>Research References</th>
<th>Research Incorporate Sources Appropriately</th>
<th>Critical Thinking Broad Perspective</th>
<th>Critical Thinking More than One Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceed Expectation</td>
<td>10</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Meet Expectation</td>
<td>9</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Need Improvement</td>
<td>1</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

The reviewers indicated the more data collected, and assessed, the more confidence they could place in the accuracy of these measures of reliability. Artifacts assessed by the research reference rubric exceeded expectations. Thirty-five percent of the artifacts exceeded expectation on the ability to incorporate sources. In general terms, most artifacts assessed using the research rubrics met expectation or exceeded expectations (95% for the research reference rubric and the incorporate research appropriately rubric). Fifty-five percent of the artifacts assessed met or exceeded expectation for Critical thinking (broad perspective). A significant number of artifacts (45%) were rated as needing improvement. In terms of providing more than one perspective, most artifacts (60%) were rated as needing improvement. These data highlight the need for ongoing development and augmentation of student skills in the area of critical thinking.

All papers had a works cited or reference page, which assisted the readers in evaluating the source materials. Reviewers suggested that the first writing rubric contained too many items, and some seem random (why, for example, is five sources better than three or four if there are no scholarly sources present?) Others seemed to beg greater questions (students with no main point can sometimes “effectively integrate sources,” into their topic, but to what end?) There were also questions about what makes a source “scholarly” and whether or not “credible” might be a better term. The critical thinking rubric was generally effective, but some reviewers questioned whether the expectations for a “3” were too high for first-year students to reasonably achieve.

Troubling to reviewers was the lack of citations in some papers and the numbers of errors in the citations. Even among students who were able to integrate sources into their papers smoothly, sources were not cited properly (or even at all, in some cases). Several things stood out for the reviewers: 1) the number of times a student used in-text citations for the same source; 2) sources listed in the bibliography that are not referenced in-text; 3) the brevity of the URLs in web sources; 4) the lack of full referencing (author, title, journal, page) and 5) heavy reliance on unreliable websites, despite the presence of some scholarly resources.
Recommendations

Reviewers indicated that the writing and critical thinking rubrics must continue to be revised. The norming session was helpful in ensuring high inter-rater reliability, and reviewers recommended that these sessions be continued in future assessments. Reviewers appreciated having time to do the reading; having two months (rather than two weeks) to assess the essays cut down on “reader fatigue” and allowed for better focus. Reviewers indicated that students in their first year of the Integrative Studies Program were struggling with how to demonstrate an understanding of multiple perspectives on an issue.

Being able to locate, synthesize, and integrate conflicting perspectives is a higher-order skill that ITW is responsible for beginning to develop, and that upper-level ISP courses need to emphasize. Reviewers cautioned that this does not preclude these skills being developed more deliberately in ITW courses, however. Often students will not seek out (or not include) a piece of information that conflicts with what they are arguing in their papers because they think that it will weaken their argument. Students need to be encouraged (and required) to seek out—and take seriously—views that do not reflect their own. A new workbook created specifically for the course, Think, Write, Learn: A User’s Guide to Sustained Writing Projects contains material that helps students learn how to incorporate multiple perspectives into their writing projects.

Despite what these assessment results indicated, the problem of incorrect citation and documentation, coupled with an unprecedented number of plagiarism cases in ITW courses, warranted immediate action. Reviewers suggested that a strong effort needed to be made to understand why these problems exist, and to help faculty teach citation and documentation so that students can demonstrate these skills. Collaboration on all three fronts—ITW faculty, the library faculty, and the Center for Writing staff—was recommended as essential.

Reviewers suggested a need to consider how we assess what a student is attempting to do, in addition to how the student scores on the rubric. The intellectual work that ITW asks of students is much more rigorous than what they were asked to do in English 101 (College Essay Writing). However, a standard 101 paper (i.e. a “report” on the effort to legalize marijuana) might score higher on these rubrics than one of the papers in the sample (i.e. a discussion of child soldiers in Africa), even though the ITW papers aim for something much higher.

Missing from this assessment, as well, was a key part of what makes an academic paper successful—a main claim or thesis that is argued effectively and supported with evidence. Reviewers encouraged their colleagues to make sure that this happens, because it is what sets these papers apart from traditional “reports.”

A great deal of the growth in critical thinking and writing ability in the first semester is not demonstrated as a finished product but is evident in the sometimes monumental gains
that students make in understanding the complexity of the issues they are investigating. Sometimes this only begins to show up after fifteen weeks, and it can be evidenced at many times and places, including reflective writings that many ITW faculty have students complete throughout the semester that accompany their formal papers. Because ITW is just the first step in the program, reviewers indicated it is essential that we begin assessing work from upper-level ISP courses or else we would not get a truly accurate picture of whether students in the Integrative Studies Program are meeting the writing and critical thinking outcomes.

Reviewers reported that we need to find a way to assess writing and critical thinking that measures real intellectual growth. This is not to say that the assessment that we have undertaken tells us nothing; it does provide valuable information. However, it only tells us a small part of what we want and need to know about how this course and program impacts student learning.

Fall 2008

Procedures

There were 813 students enrolled in ITW courses during the fall 2008 semester. Of these 813 students, 659 writing artifacts (81.1%) were submitted to Blackboard (Bb). A random sample of 60 artifacts was selected to be assessed. Each of three reviewers read a randomly selected subset of 20 papers.

In order to ensure subject anonymity, student IDs were the only identifying data included on the artifact. Though reference sections were supposed to have been included with the papers, several students failed to submit them. For this assessment, the writing skills outcome was operationalized as three components: writing using research, incorporating research into the written product, and writing with grammatical and syntactical competence.

Results

Each of the three reviewers assessed a different set of twenty artifacts using the rubrics as shown in Table 8. The reviewers were the same as those from fall 2007 and spring 2008; all were trained and experienced assessors.
## Table 8

**ITW Fall 08 Artifact Grading Rubrics**

<table>
<thead>
<tr>
<th>Writing Outcomes</th>
<th>Needs Improvement (1)</th>
<th>Meets Expectation (2)</th>
<th>Exceeds Expectation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cultivate</strong> disciplinary and interdisciplinary expertise necessary to question sources, develop ideas, and offer interpretations</td>
<td>Citations absent, or in-text citations are largely from popular, nonacademic sources.</td>
<td>In-text citations reference three or more credible sources</td>
<td>In-text citations reference more than three credible or academic sources.</td>
</tr>
<tr>
<td><strong>Incorporate</strong> sources appropriately</td>
<td>No sources, or sources present, but cut-and paste</td>
<td>Sources present and mostly connected to position</td>
<td>Sources integrated into the argument smoothly and effectively</td>
</tr>
<tr>
<td><strong>Write with syntactical and grammatical competence</strong></td>
<td>Mechanics not controlled; multiple errors throughout paper.</td>
<td>Mechanics controlled; a few minor errors, but no stigmatized errors.</td>
<td>Mechanics highly controlled; few to no errors.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Critical Thinking Outcomes</th>
<th>Needs Improvement (1)</th>
<th>Met Expectation (2)</th>
<th>Exceeded Expectation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrate the ability to <strong>examine</strong> an issue within a broader context (for example, where does this issue fit within a larger social, political, philosophical, or historical framework?)</td>
<td>No attempt to examine the issue within a broader context.</td>
<td>Attempt to examine the topic within a broader context (i.e. it may be located to one paragraph or section and may not be fully integrated into the paper as a whole)</td>
<td>Demonstrates, throughout the paper, an awareness of the broader context(s) surrounding the issue.</td>
</tr>
<tr>
<td>Demonstrate the ability and willingness to <strong>approach</strong> a particular idea, problem, task, or goal from multiple perspectives.</td>
<td>Examines the issue from one perspective only.</td>
<td>Attempt to examine another side of or perspective on the issue; this is confined to one paragraph or section and may not be fully integrated into the argument.</td>
<td>Demonstrates knowledge of multiple perspectives on the issue and integrates them effectively.</td>
</tr>
</tbody>
</table>
Rubrics used had been revised since the spring 2008 assessment. Most student work appeared to meet and/or exceed expectations for both writing and critical thinking skills during this iteration of the assessment process. In terms of writing skills, approximately seventy percent of the artifacts assessed met or exceeded expectations for the writing rubrics. Most artifacts (70%) demonstrated effective use of research. Approximately eighty percent (80%) of artifacts met or exceeded expectations for incorporating research appropriately. Most artifacts (approximately 70%) displayed syntactical and grammatical competence. These findings reflected improvements in writing skills in comparison to previous assessment findings. Troubling to reviewers was the lack of citations in some papers and the numbers of errors in the citations. This was similar to findings in fall 2007 and spring 2008. The number of plagiarism cases in ITW decreased substantially in fall 2008, mostly as a result of increased awareness of this issue. However, reviewers indicated that students still need more instruction on citation methods. They also recommended more support for faculty teaching the course in the form of teaching materials, and clearer communication from the ITW coordinator.

In terms of the critical thinking skills, inter-rater reliability analyses were not conducted. Given the absence of this data, it is important to carefully interpret these results. Most artifacts met or exceeded expectations (80%) with regard to examining an issue within a broader context. Approximately sixty-five percent (65%) of the artifacts met or exceeded the outcome for examining an issue from multiple contexts. Reviewers suggested these tentative findings might represent an important shift in the conceptualization of critical thinking. Reviewers found the new rubrics easier to read and score and said that they reflect the course outcomes more clearly. They recommended continued analysis of the rubrics as essential.

**Recommendations**

Reviewers found it difficult reading lengthy electronic texts and suggested a need to find a way to provide paper copies. Realizing it might be expensive to copy and print papers, they suggested, none the less, that an alternative to reading papers on line needed to be discussed.

Reviewers indicated that too many students were still not submitting works cited or reference pages with their papers. They recommended that this issue needed to be communicated to faculty and to the students, especially those faculty who ask their students to hand in their reference pages as a separate document (which is a relatively common practice.)

Reviewers continued to support the need for norming sessions occurring before each round of assessment to ensure that the reviewers are consistent in their scoring.
Assessment of Information Literacy

Description of Project SAILS

In the fall of 2008, the library faculty chose Project SAILS (Standardized Assessment of Information Literacy Skills), a standard measurement test originated at Kent State University and sponsored by the ARL (Association of Research Libraries), as the tool that most closely met their needs as a pre-test to establish a benchmark before library information sessions are taught. Project SAILS is a knowledge test, consisting of 45 randomly-generated multiple choice questions that target a variety of information skills. Test items are based directly on two documents authored by the ACRL (Association of College & Research Libraries): (1) Information Literacy Competency Standards for Higher Education Performance Indicators, and Outcomes; and (2) Objectives for Information Literacy Instruction; A Model Statement for Academic Librarians. In those documents, each of five information literacy competency standards is expanded to include performance indicators, outcomes, and objectives. The SAILS test questions are derived from the outcomes and objectives (see p. 1 of SAILS report on The Test and how it is Scored for further information). The test, in existence since 2003, has gone through revisions and has been determined to be reliable and valid.¹

¹For more information, see https://www.projectsails.org/abouttest/validation.php.

Procedures

In the 2008 fall semester, the test was administered by the eight library faculty to 300 entry level students in various ITW classes prior to any library instruction in the fall of 2008. 292 were usable.

The ITW classes were selected because an information literacy component had already been incorporated into them and because the library faculty has an established working relationship with the majority of ITW instructors. Answer sheets were collected by the library faculty and forwarded to Kent State University for analysis. KSC’s results were compared to those of the selected cohort; individual test-takers were not scored. After all the institutions involved in the fall administration had submitted their tests a report was sent out in January of 2009.

The SAILS item bank has 157 items in American English. Each student answers 40 items from the item bank and 5 items that are in development. Test takers are instructed to select the one best answer. The items span the eight SAILS skill sets and the four ACRL standards targeted by the test. Students respond to different sets of items, with some common items shared across the individual tests.

The following are the eight skill sets that were tested: 1) developing a research strategy, 2) selecting finding tools, 3) searching, 4) using finding tool features, 5) retrieving
sources, 6) evaluating sources, 7) documenting sources, 8) understanding economic, legal, and social issues. To identify which skill sets were easier and which were more difficult, the test was administered to over 300 students with 292 being valid.

In addition to administering Project SAILS, the following outcomes were assessed in 2008-2009: 1) determine appropriate information by evaluating the relevance and usefulness of the information (based on ACRL Information Literacy Standard 3); 2) evaluate credibility of sources by applying criteria such as currency, authority and objectivity (based on ACRL Information Literacy Standard 3); and 3) cite completely all sources used (based on ACRL Information Literacy Standard 5). Library faculty examined a random sample (41) of ITW bibliographies.

Bibliographies

In order to assess the student learning outcomes that were identified in the library’s assessment plan, the library faculty examined a random sample of 41 ITW bibliographies. The library faculty developed a citation rubric (Table 9) to assess each bibliography. Bibliographies were divided among the faculty and individually assessed. Results were discussed, tabulated, and analyzed. For the first year, library faculty assessed the bibliographies under a weighted rubric. A weighted rubric is an analytic rubric in which certain concepts are judged more heavily than others. The library faculty used three objectives in the rubric to score each bibliography. 1) Applicable to project (title, argument, etc.); 2) Credible sources (scholarly and/or reliable sources; 3) Completeness (elements of citation). The second and third objectives were judged as the most important concepts, and the first objective as the least important concept. Using a point system the first objective was assigned the least points(I, 2 and 3 points) for the three criteria (does not meet the objectives, partially meets the objectives and meets the objectives) and the second and third objectives the most points (3, 6, and 9 points).

Results

PROJECT SAILS

Though the SAILS report goes into detailed results by each skill set for the purposes of this paper a summary of results is provided. Students performed about the same as the institution-type benchmark on using finding tool features. They performed worse than the institution-type benchmark on developing a research strategy; selecting finding tools; searching; retrieving sources; evaluating sources; documenting sources; and understanding economic, legal, and social issues.

When asked to identify which skill sets were easier and which were more difficult students identified the skill sets ordered by performance, from best to worst. Best: using finding tool features; evaluating sources; documenting sources; developing a research strategy; understanding economic, legal, and social issues; retrieving sources; and searching. Worst: selecting finding tools.
In summary the students did not do very well. Out of the eight skill sets students only performed” about the same” as similar institutions. It should be noted, however that the test was given before they had any library instruction sessions so the results were not unexpected. By administering the Project SAILS test, the librarians are able to measure students’ entry level skills, which then indicate areas of emphasis needed in library instruction. In AY 2011, the skills of third year students will be assessed and those results will be compared to results of the Project SAILS’ measure of first-year students’ entry level skills. Results will allow library faculty to ascertain the effectiveness of the instruction program, determine areas that need improvement, and modify instruction to meet goal of delivering high-quality instruction program.

Bibliographies

Results are summarized in Table 9.

A score between 18 and 21 indicated that the bibliography met all the objectives. 46% of the 41 bibliographies that were examined met all the objectives. 17% partially met the objectives (score between 13 and 17 points), and 37% did not meet the objectives (score between 7 and 12 points).

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Does Not Meet Objectives Criteria</th>
<th>Partially Meets Objectives Criteria</th>
<th>Meets Objectives Criteria</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicable to project (title, argument, etc.)</td>
<td>Fewer than 50% sources are applicable (1 point)</td>
<td>51% - 75% sources are applicable (2 points)</td>
<td>More than 75% sources are applicable (3 points)</td>
<td></td>
</tr>
<tr>
<td>Credible sources (scholarly and/or reliable sources)</td>
<td>Fewer than 50% sources are credible (3 points)</td>
<td>51% - 75% sources are credible (6 points)</td>
<td>More than 75% sources are credible (9 points)</td>
<td></td>
</tr>
<tr>
<td>Completeness (elements of citation)</td>
<td>Fewer than 50% sources have complete citations (3 points)</td>
<td>51% - 75% sources have complete citations (6 points)</td>
<td>More than 75% sources have complete citations (9 points)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rubric Score</td>
<td></td>
</tr>
</tbody>
</table>

7-12: poor (does not meet objectives)

13-17: average (partially meets objectives)

18-21: excellent (meets objectives)
Recommendations

Based on the citation analysis findings, the library faculty suggested that more emphasis need to be placed on explaining why citations are important in conducting and recording research, and on teaching students to evaluate the credibility and veracity of websites. Greater collaboration with the classroom ITW faculty is suggested particularly in examining student bibliographies as they are in progress, and in emphasizing the importance of the research process in collaboration with the library faculty.

IQL Assessment

Fall 2007 IQL 101 Pre-test and Post-test Results

There were 393 students who completed the IQL 101 pre-test and 297 students who completed the post-test in the fall of 2007. The results of student responses on the five attitudinal questions seemed to be positive in that students’ attitudes had changed slightly after taking a semester course on quantitative literacy. Students felt less confused, more confident and less insecure in their ability to do mathematics. The results of the fifteen quantitative skills questions that relate directly to the QL outcomes were not so favorable. Students’ scores from the pre-test to the post-test increased on seven of the questions, but decreased on eight of the questions. However, the change in scores was not significant enough to gain insight into the success the IQL 101 course had on the students’ abilities to meet the QL outcomes assessed on this testing instrument.

Fall 2007 IQL Student Projects

Procedures

There were 401 students enrolled in QL courses during the Fall 2007 semester. Of those, 283 (71%) submitted projects to the blackboard site. Fifty-seven artifacts were randomly selected to be assessed.

The projects were assessed using the rubrics included in Table 10 and checklist criteria in Table 11. The rubric was tailored to match the approved student learning outcomes to be addressed by students in the projects. As part of their QL faculty development, QL faculty members were informed of the learning outcomes to be assessed on the projects and were provided copies of the rubrics. In a very real sense, the rubrics help us to understand what is meant by each of the outcomes, and define what is meant for a student to meet or exceed expectations for the outcome and for the project.
<table>
<thead>
<tr>
<th></th>
<th>Needs Improvement</th>
<th>Met Expectation</th>
<th>Exceeded Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply the basic methods of</td>
<td>Pictorial representation or numerical summary measures not included, inappropriate,</td>
<td>Both an appropriate pictorial representation and an appropriate numerical summary</td>
<td>Appropriate pictorial representations and numerical summary measures are used and discussed effectively to support conclusions.</td>
</tr>
<tr>
<td>descriptive statistics,</td>
<td>or not used in a meaningful way. No discussion of the graphical or numerical</td>
<td>measure included and used in meaningful ways. Discussion of the graph(s) or</td>
<td></td>
</tr>
<tr>
<td>including both pictorial</td>
<td>representations included in the text.</td>
<td>numerical measures included.</td>
<td></td>
</tr>
<tr>
<td>representations and numerical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>summary measures, to analyze</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>data.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use appropriate software to</td>
<td>No or inappropriate tables, graphs, or charts. Tables, charts, or graphs</td>
<td>Software is used to tabulate numerical computations and to create an appropriately</td>
<td>Software used to tabulate numerical computations. Incorporates appropriately labeled</td>
</tr>
<tr>
<td>create spreadsheets, tables,</td>
<td>incorrectly labeled.</td>
<td>labeled graph or chart.</td>
<td>graphs or charts in a creative and effective way.</td>
</tr>
<tr>
<td>graphs, and charts.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read and interpret visually</td>
<td>Inaccurate or missing interpretation of visually represented data.</td>
<td>Accurate interpretation of visually represented data.</td>
<td>Accurate interpretation of visually represented data and clear, concise explanation of</td>
</tr>
<tr>
<td>represented data.</td>
<td></td>
<td></td>
<td>the interpretation.</td>
</tr>
<tr>
<td>Apply prior knowledge to</td>
<td>Weak or no connection to information learned in the course. Quantitative</td>
<td>Connection to information learned in the course. Quantitative language used</td>
<td>Extended information learned in the course. Used quantitative language correctly.</td>
</tr>
<tr>
<td>solve a new problem.</td>
<td>language used incorrectly. Weak or no synthesis of quantitative and contextual</td>
<td>correctly. Synthesized quantitative and contextual aspects of the course.</td>
<td>Synthesized quantitative and contextual aspects of the course.</td>
</tr>
<tr>
<td></td>
<td>aspects of the course.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11 Checklist Criteria

<table>
<thead>
<tr>
<th>Needs Improvement</th>
<th>Met Expectation</th>
<th>Exceeded Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contains many of the following</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pictorial representation or numerical summary measures not included, inappropriate, or not used in a meaningful way.</td>
<td>Both an appropriate pictorial representation and an appropriate numerical summary measure included and used in meaningful ways.</td>
<td>Appropriate pictorial representations and numerical summary measures are used and discussed effectively to support conclusions.</td>
</tr>
<tr>
<td>No discussion of the graphical or numerical representations included in the text.</td>
<td>Discussion of the graph(s) or numerical measures included.</td>
<td>Software used to tabulate numerical computations.</td>
</tr>
<tr>
<td>No or inappropriate tables, graphs, or charts.</td>
<td>Software is used to tabulate numerical computations and to create an appropriately labeled graph or chart.</td>
<td>Incorporates appropriately labeled graphs or charts in a creative and effective way.</td>
</tr>
<tr>
<td>Tables, charts, or graphs incorrectly labeled.</td>
<td>Accurate interpretation of visually represented data.</td>
<td>Accurate interpretation of visually represented data and clear, concise explanation of the interpretation.</td>
</tr>
<tr>
<td>Inaccurate or missing interpretation of visually represented data.</td>
<td>Connection to information learned in the course.</td>
<td>Extended information learned in the course.</td>
</tr>
<tr>
<td>Weak or no connection to information learned in the course.</td>
<td>Quantitative language used correctly.</td>
<td>Used quantitative language correctly.</td>
</tr>
<tr>
<td>Quantitative language used incorrectly.</td>
<td>Synthesized quantitative and contextual aspects of the course.</td>
<td>Synthesized quantitative and contextual aspects of the course.</td>
</tr>
<tr>
<td>Weak or no synthesis of quantitative and contextual aspects of the course.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results

Numerical results of the assessment of the respective learning outcomes are summarized in Table 12.

<table>
<thead>
<tr>
<th>TABLE 12  Assessment results</th>
<th>Outcome:</th>
<th>Needs improvement</th>
<th>Met expectation</th>
<th>Exceeded expectation</th>
<th>% Met/Exc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Apply the basic methods of descriptive statistics, including both pictorial representations and numerical summary measures, to analyze data.</td>
<td></td>
<td>34</td>
<td>17</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>2. Use appropriate software to create spreadsheets, tables, graphs, and charts.</td>
<td></td>
<td>27</td>
<td>24</td>
<td>0</td>
<td>47</td>
</tr>
<tr>
<td>3. Read and interpret visually represented data.</td>
<td></td>
<td>35</td>
<td>16</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>4. Apply prior knowledge to solve a new problem.</td>
<td></td>
<td>10</td>
<td>41</td>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>5. Overall project assessment</td>
<td></td>
<td>34</td>
<td>17</td>
<td>0</td>
<td>33</td>
</tr>
</tbody>
</table>

In this assessment, no student projects exceeded expectations. It is clear from the results that the projects met expectations at a high rate for one outcome but did not in the others.

Also clear is that most students met expectations at a reasonably high level (80%) on the outcome addressing the students’ ability to apply quantitative skills to the context of their course. The assessment also indicates that most of the projects did not address the other three outcomes at a level deemed sufficient to meet expectations as described in the rubric, if they were addressed at all. In general, two-thirds of the projects failed to meet expectations, for reasons described in the comments that follow.

Some project assignments did not require students to demonstrate proficiency in meeting the designated learning outcomes (Projects contained no graphs and/or very little statistical work with discussion; projects contained no analysis of statistics/graphs produced by software, or no evidence of software usage). Student analysis of graphs or numerical statistics was often weak or non-existent. In some cases, it was apparent that students can “do the math,” but have difficulty describing what the resulting numbers mean. Many projects had pages and pages of Excel output with no discussion, not even mention, of the reason why the spreadsheets were included. One student submitted a PowerPoint presentation, another an Excel file with only statistics and graphs—no explanations! Some student submissions were incomplete, as they mistakenly submitted a portion of the project and were subsequently unable to upload all they intended.

Not all students submitted projects to the Bb site (116 of 401 did not). This tended to skew the results. For example, the students of one faculty member who taught three sections of QL submitted projects, so those sections were over-represented in the sample.
Of the 283 that did submit, a simple random sample of 57 was chosen. This number was based on the assumption that the desired population parameter was a proportion and that the expected proportion was 0.7 (or 70%) of the projects would meet or exceed expectation.

Reviewers found that retrieving the projects from Bb is a time-consuming process—the first five took well over a half hour due to file conversions, numbering, multiple files per student, consulting the Bb administrator and the Helpdesk, dealing with corrupt student files, etc. Reviewers said it took a while to get into the rhythm, and overall it took hours.

Having little experience with grading from a rubric of this nature, it took some adjustment for one faculty member to evaluate using the rubric criteria rather than preconceived values. It took reading seven projects over 1½ hours before reviewers felt comfortable that they would evaluate consistently and similarly. Reviewers started by assessing several projects together, then switched to evaluating separately and comparing results.

**Recommendations**

A proposed additional learning objective for the foundations courses is that students learn to compile a coherent electronic document. That would require some administrative instruction (perhaps an online Bb course/tutorial) that would help students learn to cut-and-paste figures, graphs, spreadsheet output into a single Word document, or create one .pdf file from multiple sources. That is an important skill for life outside the college, as well as inside.

Faculty teaching the QL course should be well-versed in the QL outcomes and be mindful of those outcomes when designing project requirements and other course activities. Reviewers stated that though many faculty members have demonstrated they are very capable of designing assignments which effectively address the QL outcomes, more faculty development would help those who are not yet comfortable connecting, learning outcomes to pedagogical practices.

There is a wide range in the nature of the quantitative experience students are getting in a QL course—some are getting a full-blown introductory statistics experience, the rest are getting something less, and in some cases much less. This is connected to the recommendation above, in which QL faculty development helps instructors understand that students are to be challenged to meet all the QL outcomes and skills at some point in the course.

The assessment process could be administratively expedited if there was some quality control on the project submissions (ensuring that all submissions are in an uncorrupted format, increasing the number of students submitting projects, “mistake-proofing” upload instructions, etc.) and if the sample was determined and projects printed prior to faculty evaluators becoming involved.
In summary, the reviewers indicated that this assessment revealed more about our process than what students have learned in their QL courses, but they believed that was to be expected in a new program. The assessment was “doable” in a reasonable amount of time by faculty. The viability of the electronic submission process was validated, although there were bugs identified that needed to be eliminated. Including an adjunct in the assessment process was important as adjuncts are bearing a large load in QL instruction at KSC. Faculty have much to learn about including projects (in some sense, technical writing) as a learning vehicle in QL courses. Students must be challenged in QL courses to not only compute statistics and create graphs, but interpret what the specific statistics mean and what the graphs reveal. Continued faculty development and administrative support can close the loop to improve the current QL program.

**Spring 2008 IQL 101 Pre-test and Post-test Results**

There were 256 students who completed the IQL 101 pre-test and 180 students who completed the post-test in the spring of 2008. Once again, the results of student responses on the five attitudinal questions seemed to be positive in that students reported feeling less confused, more confident and more secure in their ability to do mathematics. The results of the fifteen quantitative skills questions that relate directly to the QL outcomes were more favorable in this assessment because a higher percentage of students answered 14 of the 15 questions correctly on the post-test than they did on the pre-test. Although there was improvement, the overall percentage of students answering questions correctly was low and reviewers indicated this was an area of concern that needed to be addressed.

**Spring 2008 QL Student Projects**

**Procedures**

Fifty projects were selected to be assessed. The projects were assessed using the same rubric as for the Fall 2007 semester. Results of the assessment of the respective learning outcomes are summarized in Table 13.
Results

Results are summarized in Table 13.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Needs improvement</th>
<th>Met expectation</th>
<th>Exceeded expectation</th>
<th>% Met/Exc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Apply the basic methods of descriptive statistics, including both pictorial representations and numerical summary measures, to analyze data.</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>2. Use appropriate software to create spreadsheets, tables, graphs, and charts.</td>
<td>16</td>
<td>34</td>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>3. Read and interpret visually represented data.</td>
<td>28</td>
<td>22</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>4. Apply prior knowledge to solve a new problem.</td>
<td>4</td>
<td>46</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>5. Overall project assessment</td>
<td>26</td>
<td>24</td>
<td>0</td>
<td>48</td>
</tr>
</tbody>
</table>

As was the case with results from fall 2007, no student projects exceeded expectations. It is clear from the results that the projects met expectations at a high rate for one outcome but did not in the others.

It is important to compare the Spring 2008 results with those of the Fall 2007 semester, which is done in Figure 1 below.

![Percentage of Students Meeting & Exceeding Expectations](image)

Once again, no student projects exceeded expectations. Reviewers noted that a greater percentage of students demonstrated achievement of the outcomes in all areas than did students in the fall semester. Reviewers attributed this primarily to the efforts of QL faculty to more clearly communicate project expectations to students, and to the efforts of
students to, in turn, perform to those standards. This was a very reassuring result in our effort to encourage faculty to use authentic assessment and for students to respond positively to the challenge. It remains clear that some project assignments did not require students to demonstrate proficiency in meeting the designated learning outcomes. Some projects contained no graphs and/or very little statistical work in the narratives, and some projects contained no analysis of statistics/graphs produced by software, or no evidence of software usage.

Student analysis of graphs or numerical statistics was often weak or non-existent. In some cases, it was apparent that students can “do the math”, but have difficulty describing what the resulting numbers mean. Many project submissions still had pages and pages of Excel output with no discussion, or even mention, of the included spreadsheets. Some students produced graphs with Excel, but they were poorly labeled to the point of not conveying the information the student hoped to communicate. The project submissions reviewers evaluated were principally just one file. That is a vast improvement. The exceptions were submissions that included a Word document and Excel spreadsheet.

The administration of the assessment went very smoothly. The calibration session was much shorter than last time, less than a half-hour, since the same evaluators conducted the assessment.

Reviewers noted that several projects were not written in narrative form, and appeared to be not much more than PowerPoint slides. Reviewers suggested that the rubric be modified to clearly state that a narrative report is expected, written in complete sentences and paragraph form in an appropriate report format. That correction was implemented and is addressed in the rubric and checklist included in the Appendices.

**Recommendations**

Students apparently were getting the message from faculty to learn to combine their various files into one document. That practice should be continued. Faculty teaching the QL course should continue to be mindful of the QL outcomes when designing project requirements. It is strongly recommended that the project rubric be shared with students prior to their starting the assignment.

Students continued to struggle with using quantitative information to support their arguments. In many projects, there was only a presentation of “the numbers” without much analysis or demonstration of where the numbers came from or their connection to the context of the course. It is suggested that faculty provide many opportunities for students to practice that skill, with both in- and out-of-class activities designed for that purpose. Providing students with examples of “how it is done” is an important modeling exercise to facilitate student learning.

All project submissions should include well-labeled graphs and tables that are clearly produced by a software package, such as Excel or SPSS. This is a clearly stated QL requirement that is not appearing in most of the submissions.
Any difficulties in collecting, sampling, or having the projects available were totally hidden from the reviewers, making their task simple. The time required for faculty to do the assessment was significantly reduced. Reviewers recommended that this administrative support should continue.

In summary, this assessment reveals that faculty are increasingly understanding the expectations of the QL outcomes assessment and that student submissions are improving. The assessment was readily accomplished in a reasonable amount of time by faculty. Including an adjunct professor in the assessment process continued to be important as adjuncts are bearing a large load in QL instruction at KSC. Faculty should continue to discuss how to include projects (in some sense, technical writing) as a learning vehicle in QL courses. Students must be challenged in QL courses to not only compute statistics and create well-labeled graphs, but interpret what the specific statistics mean and what the graphs reveal in the context of the specific QL course subject. Continued faculty development will continue to improve student learning in the QL program.

Spring 2009

In the 2009 spring semester a pilot assessment of the writing, critical thinking and quantitative reasoning outcomes was conducted using artifacts from perspectives and interdisciplinary courses.

Critical Thinking (Across the Program)

Procedures

577 students out of 1419 (40.7%) taking perspectives and interdisciplinary courses (identifying as addressing critical thinking) submitted artifacts. Sixty artifacts were randomly selected for assessment. Each of three reviewers was responsible for assessing twenty artifacts. All of the artifacts were rendered as Word documents; the authors’ names and course information were removed and student identification numbers were assigned to each to preserve anonymity and ensure unbiased review. Assignment information was not provided, nor was information as to which outcome(s) had been chosen by the instructor. While assessment was possible without the assignment and desired outcome(s), reviewers found that such information could have clarified numerous issues that arose during the norming session and actual assessment.

The norming session was conducted by the chair of the ISP Assessment Committee using ten artifacts from the Perspectives/Critical Thinking assessment sample. The session lasted two hours, in which time reviewers were able to assess individually each artifact and compare experiences with the artifact and the scores assigned. Scoring was mostly consistent among the three evaluators. The most significant issues faced came from the interpretation of certain terms used in the outcomes (Table 14) and the extent to which each outcome was indicative of critical thinking skills. Both issues were discussed in the reviewers’ final meeting.
<table>
<thead>
<tr>
<th>Table 14</th>
<th>Critical Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 Needs Improvement</td>
</tr>
<tr>
<td><strong>Outcome A:</strong> Uses credible evidence to support or refute an idea</td>
<td>Does not use credible evidence to support or refute an idea</td>
</tr>
<tr>
<td><strong>Outcome B:</strong> Incorporates multiple perspectives in examining an issue</td>
<td>Does not incorporate multiple perspectives</td>
</tr>
<tr>
<td><strong>Outcome C:</strong> Evaluates a source's use of evidence to support an idea</td>
<td>Does not accurately identify a source’s main idea, or does not adequately evaluate the source’s use of evidence</td>
</tr>
</tbody>
</table>
Results

Results are summarized in Table 15

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Needs improvement</th>
<th>Meets expectation</th>
<th>Exceeds expectation</th>
<th>% Met/Exc</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Uses credible evidence to support or refute an idea</td>
<td>23</td>
<td>28</td>
<td>9</td>
<td>62</td>
</tr>
<tr>
<td>B: Incorporates multiple perspectives in examining an issue</td>
<td>21</td>
<td>30</td>
<td>9</td>
<td>65</td>
</tr>
<tr>
<td>C: Evaluates a source’s use of evidence to support an idea</td>
<td>32</td>
<td>20</td>
<td>8</td>
<td>47</td>
</tr>
</tbody>
</table>

Regarding Outcome A (Uses credible evidence to support or refute an idea)

Reviewers reported that what students consider to be “credible evidence” ranged from unacceptable online dictionary definitions to thoughtful quotations from scholarly sources (books, journals, etc.). Many students seemed to interpret and/or limit the use of credible evidence to factual snippets that merely “flesh out” the narrative. Many of these facts/quotes were liberally inserted without citation, though bibliographies often appeared at the end of each artifact. In terms of writing, connections between the evidence presented and the overall issue/topic were generally weak. Reviewers were often left to make assumptions about the viability of the evidence and whether or how it was germane to the argument.

Regarding Outcome B (Incorporates multiple perspectives in examining an issue)

Reviewers reported that students’ frequent use of the unsubstantiated “I think…” was not enough to constitute critical thinking aptitude. They also determined that there was a clear disconnect for many students between the concept of personal opinion vs. informed perspective. Reviewers concluded that students seemed to believe that mere mention of what they consider an authoritative source constituted agreement and basis for a “perspective.” Rarely did students question or examine a source in a more meaningful way. Too many of the artifacts listed a number of sources that were all in agreement with each other; clearly the term “multiple” was understood to be “more than one” rather than “various” or “different” perspectives.

Regarding Outcome C (Evaluates a source’s use of evidence to support an idea)

Reviewers indicated that students who were able to meet or exceed expectations for this outcome were clearly of a different caliber than their colleagues. Reviewers envisioned these artifacts as having come from seniors, majors, or students with strong backgrounds in writing. The number of students who actually achieved a score of 2 or 3 for this
outcome was surprising, given the fact that so few artifacts seemed to have had this outcome as part of the assignment.

**Recommendations**

Reviewers agreed that faculty should design assignments that make explicit the critical thinking outcomes. They suggested a workshop/meeting of faculty teaching courses with critical thinking content would facilitate learning how to craft assignments in which students demonstrate their ability to meet these outcomes.

Students struggled with using critical thinking skills to construct and support their arguments. In many submissions, there was only a presentation of information without application, analysis or evaluation of content. Reviewers suggested that faculty provide many opportunities for students to practice that skill, with both in- and out-of-class activities designed for that purpose. Providing students with examples of “how it is done” is an important modeling exercise to facilitate student learning.

**Outcome A:**

From an assessment standpoint, Outcome A is somewhat problematic in that it does not allow for a range of quantification. Reviewers questioned whether “adequate” meant one connection? Two? When one strong connection was made between evidence and argument, reviewers opted for a score of 2, though they indicated that it should have merited something lower than that. Reviewers suggested clarification of the rubric with the following or similar quantifiers: “none,” “one to a few,” and “many.”

In the norming session and final meeting, reviewers found that their individual understanding of the term “credible” varied. Does information posted on websites or dictionary definitions constitute “credible” evidence? Are students taught how to discern the credibility of a source? Reviewers suggested that knowing the research expectations/parameters of the assignment would help evaluators.

In a similar vein, reviewers wondered whether “credible evidence” was too broad a concept and thus easily misconstrued. Can it be clarified in some way allowing for flexible language and thus disciplinary variety? For instance, consider the following:

- English: “Uses literary passages from a novel to support or refute an idea.”
- Biology: “Uses scientific data from a journal article to support or refute an idea.”

In any case, reviewers determined that the faculty should be able to clarify the outcome in some way to be inclusive of their disciplinary methods and sources. Such specificity would aid in the teaching, learning, and assessment of this outcome.

Reviewers were also in disagreement as to how to measure the extent of evidence “articulation.” They asked whether the introduction of a basic quote or statistic was enough to constitute an “adequate connection” though the writer does not go into any great depth. Owing to the vague nature of the rubric, reviewers often scored with a 2 when a score of 1 was merited. In addition to changing the language of the rubric, faculty
need to impress upon students the notion of “strong reading,” in which an idea/evidence is evaluated within, questioned through, or connected to the larger context of an argument.

Outcome B:

Outcome B should be refined to include or exclude the writer’s perspective. It may be unclear to someone who is assessing an artifact whether one should consider the writer’s perspective as one of the “multiple.” Reviewers felt it was not and assessed the artifacts with that in mind. Some type of clarification in the rubric – “in addition to the writer’s perspective” – may help. Or in some way, the notion of the writer’s perspective as part of the critical thinking process should be included, i.e. “The writer’s perspective is based on evidence taken from credible sources and includes mention and/or citation of these sources.”

Again, the use of a term – in this case “multiple” – was problematic. The reviewers questioned whether the mention (incorporation) of more than one perspective is enough to constitute critical thinking? Reviewers felt that critical thinking meant some type of analysis or conversation, and thus multiple perspectives needed to be put into some form of confrontation with one another. Reviewers discussed using terms such as “divergent” and “contradictory” that would ask students to create such a “conversation.” Perhaps the outcome should read, “Incorporates opposing perspectives while examining an issue.” Reviewers suggested considering the following:

1) Need Improvement: “Does not incorporate opposing perspectives”
2) Meets Expectation: “Incorporates opposing perspectives…”
3) Exceeds Expectation: “Incorporates and analyzes opposing perspectives…”

Again, allowing for flexibility in language may help here as well. “Opposing perspectives” could also read as “conflicting data,” etc.

Outcome C:

Outcome C is problematic because it assumes Outcome A meets or exceeds expectation – that there is in fact a source or evidence to evaluate. During the norming session, reviewers found that Outcome C also seems to be a higher order outcome more suited to upper-level courses. Though a few artifacts did attain an “exceeds expectation” score for this outcome, it was more “bottom-heavy” than Outcomes A and B.

Reviewers were also concerned that “identifying a main idea” was not a function of critical thinking and that “evaluating a source’s use of evidence” would be difficult for any undergraduate in a lower-level course. Again, reviewers considered analysis a stronger indicator of critical thinking skills.

Reviewers determined a more suitable rubric needed to be created. They questioned that if analysis stems from an inquiry, then why not assess the ability of the student to
formulate a question or suggest alternative modes of thinking based on the evidence presented? Consider the following: “Poses a substantive question based on varied evidence...” “Suggests a way to reconcile opposing viewpoints…”

Reviewers were dissatisfied with the critical thinking rubrics as they stand, both in terms of language and actual assessibility. However, they believed there is definite room for improvement in terms of how faculty teach and create assignments with the current outcomes in mind.

**Writing (Across the Program)**

**Procedures**

279 students out of 924 (30.2%) taking perspectives and interdisciplinary courses (identifying as addressing writing) submitted artifacts. Following the norming session a random sample of 59 projects was drawn from perspectives and interdisciplinary courses and made available to the three reviewers. Each of three reviewers was responsible for assessing twenty artifacts. All of the artifacts were rendered as Word documents; the authors’ names and course information were removed and student identification numbers were assigned to each to preserve anonymity and ensure unbiased review. Assignment information was not provided, nor was information as to which outcome(s) had been chosen by the instructor.

During the norming session, it was determined that it was impossible to assess Outcome #3 (Use revision effectively as part of the writing process.) since the submissions did not indicate whether multiple drafts of the written projects were required. Consequently, Outcomes 1, 2, 4, and 5 (Table 16) were assessed with results presented in Table 1.

**Table 16 Writing Outcomes**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Doesn't Meet</th>
<th>Meets</th>
<th>Exceeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 Develop complex perspectives, positions, and/or arguments.</td>
<td>Only the writer’s perspective provided. No clear evidence of any perspective, position, or argument.</td>
<td>Evidence of at least one other perspective, position, or argument.</td>
<td>Evidence of multiple perspectives, positions, or arguments.</td>
</tr>
<tr>
<td>#2 Support complex perspectives, positions, and/or arguments.</td>
<td>Little or no supporting evidence (i.e. studies, interviews, observations, examples, quotations,</td>
<td>Moderate / adequate evidence present.</td>
<td>Evidence of in-depth, detailed, supporting data.</td>
</tr>
</tbody>
</table>
A far greater proportion of students met or exceeded expectations for Outcomes 2, 4, and 5 than for Outcome 1. Some of this discrepancy can likely be attributed to the nature of the assignment. While the reviewers did not have descriptions of the assignments, it was fairly apparent in many cases that the students were only asked to provide their own interpretation or perspective of a reading or film, for example, and that they were likely not asked to provide more than one perspective, position, or argument. This assumption is supported by the differences in scores for the outcomes. For many of the writing artifacts, all other aspects of the students’ writing, for the most part, met or exceeded expectations, so it is likely that had they been asked to provide more than one perspective, position, or argument, they would have done so.
perspective, many of these students would have done so. Of course, the only way this supposition can be supported would be to ask for the original assignments.

The administration of the assessment went smoothly. Students were successful in compiling the various files into a coherent final file, with few exceptions (e.g. one file still had the student name and class included).

The calibration session was very important to the process of standardizing the use of the rubric for assessing the outcomes as well as for discussing how the rubric might be improved. The assessors came to an understanding of the rubric during the calibration session, and as a result the rubric was an effective tool for assessing the outcomes.

Some minor clarifications to the rubric were recommended. For example, in the wording for Outcome #4, “multiple errors” implies only a number. However, a student could repeat the same error over and over throughout a paper or a student could have a number of different types of errors in a paper. Reviewers chose to interpret that outcome using the latter description.

The reviewers questioned why only grammatical errors were included in Outcome #4. The ISP Council may want to consider additional types of errors that impact the clarity of ideas.

In summary, this assessment reveals that 80% or more of students in ISP Perspective courses, where faculty identified writing as an outcome, demonstrated that they met or exceeded 3 out of 5 writing outcomes including the ability to support complex perspectives, positions, and/or arguments; use grammar effectively to communicate ideas; and use organization effectively to communicate ideas. While only 42% demonstrated an ability to provide more than one perspective, reviewers hypothesize that this may be due to the nature of the assignments. Finally, it wasn’t possible to assess one of the outcomes (using revision effectively as part of the writing process) because only one draft of each assignment was submitted. Some minor revisions were suggested to the writing rubric.

**Recommendations**

Either faculty should be asked to design an assignment that makes explicit Writing Outcome #1 (requiring multiple perspectives) or a different, more universal outcome should be substituted to include assignments where only the student’s perspective is required.

While Outcome #3 (requiring multiple drafts) is certainly laudable, it is not readily assessable using the current method of artifact submission. There was no way for the reviewers to determine which assignments represented final drafts and which required no revision. Even if drafts were made available for review, assessing multiple drafts might prove too onerous for reviewers.
Quantitative Reasoning (Across the Program)

Procedures

108 students out of 503 (21.5%) taking perspectives and interdisciplinary courses (identifying as addressing quantitative reasoning) submitted artifacts. Seventy randomly chosen projects from perspectives and interdisciplinary courses were made available to the three reviewers. Each reviewer was responsible for assessing twenty-three artifacts. Student identification numbers were assigned to each to preserve anonymity and ensure unbiased review. Assignment information was not provided, nor was information as to which outcome(s) had been chosen by the instructor. Assessments were conducted using the following rubric (Table 18).

Table 18  Quantitative Reasoning Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Needs Improvement</th>
<th>Meets Expectations</th>
<th>Exceeded Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Does not use a quantitative problem solving strategy</td>
<td>Uses a quantitative problem solving strategy appropriately.</td>
<td>Demonstrates mastery of quantitative problem-solving strategies. Shows creative and innovative uses of quantitative tools such as statistical methods, mathematical and/or graphical representations.</td>
</tr>
<tr>
<td></td>
<td>appropriately. No or weak connection shown between</td>
<td>Successfully applies quantitative tools such as statistical methods, mathematical and/or graphical representations of real world cases. Uses quantitative language appropriately.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>quantitative tools and real world cases. Fails to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>appropriate quantitative language.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>Uses inappropriate pictorial, verbal, and/or numerical</td>
<td>Uses pictorial, verbal, and/or numerical representation appropriately to demonstrate an understanding of the data</td>
<td>Uses appropriate pictorial, verbal, and/or numerical representations. Analysis is particularly creative and effective or demonstrates unique insights about the data</td>
</tr>
<tr>
<td></td>
<td>representation and/or shows a misunderstanding or limited</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>understanding of the data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>(as appropriate for the assignment)</td>
<td>(as appropriate for the assignment)</td>
<td>(as appropriate for the assignment)</td>
</tr>
<tr>
<td></td>
<td>• Fails to suggest or incorrectly/inadequately suggests</td>
<td>• Suggests possible errors in data collection, analysis, interpretation, predictions or conclusions.</td>
<td>• Suggests corrections to possible errors in data collection, analysis, interpretation, predictions or conclusions.</td>
</tr>
<tr>
<td></td>
<td>possible errors in data collection, analysis, interpretation,</td>
<td>• Relates errors in methodology to errors in prediction/conclusion.</td>
<td>• Suggests corrections to the methodology or to prediction/conclusion.</td>
</tr>
<tr>
<td></td>
<td>predictions or conclusions.</td>
<td>• Identifies when choice of quantitative</td>
<td>• Suggests different</td>
</tr>
<tr>
<td></td>
<td>• Fails to relate errors in methodology to errors in</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>prediction/conclusion.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Fails to identify when</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
choice of quantitative information or presentation is used to emphasize a particular point of view.

information or presentation is used to emphasize a particular point of view.

presentation format to suggest an alternate point of view.

## Results

A summary of results is presented in Table 19.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Needs improvement</th>
<th>Met expectation</th>
<th>Exceeded expectation</th>
<th>% Met/Exc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Application</strong>. Develop appropriate quantitative problem solving strategies and apply them to real-world phenomena.</td>
<td>55</td>
<td>14</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>2. <strong>Analysis</strong>. Analyze, summarize, and interpret quantitative data and effectively communicate the findings.</td>
<td>60</td>
<td>19</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>3. <strong>Evaluation</strong>. Critically evaluate quantitative processes and results, which could include data collection, analysis, interpretation, predictions and conclusions.</td>
<td>51</td>
<td>3</td>
<td>6</td>
<td>13</td>
</tr>
</tbody>
</table>

Though the calibration session was not as effective as in QL norming sessions, the reviewers came to an understanding of the rubric during the session, and as a result the rubric was an effective tool for assessing the outcomes. What contributed to confusion was that the ten artifacts in the sample turned out to not represent the population of artifacts submitted, as all the artifacts in that sample were evaluated to indicate that all students needed improvement in all three outcomes. Reviewers acknowledged that this happens sometimes with random selection. Some projects evaluated subsequent to the calibration had more substantial QR content, and communication about those projects by the three reviewers would have been valuable. Communicating about those projects via email with reviewers’ disparate summer schedules proved ineffective.

That so few students met or exceeded expectations is problematic. In general, while many of the student submissions had quantitative information, few included quantitative methods to solve a problem or support an argument, few used any representation to describe data, and few evaluated the quantitative process or results.

In many submissions, students repeated quantitative information found in other sources. Reviewers found that most students are making an attempt to connect quantitative and contextual aspects of the course; however there was little evidence of application, analysis, and evaluation as described in the rubric. Reviewers concluded that students are not learning to use the quantitative skills to support an argument or make a case.
Reviewers found little or no evidence of carry-over of QL skills and outcomes. A general observation by reviewers was that most students demonstrate uniformly and disturbingly weak writing and quantitative reasoning skills. An obvious exception identified by reviewers was work submitted in one course. Submissions from that course were long (in excess of 60 pages) and comprehensive reports in which the students generated data and effectively applied, analyzed, and evaluated quantitative methods. Student success was achieved because of multiple drafts of the reports with feedback from peers and the instructors. Reviewers suggested that this could be a good model for emulation in other courses.

The administration of the assessment went very smoothly. Students were successful in compiling the various files into a coherent final file, with few exceptions. While it would have been ideal to have the assignment information, it was possible to do the assessment without those details.

In summary, reviewers suggested that this assessment revealed that faculty and students have room for improvement in demonstrating that they understand and can meet the QR outcomes. It is important to include an adjunct professor in the assessment process as adjuncts are bearing a large load in QR instruction. Faculty discussions involving the results of this assessment and the teaching of QR have potential to improve students’ demonstrating the desired QR abilities.

**Recommendations**

Reviewers suggested that faculty should design assignments that make explicit the QR outcomes. They thought a workshop/meeting of faculty teaching courses with QR content would facilitate learning how to craft assignments in which students demonstrate their ability to meet QR outcomes. Reviewers indicated that it is not appropriate for students to highlight sections of their completed assignment to indicate where they feel they addressed the outcome. As described in the rubric, the outcomes are clear and assessable.

Reviewers found that students struggle with using quantitative information to support their arguments. In many submissions, there was only a presentation of numerical information without application, analysis or evaluation of quantitative processes. Reviewers suggested that faculty provide many opportunities for students to practice that skill, with both in- and out-of-class activities designed for that purpose. Providing students with examples of “how it is done” is an important modeling exercise to facilitate student learning.

Multiple submissions with constructive peer/faculty feedback seems to be a key to success in students producing a report which meets/exceeds expectations as described in the rubric.
Finally, reviewers recommended that the assessment be completed in a shorter time window, to promote communication between assessors and preclude disparities in assessment ratings.

Discussion

“Over the past thirty-five years, state and federal policy makers, as well as the general public, have increasingly been pressuring higher education to account for student learning and to create a culture of evidence” (Shavelson 2007 p.1). At Keene State, we finally have evidence regarding programmatic outcomes. We have learned much over the last four years. Our assessment process is effective and is producing useful and reliable information. When this information has been discussed by faculty cohorts it has resulted in revisions to outcomes and criteria, and most importantly in discussions about improving program and course design and revising pedagogical strategies.

As a result of our assessment efforts, we have determined that there is much that needs to be addressed to improve student learning and that the quality of student learning will not improve to desirable levels without faculty commitment to ongoing instructional development. Improvement in student learning is less likely without this commitment. In outcomes based programs, alignment between assignments/experiences and program outcomes is necessary and it is challenging. In programs that have identified outcomes and in which assessment is occurring, faculty need to discuss findings, their teaching and their students’ learning, and they need to share the assignments and approaches they are using to improve learning.

Developing outcomes and effectively assessing them will not result in improvement in learning unless faculty are prepared to engage in ongoing discussions about curriculum development and effective pedagogical methods, and unless they are willing to participate in ongoing instructional development as part of their professional responsibilities. However, for us to move in this direction a significant change must occur in higher education. Faculty must see themselves as educators, not just scholars in disciplinary or interdisciplinary areas. They must be willing to learn some things about teaching and learning. As Bok (2006) states, higher education has been and continues to be out of step with the times. He says that “most successful organizations today, regardless of the work they do, are trying hard to become effective ‘learning organizations’ that engage in an ongoing process of improvement by constantly evaluating their performance, identifying problems, trying various remedies, measuring their success, discarding those that do not work, and incorporating those that do. Unfortunately, universities leave a lot to be desired when it comes to working systematically to improve their own performance” (p.316).

AAC&U (2007) purports that those in higher education will need to become “more intentional about essential learning outcomes and effective educational practices” (p.25). Bok (2006) suggests that faculty will have to move beyond past practices and experiment with new pedagogical methods. “Instructors have to change long-standing habits and master new skills for which many of them have little preparation. To avoid such difficult
ties, faculties have taken the principle of academic freedom and stretched it well beyond its original meaning to gain immunity from interference with how their courses should be taught” (Bok 2006. p.49).

Students will not perform as expected unless they have the opportunity to repeatedly practice and unless they get consistent and useful feedback. So, too, it is with pedagogy. We will not be successful when only a few faculty randomly participate in instructional development. To be more effective in helping students develop intellectual skills, faculty must see instructional development as integral to their teaching. Outcomes, instruction, and learning are intricately integrated.

We continue our journey at Keene State. We will continue to revise outcomes and refine our assessment process. We likely will continue to get useful and rich information that can lead to improved teaching and learning. We will continue to provide instructional development opportunities. We have rounded an important corner – we are assessing student work to help us determine program effectiveness. The next corner is equally challenging – will we continue to discuss the findings and determine what needs to be done to improve teaching and learning?
References


Association of American Colleges and Universities 2006. *Communicating Commitment to Liberal Education*. Washington, D.C.


Kent State University. *Project SAILS (Standardized Assessment of Information Literacy Skills)*. https://www.projectfails.org/abouttest/validation.php


*Significant material in the semester assessment reports were provided by Yi Gong, Chair of the Assessment Committee, Kirsti Sandy, Coordinator of ITW, Kathleen Halverson, Library faculty, Dick Jardine and Eileen Phillips, Coordinators of IQL, and Dick Jardine (Quantitative Reasoning), Susan Whittemore (Writing) and Stephen Lucey (Critical Thinking)*