

## **Developing and Field-Testing an Instrument for Assessing Information Competency Learning Outcomes of Community College Students: Final Report of the Bay Area Community Colleges Information Competency Assessment Project**

**September 2003**

### **EXECUTIVE SUMMARY**

The Bay Area Community Colleges Information Competency Assessment Project has developed and field-tested a criterion-referenced, two-part information competency assessment instrument based on specific performance outcomes derived from two national information literacy standards publications. Part A consists of 47 multiple choice, matching, and short answer items. Part B is composed of 12 performance-based exercises with sub-parts. The project's purpose was to develop a challenge-out or credit-by-exam instrument that might be used and/or modified at community colleges having an information competency requirement. Other uses of the instrument are possible, as long as librarians who use or modify the exam carefully read the Test Specifications document which explains in detail how exam items were developed. The exam was not sufficiently field-tested to establish validity and reliability measures.

### **BACKGROUND**

The Bay Area Community Colleges Information Competency Assessment Project evolved out of a number of activities underway at one of the San Francisco Bay Area community colleges, Diablo Valley College (DVC), and important statewide initiatives related to making information competency (IC) a graduation requirement for AA/AS degrees at all California community colleges. Diablo Valley College was among a small number of California community colleges awarded, in Fall 1999, three Student Success grants from the California Community Colleges' Chancellor's Office to integrate information competency into the college's instructional offerings and to promote it to other educational institutions. One of the three grants stressed collaboration between educational institutions at various levels with regard to information competency program planning. As part of this grant, DVC librarians sponsored three colloquia that brought members of what would become the IC test development project team together for stimulating presentations and discussions on issues related to information competency. During this time period, DVC and other Bay Area community colleges were also working with their campus curriculum committees and other governance groups to become informed about the national information literacy standards and to review and revise general education requirements with information competency outcomes in mind.

At the state level, librarians and other faculty and administrators in the California Community Colleges have been active advocating for information competency for over ten years. In 1996 the State Chancellor's Office of the California Community Colleges identified information competency as a system-wide priority. Between 1996 and 1999, several reports and resolutions discussed implementation. In April 2001 the Statewide Academic Senate's Curriculum Committee passed a resolution endorsing information competency as a "locally designed graduation requirement." A Chancellor's Office Task Force prepared language, which was nearly approved by the state Board of Governors, to revise *Title 5*, the Education Code, of the *California Code of Regulations*. However, the California Department of Finance determined that the fiscal impact might be difficult to fund, so the initiative has been put on hold.<sup>1</sup>

All of these activities provided both a ripe environment for librarians to share information and strategies about their initiatives and also identify a common need – a proficiency exam that could be used for students to satisfy the anticipated graduation requirement. Therefore, several meetings were held of interested Bay Area community college librarians and librarians from the California State Universities

---

<sup>1</sup> For detailed information about the chronology of these initiatives and text of these various reports and resolutions, see <http://www.topsy.org/infocomp.html>

(CSUs) to share information and to discuss approaches for building on the information competency momentum that was underway locally, regionally, statewide and nationally. Discussions centered around how to define the information competency skills as measurable outcomes; strategies for encouraging acceptance among other faculty and administrators; and effective teaching strategies and materials. A librarian from City College of San Francisco encouraged the formation of a working group to explore assessing information competency. As this issue was explored, the need and mutual interests pointed to the desire to develop a challenge-out test for the information competency requirement since several librarians were planning for implementation of a campus information competency requirement. Librarians from five institutions, City College of San Francisco, Diablo Valley College, Cabrillo College, Santa Rosa Jr. College and CSU-Monterey Bay, formed the initial project team, with librarians from several other community colleges and neighboring CSU universities serving as observers and contributors. This group did not define its charge to reflect the needs or situations of all California community colleges; instead, to maintain a manageable project, the scope was limited to developing and field-testing an instrument for use as a challenge-out exam for interested Bay Area community colleges with the understanding that the final instrument would be of interest to and adaptable by others.

The project team members<sup>2</sup> outlined a project plan with a draft timeline. Since the purpose of developing a challenge-out/credit-by-exam test had already been established, the next step was to determine the type of proficiency exam. The team decided to work on a criterion-referenced, two-part exam, composed of a cognitive test evaluating what students know, and a performance-based part assessing what students can do. Aware of the need to have a reliable and valid exam, the project team applied for and was awarded a California Academic and Research Libraries research grant for \$2,000 for the testing consultant's fees. Much later in the project an additional \$600 was provided from part of a Student Success Grant by Diablo Valley College to complete payment to the project consultant. The project team members accomplished many of the goals of the project by working independently and in two sub-groups. Numerous e-mail and conference call discussions occurred throughout the three-year project. Several half or full-day meetings, usually during the summer months, were also held. This loose collaboration proved to be effective, but only because of the dedication and commitment to the project goals of the team members who donated their "free" time to accomplish them. The supporting documentation outlined at the Web site <<http://www.topsy.org/ICAP/ICAPProject.html>> reflect final revisions after the second field test.

## **EXAM DESIGN AND DEVELOPMENT: PHASE I**

The initial planning, exam development, two field tests and revisions occurred in four phases. Phase one was devoted to identifying the specific performance outcomes for which test items would be later developed, the design of the first draft of the two-part exam, and drafting a preliminary test specifications document. Fortunately, two major ACRL-sponsored information literacy documents had been recently published: *Information Literacy Competency Standards for Higher Education* (ACRL, 2000 and endorsed by AAHE) and *Objectives for Information Literacy Instruction: A Model Statement for Academic Librarians* (ACRL, 2001). The project team spent several months discussing how to adapt the indicators and outcomes from the two ACRL published documents. Decisions were based on which outcomes were appropriate for community college students. The team also closely reviewed two local adaptations of the ACRL documents. These local statements of information competency standards, indicators, and performance outcomes were created with input from discipline-based faculty at Diablo Valley College and Santa Rosa Jr. College. This process resulted in several drafts and revisions, using the two national documents and the two local ones. The consensus document, "[Bay Area Community College Information Competency Standards, Performance Indicators and Outcomes](#)," reflects our final consensus.

---

<sup>2</sup> The project team included these Bay Area librarians: Bonnie Gratch-Lindauer and Brian Lym, City College of San Francisco; Amelie Brown and Andy Kivel, Diablo Valley College; Micca Gray, Santa Rosa Jr. College; Topsy Smalley, Cabrillo College; and Pam Baker, CSU-Monterey Bay. Active contributors were Evelyn Posamentier, Skyline College; Jun Wang, Delta College; and Karen Gillette, Foothill College. All these institutions are thanked for their support. The project team acknowledges CARL for the research grant award and DVC for grant monies donated for consulting fees for Steven Spurling, institutional researcher at City College of San Francisco. We owe a huge debt of gratitude to team member Topsy Smalley, for posting and maintaining the project documents on her Web site.

The design of exam items for the first field-test was time-consuming and involved several drafts and discussions. To make the task more manageable, the team sub-divided into two groups: a sub-group for the cognitive items, which became Part A and includes multiple choice, matching, and short answer items, and one for the performance tasks, which became Part B. Several issues arose which required decisions that affected how the project team conceptualized and phrased exam items. The following points illustrate their decisions:

- use as little library jargon as possible and be aware of language relating to cultural/ethnic differences and English language proficiency;
- when appropriate, phrase items within a specific information need or scenario to provide context; include academic type scenarios of information needs as well as vocational and practical ones to reflect the diversity of community college students' educational goals and experiences and their personal needs for information;
- use the active voice, rather than the passive voice;
- include a mixture of format types, such as multiple choice and short answer, to best match the outcome assessed; limit the number of answer choices for the multiple choice items to no more than five, but no true/false or yes/no answers since it was felt this format is too easy; and
- ensure that the item adequately measures what the outcome states.

The Part A sub-group divided the task of writing test items among its four members. Each one was assigned one of the four standards and he/she wrote test items based on the specific performance outcomes from the consensus document. Copies of locally-developed tests were studied to get ideas for item types and formats. Two test items for each outcome were created so question variations could be evaluated in the first field test. The Part A sub-group, over the course of several drafts and communications, made decisions about which scenarios would be appropriate for a specific outcome and which outcomes required a different format than multiple choice. Occasionally there was disagreement about the correct answer of the test item. After the consultant reviewed the draft, more changes, based on good test design principles, were made. They included:

- a complete reorganization of the items so that those of the same format type were grouped together (i.e., all the short-answer items, all the multiple choice items, all the items that include screen shots or boxed text);
- within the multiple-choice section, the "circle all that apply" items were grouped separately from the "circle the best answer" items;
- a few of the items were revised to improve their clarity;
- the number of negatively phrased items was reduced, e.g.,: "Which of the following is NOT an example of a primary source?"
- consistent use of shading or boxes was used for screen shots and quoted text, which were located before the question; and
- some "circle all that apply" items were changed into "select the best answer" ones.

The final step in this phase was the production of an answer key for Part A and a rough draft of the scoring rubric for the short answer items, with the idea that it would be expanded based on an analysis of the actual responses from the first field test.

The Part B sub-group also worked over several months to produce the draft used for the first field test. They reviewed some library research assignments for ideas for item content and format, as well as referred to the outcomes listed in the consensus document "[Bay Area Community College Information Competency Standards, Performance Indicators and Outcomes](#)." After several drafts they created several exercises which require students to:

- narrow a broad topic and explain the steps taken to do so;
- pose a research question;
- identify a research question's main concepts and synonyms/related terms;
- do research using three different search tools to identify three relevant information sources for the research question;
- cite each source using a standard citation format;
- describe how each source was found and explain why it was selected;
- present additional research steps that might be taken to further address the research question;
- evaluate two Web pages according to standard criteria;
- paraphrase main ideas from the information provided at the two Web pages.

The performance part was constructed to evaluate a student's abilities to formulate correct/appropriate responses after completing various research steps. Thus, many of the exam questions require an answer, as well as an explanation of how the student arrived at that answer. For example: "Narrow this subject to a manageable topic for the assignment" (Item B.1.a); or "In two or three sentences, describe in detail what you did to develop your narrowed topic" (Item B.1.b). Part B was designed as a Web-based form to facilitate the capture and sharing of responses. Although it does not adequately assess technology competency – a job more suited to portfolios, or project-based assessments – the group felt it important that students demonstrate basic information technology skills necessary to successfully navigate the online format.

Perhaps the most difficult aspect of developing the Part B items was the conceptualization and later development of the scoring rubric. While the sub-group drafted the Part B items, they began a discussion about how to design the scoring rubric. Rubrics, which are scoring tools, are essential for performance-based tasks because this type of assessment does not lend itself to answer keys like multiple choice tests do. A rubric provides scoring criteria, guidelines, and examples of appropriate and inappropriate responses so that different people scoring the responses will generally arrive at the same score, thus greatly reducing the subjectivity of scoring. The extent of agreement among different scorers is a measure of the reliability of the rubric.

To prepare for developing the scoring rubric, the sub-group members each completed the exam and constructed a table that mapped test items and responses back to the performance outcomes. In this side-by-side display, performance outcomes were linked to test items, right along with actual responses, and a fourth column was added for inserting ideas for what might go into the scoring rubric. This table of a preliminary scoring rubric was used throughout the process of developing the final rubric.

### **Preparation of the Test Specifications Document**

The final part of phase one involved writing a preliminary test specifications document. The consultant provided guidance and suggestions for the content and organization of this document, which could not be completed until the end of the project. Test specifications not only clarify the content and scoring of an instrument, they also serve as documentation that would allow someone else to develop or modify items while maintaining the integrity of the test design. The test specifications document includes several sections: purpose; description of the content, especially as it relates to the criterion-related levels of performance; details about the development of test items; information about format, administration of the exam, and scoring procedures; charts which map the test items to specific outcomes; and a table which shows the item mix and point values for each standard's outcomes.

Producing the mapping for each standard was a two-step process, since each sub-group mapped its items and then a combined mapping was produced later to reflect the entire exam. It cannot be stressed enough how helpful this mapping was, as it revealed which outcomes had been inadvertently overlooked when designing test items, and it uncovered which outcomes were represented by multiple items. An early example of this combined mapping follows with an extract from the table for Standard One, which states “the information literate student determines the nature and extent of the information needed.” The latest versions of the [maps](#) for each of the four standards are part of the [Test Specifications](#) document and are available on the project Web site <<http://www.topsy.org/ICAP/ICAPProject.html>>

**Table 1: Information Competency Outcomes Mapped to Exam Items: Standard One**

Test Items	Outcome(s) Assessed	Points	Test Section/Item
Which group of keywords best describes the information need for the following assignment: Prepare a presentation about the gains achieved by women in the 1980's?	Standard 1.1.5 Identifies key concepts and terms that represent the information need or research topic/question.	1 pt.	A.4.
You've been given the assignment to write a 3-5 page research paper on the following topic: "Should colleges be allowed to restrict student speech?" What are the main concepts for this research question?	Same outcome	9 pts.	B.3.a.1,b.1,c.1

**PHASE 2: FIRST FIELD TEST**

The first field test had specific objectives: to identify unclear or problematic items, to determine the amount of time required to complete the exam, and to use the actual responses from Part B to help improve the scoring rubric. Part A included 33 multiple choice items, 5 short answer and 7 matching/ranking items. Part B was composed of 7 items with many sub-parts, bringing the number of potential responses to 26. A cover page to Part A also surveyed students about their experiences using libraries and online resources; the amount of previous training/instruction in research skills; and asked them for a self-rating of their research skills.

It was field tested in April and May 2002 at four different community colleges in the greater Bay Area with a sample population of 29 for Part A and 19 for Part B. The project team did not develop instructions for administering the exam, but learned that most of the students completed it in a classroom lab or library setting. Students were given the option of doing both parts or just one part of the exam. Completion time ranged from 25 to 50 minutes on Part A, and between one hour and one hour and a half completing Part B. The sample purposefully included non-native speakers of English and students who did and did not have previous library research skills instruction. Only a few of the respondents completed both parts of the exam, and since student name or identification number was not required on the separate parts, it was not possible to match them. Three of the project team members scored the completed Part A tests and together compiled comments about problematic items. The Part B sub-group alternatively scored and used the responses to further develop the scoring rubric.

The consultant's report contained a detailed item analysis and correlation table. The item analysis revealed the Part A items that were problematic and needed revision. The correlation table helped identify which items the higher scoring and lower scoring students missed. The sample was too small to use the factor analysis data or to reliably correlate amount of previous information competency skills instruction to

performance. Based on his analysis and the project team's, the following activities occurred to improve Part A:

- several items and/or answer choices were reworded to improve clarity;
- some item formats were changed from a multiple choice to a short-answer, or vice versa, to adjust the level of difficulty;
- a few items involved revisions to make them more challenging;
- a few items were moved to a different section of Part A to keep item types together; and
- "circle all that apply" choices were eliminated.

Another concern was the need to establish empirically a valid passing score for the exam. From various approaches suggested by the consultant, the project team used a "panel of experts" approach for Part A items only. Three members of the Part A sub-group estimated the percentage of students who would answer each item correctly, and then later compared these estimates to the actual percentage of students who correctly answered each item. The consultant suggested that the estimated percentages might be averaged to arrive at a cut-off score. As it turned out, the sample wasn't really large enough to reliably use these calculations to establish a cut-off score. A final activity was the production of a scoring key and the revision of the brief rubric for the short answer items.

### **Developing the Part B Scoring Rubric**

The other major objective of the first field test was to further develop the Part B scoring rubric. During phase one, a three-level scoring rubric had been sketched out – exemplary, competent, and emergent. Discussions surrounding these distinctions were critical to decisions about performance levels. After the consultant and another project team member reviewed the first draft of the rubric and attempted to use it to score a few completed tests, the three-level scoring concept was dropped. The three-level scoring system, while useful in a course or another assessment setting, wasn't deemed suitable for the project's purpose of a challenge-out test, which needs to set a standard, showing whether the performance bar had been met or not.

There were nineteen Part B submissions in the first field test, and these provided the substance both for devising the scoring rubric and for pre-testing it. In essence, a process of deconstruction was applied. The Part B sub-group had to tackle such questions as: how could discrete and describable criteria be enumerated so that levels of correctness could be identified? And, what levels of performance should be expected, i.e., where was the bar defining performance that reflected the level of information competence appropriate for two-year college graduates?

One of the most challenging tasks was to devise language that would describe the elements distinguishing the meaning of an adequate response. For example, most scorers would correctly *recognize* a research topic that has been sufficiently narrowed to be appropriate for a 3-5 page research paper. However, coming up with a description of what characterizes a sufficiently narrowed research topic that could be applied *across* research questions is another matter. In analyzing student responses, the important questions were these: If someone responded adequately or well on an item, what specific characteristics were associated with that response? For test responses that obviously failed to hit the mark, what was it that characterized those responses and others like them? The goal was to extract these characteristics and describe them generically so that any librarian-scorer could apply them. For example, the criterion for scoring whether a student had developed an appropriately narrowed topic (Item B.1.a) reads, in part: "1. topic is narrowed by specifying time frame, or persons, or organization or group, or location, or event or incident, or some combination of these, or other similar, appropriate limiter(s) is (are) applied." Some test items and their performance outcomes were more complex and composed of subparts; not only did the level of performance required for competency need to be addressed, but also the issue of which subparts had to be demonstrated in tandem.

Working independently, and using the evolving rubric, sub-group B members scored tests, compared scores, and discussed how specific scores had been determined. Developing and refining the scoring rubric was a very iterative process. As a last step to completing the rubric, the sub-group added

examples to clarify the meaning of the criteria, supplied examples of adequate/inadequate responses, and assigned a weighted scale of numerical values for scoring. Project team members decided to weight the scores in order to have a cumulative total points possible of 100 points. For most test items all of the criteria had to be satisfied in order for the student to score any points, so that partial fulfillment of the criteria with a corresponding score was only allowed for three of the 12 items. The part B sub-group also prepared several documents to be used for the second field test: "Introductory Notes on the Performance Exam and Scoring," a scoring manual and a scoring sheet.

### PHASE 3: SECOND FIELD TEST

The preparations for the second field test involved more planning and the production of instructions and other documentation for those administering the exam. For the second field test the group invited additional institutions to participate so that there would be a sample of at least 50 respondents completing the entire exam. Another goal was to test the accuracy and usability of the newly developed scoring rubrics by asking librarians not previously associated with this project to use them.

A packet of materials was distributed both electronically and by postal mail to the six participating community colleges in early November 2002 so that the exam could be administered before the end of the fall semester. This packet contained a cover letter detailing how to administer and score the exam; exam parts A and B; scoring sheets and instructions; Part A answer key; Part B scoring manual and instructions for using the scoring rubric, and the scoring rubric mapped to performance outcomes. Unlike the first field test's sample population, this sample was purposefully designed to be students who had some previous library research skills/information competency instruction. Five of the six participating institutions used the exam in the context of an information competency class which provided an actual test taking environment. The other one posted a sign in the library and an enticement of a \$5.00 copy card. To encourage the timely return of the completed tests and reduce the burden on using class time, instructions for administering the exam stipulated that students should complete the paper-copy Part A in a classroom setting, while they were allowed to complete the online Part B at home or in another setting where they had Internet access. Participating librarians were asked to score both parts of the exams taken by students at their institutions and return them by the end of January 2003. To assess inter-rater reliability, Part Bs were additionally scored by two project team-members so that each completed Part B was scored by three individuals.

While there were many more submissions, the following shows the originating institutions for the 57 completed exams included in the analysis:

City College of San Francisco	12 exams	21 %
Diablo Valley College	11 exams	19 %
Foothill College	4 exams	7 %
Glendale Community College	19 exams	33 %
San Joaquin Delta College	5 exams	9 %
Santa Rosa Junior College	6 exams	11 %
Total	57 exams	100%

Two project team members coordinated the online submissions of Part B and the paper-copy Part As to check for completeness of responses and to match the student ID number for both parts. About mid-way during the Part B scoring process, the person coordinating the online submissions of Part B and the online submissions of the scoring sheets became aware of a fairly large range of scores among the three librarian-scorers for certain items. She suggested a phone conference to clarify the use and interpretation of the scoring rubric. The phone conference allowed project team members to better understand their interpretations of responses and made it clearer which aspects each scorer had focused on. As a result of the phone conference, their observations were recorded for the later revision of the rubric and for scorer training. In order to preserve the integrity of the second field-test, no scores were changed and the rubric

remained as it was for the completion of the second field-test. The project team began to articulate the feeling that scorer training had been insufficient and that each institution using this exam might need to develop its own “group mind” about how to evaluate and score the range of answers depending on its information competency instructional program and local setting.

## **Findings from the Second Field Test**

Findings from the second field test came from the consultant’s analysis as well as from the experiences and reflections of the project team members and participating librarians. The consultant’s overall analysis found that the instrument needed some further work, a finding that confirmed much of the project team’s own observations. Selected findings from the consultant’s descriptive statistics and item analysis follow:

- Part A - a reliability score of .76 on the Kuder-Richardson 20 scale, which is not sufficiently high; identified eight “problem” items needing further revision;
- Part A – inter-rater reliability on Part A was .99, which means that the answer key and rubric worked well for the eight short-answer items;
- Part A score range was from 61 high (out of 64 unweighted points possible) to 21 low; mean score was 45.2 points (Std = 8.57); and the median score was 45 points; 31 students scored 45 or higher, which is 54.4% of the total number of respondents; a score of 45 or higher reflects getting 70.3% of the items correct; and 18 students (which is 31.6% of the total) scored 51 or higher, reflecting 80% or more of the items correct;
- Part B – inter-rater reliability of .71-.77, which is insufficient for overall test score reliability;
- Part B, averaging the 3 scores, the range was from 90.3 high to 11.3 low; mean score of the three scorers totals was 51.4 points out of a possible 100 weighted points (Std = 18.43); and the median score was 47.3;
- the relationship between how students rate their IC skills and their actual performance was not particularly strong (for Part A  $Rho = .275$ ,  $p < .04$  and for Part B  $Rho = .10$ ,  $p < .04$ ), but the mean self-rating was 7.1 on a 10 point scale of 10 being the highest;
- a positive and fairly high correlation was found between the total score on Part A and Part B ( $Rho = .588$ ,  $p < .0001$ )

## **PHASE 4: REVISION OF THE EXAM AND SUPPORTING DOCUMENTATION**

Absorbing the consultant’s report and deciding how to complete the project required a great deal of study and discussion. The project team had several e-mail exchanges, two phone conferences and an all day meeting to analyze the consultant’s report and make plans for how to address the findings and complete the project. They used the summer months of 2003 to carry out the decisions reached during these various conversations. They decided not to do a third field test, but instead to concentrate on revising the problem items identified in both parts of the instrument, review the original mapping of test items to performance outcomes, and to prepare thorough documentation to explain to others how test items were developed. The decision not to conduct a third field test was reached partially because of time constraints, but primarily because team members were convinced that local use of the instrument would likely result in some modifications. In addition to the revision activities, they added more detailed explanations in the “Test Specifications” documents about how items were developed so that local modifications would be guided by good test design principles. This decision means that the exam lacks statistical reliability and validity, since the final revisions were not further field-tested. The final revision of the instrument benefited from several inputs. One input came during an all-day meeting of the project team members when another examination of the content validity of test items was made by reviewing the mapping of test items to performance outcomes. Another input came from the consultant’s item analysis where he had flagged items which the better performing students missed. These items were examined to ensure that they were clearly phrased and not misleading. In some cases the items were rephrased or completely redesigned, in others the project team conceded that it was an appropriately difficult item and didn’t need revision. A third input came from an ESL instructor’s examination of both parts of the exam. Her involvement was requested by one of the project team members after the draft instrumented was

shared with a large group of faculty at her institution. Faculty comments identified the possible difficulty of some items for non-native speakers of English and potential length and difficulty of the exam. Several items were rewritten to simplify and clarify the language for non-native speakers of English.

Part B items benefited from the same three inputs. The major issues for Part B revision centered on finding ways to encourage students to write more complete explanations of their research steps and reasons for their responses, and increasing the inter-rater reliability for scoring. A few items were slightly revised to provide more prompts for students. The scoring manual rubric was improved to include many more examples of adequate and inadequate student responses, as well as more notes to the scorer. In response to the team's concern about item 4 of Part B, which required students to conduct searches and identify three relevant sources for the research question, the scoring and weighted points were reallocated to emphasize the selection of the source and its quality over the formulation of the citation. Thus, a new criterion was added and 3 weighted scores were created. Moreover, suggestions were made in the "Test Specifications" document about conducting training for the scorers before using the rubric.

In addition to concerns about ensuring content validity, unintentional weighting of certain outcomes over others concerned the team. They looked at the combined points and identified some outcomes which were mapped to two or more test items. The point of this activity was to more consciously decide whether they wanted to weight certain outcomes more heavily than others. For example, they discovered that the outcomes associated with posing and modifying a research question and identifying the major concepts and synonyms for these keywords in the research question represented 44.5 points of the total exam of 164 points. As a result of this activity, two items were eliminated from Part A and the assignment of weighted points in Part B was revised.

The project team discussed in length the overall scoring of the exam. They dismissed the idea of trying to change the scoring so that each part of the exam had the same number of points, since there was agreement about not wanting to assign more than one point for Part A's multiple choice items. Thus, the final instrument includes two parts with total possible points allocated as follows: Part A has 65 points and Part B has 100 weighted points. They also discussed the notion that a passing score for a challenge-out exam should not be higher than C-level performance expected of students in typical information competency credit courses. Rather than try to reach a decision about the best way to score the exam and determine a passing score, they decided to present several possible scoring options so that a local institution could select one that was appropriate for their needs. These three options include:

1. A passing score could be expressed in total points necessary to pass from both parts of the exam. One can easily calculate a percentage of correct answers for both parts of the exam by setting the total possible points equal to 100%. Example: total points = 165 = 100% and passing score is decided to be 70%. Thus, a student would need to score 116 total points or more in order to pass the exam. This option allows a student to compensate for poorer performance in one part by better performance in the other.

Variation to option #1: The score for Part A is changed to a percentage. Part B is already 100 points, so it is readily converted to a percentage. The two percentages are added together and divided by two which becomes the "final score." A threshold is established for a passing grade. Example: a student could score 32.5 on Part A or 50% of the items correct and 94% on Part B and the average would be 72%, a passing score if the threshold were 70%.

2. A passing score is established for each part of the exam. Each part maintains the same threshold (for example, 70%). In this scenario, a student has to perform at a passing level in both parts in order to pass the whole exam.

3. A passing score or percentage is separately established for each part of the exam, but in this case one part could be different than the other; e.g., Part A = 70%; Part B = 80%. In this scenario, a student has to perform at a passing level in Part A as well as in Part B, but one part could be weighted more heavily.

The final tasks of the project involved another lengthy conversation of the project team members to review the last revisions to the exam and scoring rubric, as well as all the supporting documentation. After this final review, they reached consensus about the rewording of Part A and Part B items as well as the detailed explanations about how the items were developed. They also decided to change the weighting of the Part B items so that the weighted points assigned to the three major sections would be relatively proportionate and also reflect the group consensus about which section should have the most weighted points. Thus, the distribution of weighted points is as follows:

Tasks	Weighted Points
<b>Exercise 1 Preparing for research</b>	<b>30 points</b>
B.1.a. narrowing a topic	6 points
B.1.b. steps taken to narrow topic	6 points
B.2. posing the research question	6 points
B.3. identifying key concepts and synonyms	12points
<b>Exercise 2 Conducting searches and preparing citations</b>	<b>41 points</b>
B.4.a.1.-c.1. identifying 3 relevant sources for research question and specifying citation guide and formatting 3 citations	12 points
B.4.a.2.-c.2. explaining how found each of the 3 sources	9 points
B.4.a.3.-c.3. explaining why chose each of the 3 sources	9 points
B.5. describing additional research steps to address research question	7 points
<b>Exercise 3 Evaluating Web documents &amp; Synthesizing Information in own words</b>	<b>29 points</b>
B.6.a. & b. evaluating two Web documents	20 points
B.7. synthesizing information from Web document in own words	9 points
Total	100

The team delegated among several members the final revisions to the maps and all the scoring documentation. Finally, all the project's products were completed and most of the documents were posted to the project's Web page <<http://topsy.org/ICAP/ICAPProject.html>>

The project team plans to leave these digital materials at the project Web page at least through 2004. The idea is to share the "how-to" information found in the "Test Specifications" document so that others can modify the exam for their local purposes, while maintaining the integrity of good design of exam items. A project team member will provide a copy of the exam to requesting librarians. The project team requests that anyone administering and/or modifying the exam study the explanatory information and advice found in the "Test Specifications" document. In addition, the project team hopes that anyone who uses the exam will acknowledge the source and share their results and experiences with the project team.

## **DISSEMINATING INFORMATION ABOUT THE PROJECT**

There were several informal and formal opportunities to share information about the project and the results of the two field tests over the course of this three-year project:

- informal sharing of information with library and discipline-based faculty colleagues on the community college campuses of each of the project team's members;
- frequent sharing of drafts and ideas among the members of both sub-groups and between the two sub-groups of the project team and observers/reactors via e-mail and phone conferences;
- periodic project updates to a contact person at a neighboring California State University;
- presentations at four conferences: the pre-conference on Information Literacy at the annual California Community Colleges Chancellor's Office Conference (March, 2002); the California Academic and Research Librarians annual conference (CARL, May, 2002); the California Library Association Conference, Academic Libraries section program (November, 2002); at two California Community College Library workshops (March 2003); and the ACRL national conference (April, 2003);

- a progress and final report published in the CARL Newsletter (September, 2002 and September, 2003);
- project documents and sample test items published on the Web site of one of the project team members; and
- postings to California academic library and national library instruction and information competency listservs about the project.

Collaboration with representatives of the California State University (CSU) librarians was expanded about midway through the project. Quite fortuitously, the Manager for Information Competence Initiative from the CSU system's Chancellor's Office made contact with a member of the project team. Thanks to her interest and initiative, early drafts of test items were shared with the CSU Information Competence Test Development Task Force. This group of CSU librarians was involved in a similar task. A productive meeting between the two groups was held in Spring 2002, which resulted in an improved understanding of our respective purposes and a commitment to continue to share all drafts and results of our work. Two other opportunities afforded a joint meeting with the CSU system contact person, a representative from the Bay Area community college project, and contacts from an Association for Research Libraries-affiliated project, Project SAILS. Project SAILS, which began after the Bay Area Community College Library Assessment Project, is a large-scale project dedicated to developing a standardized information literacy assessment instrument. These informal meetings were held at the American Library Association mid-winter meeting in January 2003 and at the ACRL national conference in April 2003.

## **Lessons Learned**

This three-year undertaking was a remarkable example of dedicated librarian collaboration and commitment to seeing this project succeed. While not inter-disciplinary in the traditional sense, it was certainly inter-segmental and inter-institutional. Operating on a shoestring, all the project team librarians spent endless hours of their own time, along with some support from their respective institutions. A collegial model of interacting became the norm. Undoubtedly, what helped the project succeed was identifying a clear purpose from the start, the project team members' experience and expertise teaching and assessing information competency skills, the size of the team, and the sub-division into two working sub-groups.

The project team learned that a regional approach to a common need is a viable one which can succeed with a fairly large and complex project. However, learning from their experiences, the following suggestions are offered to anyone undertaking a similar project:

- prepare a more detailed project plan and timeline at the early stages so that the complexity and length of time required for the project phases can be better anticipated;
- secure a larger grant and hire the consultant at the very beginning so that he/she might guide the development sooner of the test specifications document and the exam items;
- hire production help for the preparation of the exam and supporting documents, especially planning ahead to put both parts online;
- expand the number of institutions for the first field test to increase the sample size and specify instructions for how to administer it to ensure that most students complete both parts of the exam;
- more consciously plan for recruiting students who have been exposed to a range of previous instruction, i.e., none, just workshops, an information competency course;
- plan for incentives for students to complete the exam to increase the numbers of test-takers from each institution;

- designate a team member to coordinate the various tasks and to be responsible for ensuring timely communication between the working sub-groups;
- prepare and use the combined mapping of test items from both parts of the exam as a tool for deciding weighting of points and possible elimination of test items between the two field tests; and
- plan to dedicate more time to training the individuals using the rubric to score the performance-based items.