

2017-18 ICaP Assessment Report: Common Assignment Pilots

Daniel Ernst, Outgoing Assessment Research Coordinator
Carrie Grant, Outgoing ICaP Assistant Director
Alisha Karabinus, Incoming ICaP Assistant Director
Derek Sherman, Incoming Assessment Research Coordinator

Table of Contents

Introduction.....	2-3
Summary	
Assessment Project Goals	
Common Assignment Development.....	3-5
Challenges	
Pilot Types and Details	
Assessment Methods and Results.....	5-13
Methods	
Results	
Recommendations and Next Steps.....	13-17
Common Assignment Structure for Fall 2018	
Common Assignment Pilot Changes	
Rating Session Recommendations	
General Recommendations	
Thank You.....	17
Appendices.....	18-24

Introduction

Summary

This report contains an overview of Introductory Composition at Purdue's (ICaP) 2018 assessment project: the creation and piloting of six English 106 common assignments. The report begins by describing the conditions that led to this project, explains the process of developing the six assignments, details the results from initial assessment efforts, and recommends changes for future iterations of ICaP common assignment projects.

Assessment Project Goals

The Council of Writing Program Administrators (CWPA) completed an external review of Purdue's composition program in early 2017 and stressed two important concepts for ICaP's continued program development: **coherency** and **consistency** in seeking to meet learning outcomes for students. While the CWPA reviewers praised the syllabus approach system implemented for ICaP courses, the reviewers noted the difficulty in assessing achievement outcomes when faced with a diverse slate of instructors and courses. Reviewers recommended a single common assignment as a method for increasing consistency (and our own ability to assess the program) while preserving instructor agency in developing the course alongside the common assignment. This single common assignment could offer instructors and administrators a variety of benefits, from a measurable point of reference across courses and approaches, as well as serving as an anchor for professional development and pedagogical instruction (as it offers built-in opportunities for grade-norming workshops, lesson plan instruction, and other instructor support measures).

In order to identify a common assignment approach that would meet these goals while maintaining the flexibility that allows ICaP instructors to develop their own courses, the common assignment pilot was developed. Syllabus approach leadership teams, as part of the Pedagogical Initiatives Committee (PIC) were each tasked with developing a potential assignment for testing by program instructors in Spring 2018, along with instructional guides and other support material allowing individual instructors to easily integrate and implement the pilot assignments. These individual assignments were meant to be reflective of the syllabus approaches from which they were developed, but also assignments strongly reflective of the pursuit of ICaP's learning outcomes.

The pilot, extended now into Fall 2018, also allows for testing of which assignments integrate most easily into a variety of syllabus approaches, and can most effectively measure program outcomes. While assessing achievement of program outcomes is the central purpose of the assessment program, it is also important to remember that ICaP instructors come from a variety of backgrounds and work from a multitude of pedagogical positions, and instructor agency is a

key feature of ICaP, a feature allowing for a chance for instructors to leverage their unique skills in the pursuit of program goals. The six pilot assignments are currently being assessed for how well they allow instructors to continue to showcase what they do best while offering a chance for administrators to assess student writing ability, student learning, and the pursuit of program learning outcomes.

After the initial pilot in Spring 2018, student papers were collected from instructors and de-identified for early assessment—both of the program and the assignments themselves, with the goal of testing these assignments to learn more about how to implement a common assignment, or set of assignments, that best fits the program and its students.

Common Assignment Development

As the assessment project began in Fall 2017, a major goal was to create diverse pilots that fit the diverse needs of the program and ICaP's various syllabus approaches, which is why PIC was tasked with this development. PIC is comprised of at least one member each syllabus approach. Each common assignment pilot was developed by at least two individuals from the PIC committee and met at least one ICaP outcome. We intentionally tested a range of assignments and assessment methods to see what would best fit our program and provide the most meaningful assessment results.

The six pilots included the following assignments: Professional Email, Rhetorical Analysis, Literature Review, Reading Annotations, Writing Portfolio, and Information Literacy Essay. A description for each of these assignments can be found at the end of this section. Each assignment also has an instructor's guide that contains a rationale, relationship to ICaP outcomes, and a teaching sequence to guide pilots. The instructor's guides were developed by the approach leaders in Fall 2017 and they can be found via the following folder: [Common Assignment Instructor Guides](#).

Challenges:

The biggest challenge for pilot assignment creators was working towards a common assignment goal. Was our task to identify a single common assignment for the program that measures all of the ICaP outcomes? Will we maintain several options that address one or two ICaP outcomes for each assignment? We navigated this challenge by creating a range of assignments, some that address all outcomes (the portfolio), some that target primarily one outcome (professional email), and some that could address a few different outcomes together (research paper). In the future, we will have to revisit this challenge when we finalize which assignments will be used program-wide. We must also consider the task of university outreach. How can we demonstrate to the rest of the university community that what we do is valuable in a Purdue context and that English 106 and 106-I are important courses to preserve in a student's curriculum?

Additionally, the Purdue context was another challenge because we needed to show that ICaP is preparing students to become better writers within their disciplines. Each pilot seemed to address one aspect of the Purdue context. For example, the Professional Email assignment was utilized to make sure students know how to properly communicate with their professors. The Literature Review, too, was developed so that students would develop the specific skills of summarizing, analyzing, and synthesizing, which are key skills to have at a research institution. In choosing the final common assignment option(s), it will be crucial to determine what narratives we want to be able to tell about the skill set students develop in ICaP courses, without compromising the humanities-based model of exploration and growth upon which composition is founded.

A final constant challenge is instructor hesitancy towards a common assignment. Since ICaP is known for its curriculum flexibility for instructors, a common assignment seems to counter this value and limit instructor freedom. An initial goal during pilot development was to have pilot leaders from different syllabus approaches collaborate on each pilot so that we could see how it worked across syllabi. We did have instructors across various approaches test the pilots, but in the future, more could be done to assess how the common assignment will work in different approaches. Instructor hesitancy will always be a challenge, but slowly incorporating the changes of a common assignment should alleviate this concern, and attention should be paid to making sure that instructors maintain the classroom agency that research shows makes for the happiest, most effective staff.

Pilot Types and Details:

1. Professional Email Assignment

- Pilot leaders: Alisha Karabinus (akarabin@purdue.edu) and Bianca Batti (bbatti1@purdue.edu)
- *Assignment overview:* Students write two email messages to different audiences
- *Timing:* First two weeks of semester
- *Grading:* Common rubric to give ICaP, flexible rubric format to give students

2. Rhetorical Analysis

Pilot leaders: Alex Long (long205@purdue.edu), Daniel Ernst (ernst9@purdue.edu), and Carrie Grant (grant34@purdue.edu)

- *Assignment overview:* Students write a ~3 page paper analyzing the rhetorical choices of a text
- *Timing:* Assignment given twice during semester--once with no instruction at beginning of the course, once with a full course unit
- *Grading:* Common rubric to give ICaP, flexible rubric format to give students

3. Research - Literature Review

Pilot leaders: Derek Sherman (sherma11@purdue.edu) and Amanda Leary (learya@purdue.edu)

- *Assignment Overview*: Students summarize, analyze, and synthesize at least five academic, peer-reviewed sources on a topic
 - *Timing*: Four week unit anytime during semester
 - *Grading*: Common rubric, adjustable point values
4. Reading Annotations
Pilot leaders: Eugie Ruiz (ruiz56@purdue.edu) and Elizabeth Geib (geibe@purdue.edu)
- *Assignment Overview*: Students annotate course texts for understanding and synthesis
 - *Timing*: Throughout the semester, any number of times.
 - *Grading*: Common rubric, adjustable point values
5. Writing Portfolio
Pilot leaders: Mitchell Jacobs (jacobs95@purdue.edu) and Phuong Tran (tran110@purdue.edu)
- *Assignment Overview*: Students compile a portfolio of their work in the course and reflect on their growth and revision processes
 - *Timing*: End of semester (using documents collected throughout semester)
 - *Grading*: Optional rubric
6. Information Literacy Essay
Pilot leaders: Daniel Ernst (ernst9@purdue.edu) and Carrie Grant (grant34@purdue.edu)
- *Assignment Overview*: Students write two in-class information literacy essays
 - *Timing*: Beginning and end of semester
 - *Grading*: Common assignment prompt, no required rubric, minor inclusion in grade

Assessment Methods and Results

The Spring 2018 Common Assignment Pilots were evaluated by a team of ICaP instructors and administrators. The team was led by Carrie Grant, the ICaP Assistant Director; Daniel Ernst, the Assessment Research Coordinator; and Dr. Bradley Dilger, the Director of ICaP. Various Purdue English graduate student instructors also helped by serving as raters on the ICaP Assessment Committee, which rated the individual assignments. The primary goal on the assessment end was to trial run different rating configurations in addition to the different assignments types and rubric styles mentioned above to note any changes necessary to conduct an eventual program-wide assessment as seamlessly and effectively as possible.

Methods

Participants

Participants included graduate student instructors and part time lecturers who volunteered to teach and rate one of six common assignments in their class. In total, 39 instructors participated in piloting the assignments, who collectively taught their selected assignment to over 780+

undergraduate students taking English 106 and 106-I during the Spring 2018 semester. Over three months, a total of 25 graduate student instructors read and rated samples from the six pilot assignments as part of the inaugural ICaP Assessment Committee.

Procedures

The pilots were incorporated into ICaP courses as typical writing assignments. There were no changes made to participating instructors' courses other than the incorporation of one of the six assignments into their English 106 sequence. The assignments were administered by instructors at various times throughout the semester, depending on assignment type, and once students completed the assignment, instructors delivered the documents to the ICaP assessment team. From there, all documents used were sorted into random samples and de-identified before being assessed.

The six assignments were collected, rated, and evaluated over the course of the Spring and Summer 2018 semesters. Each pilot was assessed over its own two day period during rating sessions lasting two hours each day. Rating sessions consisted first of norming the raters using practice sample essays followed by discussions among raters to build consensus on how best to apply the scale and interpret the language of the rubric used. A mixture of rubrics was used across all six assignments. Some rubrics were adapted versions of those used by instructors in their classroom or designed by the assignment creators themselves, while others were built directly from the six ICaP outcomes and detailed learning objectives (see appendices A-F). Each essay was then read and rated by at least two raters, with significant disagreements in scores resolved by a third. All rated essays were randomly selected from a pool of participating instructors and were de-identified to prevent raters from knowing who wrote the essay, which course or instructor the essay was written for, and whether the essay was a pre-test (rough draft) or a post-test (revised draft) when applicable.

Rating Sessions

The Assessment Research Coordinator scheduled and ran all six rating sessions. The coordinator started by emailing a listerv of all graduate student instructors to ask who was interested in rating and, depending on availability, selected anywhere from 6-9 instructors at a time to participate in the two day rating sessions. The sessions lasted two hours each day and consisted first of a brief overview of the assignment to be rated and the rubric used, followed by a discussion among raters about any ambiguous language or unclear terms in the rubric.

After the raters discussed the language of the rubric, the group practiced using it on 4-5 sample essays, with a discussion following each. The coordinator collected the practice ratings and entered them into a spreadsheet to calculate basic measures of reliability (Intraclass Correlation for the group and Pearson's Product Moment Correlation between individual raters; see appendix G). Once the group felt confident that they were normed, they began to rate the actual sample of assignments, which varied in size depending on assignment, but ranged from 20-58 essays each. Raters entered scores electronically on a Qualtrics survey-version of the rubric to

make data collection easier. Generally, the group would complete a quarter of the sample on the first day and the final three quarters on the second day.

The sessions were designed so that every essay was rated at least twice by two different raters, with significant discrepancies in scores resolved by a third. The definition of “significant discrepancies” depended on the scale of the rubric in use, but generally it meant that two ratings differing by two or fewer scale points were accepted and ratings differing by more than two scale points required a third rating.¹ Each essay’s two scores were averaged, and when essays had three ratings the two scores closest in proximity were averaged.

The coordinator exported all data from Qualtrics into the ICaP spreadsheet for basic statistical analysis. Descriptive statistics, score distributions, interrater reliability measures, correlation coefficients, and effect sizes when applicable were calculated and visualized using Microsoft Excel. The results and a brief discussion for each of the six pilots follow.

Results

1. Professional Email

The first assignment assessed was the Professional Email. The sample (N=34) averaged a score of 16.88 ± 1.67^2 out of 20 on a 16 point scale (5-20), well above the scale’s true mean of 12.5. The Pearson’s Product Moment correlation between the first and second ratings was weak ($r = .31$). As a further measure of reliability, exact-adjacent ratios were calculated, with 26/34 (76%) ratings differing by two or fewer scale points, while 8/34 (24%) assignments differed by greater than two points and thus required a third rating. The distribution (fig. 1) is negatively skewed and most of the 16 point scale went unused. The rubric used was adapted from an instructor’s.

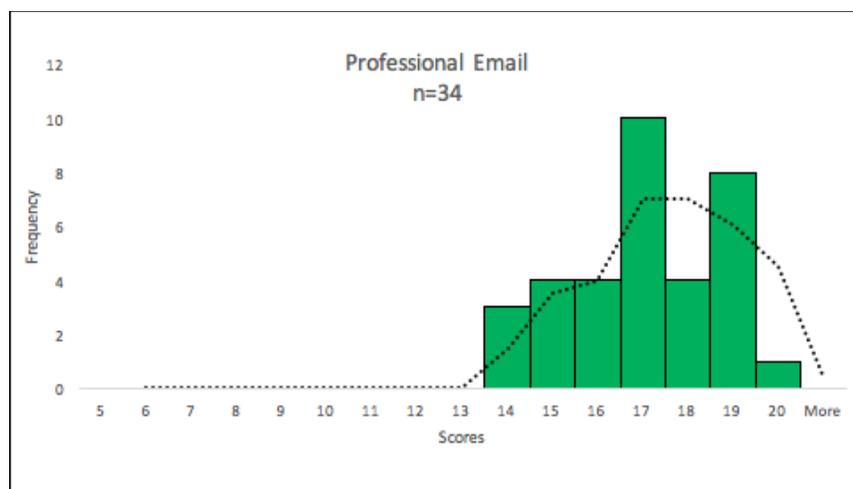


Fig. 1

¹ We decided on a threshold of two scale points because the smallest scale used was 10 points, almost twice the popular 1-6 holistic scale, which usually uses one scale point as its threshold.

² This is the mean score \pm the standard deviation, a measure of spread.

Discussion

The benefits of this assignment are its ease of use, versatility, and practicality. Additionally, such an assignment is not very disruptive to the rest of a course's assignment sequence. Per the discussion among raters, a significant challenge for rating this assignment was how to interpret what different instructors meant by and accepted as "professional." This likely affected the consistency of the ratings. Additionally, for email as a genre in general, acceptable conventions tend to vary from person to person, which might have contributed to the weak correlation coefficient as well. Because of the high mean score and because much of the scale went unused, and also because email assignments tend to be low-stakes and brief, a more narrow scale may better represent and sort email writing performance.

2. Rhetorical Analysis

Because the Rhetorical Analysis assignment is designed to be administered twice, first as a pre-test and then as a post-test, the total sample ($N=46$) is broken into two equal groups, pre-test ($n=23$) and post-test ($n=23$). The pre-test sample averaged a score of 5.78 ± 1.53 out of 12 on a 10 point scale (3-12), below the scale's true mean of 7.5. The post-test sample averaged a score of 7.63 ± 1.99 on the same scale, right at the true mean. The difference between means is statistically significant at the .001 level, $t(22)=4.25$, $p=.0003$, $d=1.04$ (see fig. 2) and the effect size is large. The Pearson's correlation for ratings was strong ($r=.82$) for the pre-test sample and moderate ($r=.60$) for the post-test sample, for an average of $r=.73$. The exact-adjacent ratio for the pre-test was 18/23 (78%) and 13/23 (58%) for the post-test. The pre-test distribution (fig. 3) is approximately normally distributed, though shifted to the low end of the scale, while the post-test distribution (fig. 4) is more spread out and flat, though the mean is shifted toward the middle of the scale; together, the samples used the entire scale. The rubric used was an ICaP outcomes-built rubric.

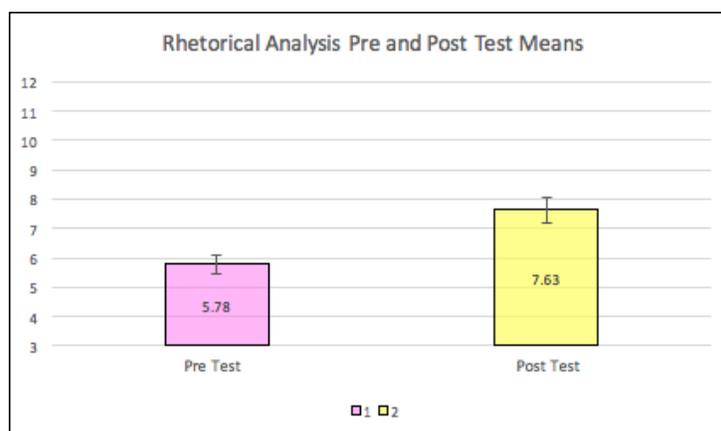


Fig. 2

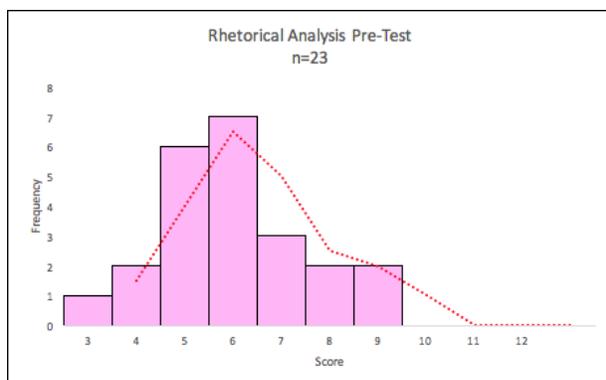


Fig. 3

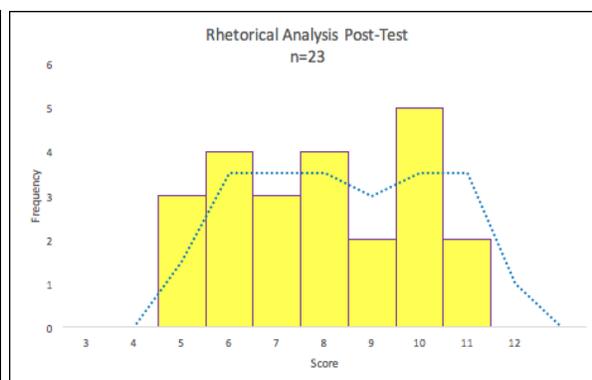


Fig. 4

Discussion

Although the large effect size is impressive, it is important to remember that the design of the assignment likely inflates it. Students were asked to complete the pre-test with little-to-no prior instruction in the principles of rhetoric and were tested again (or asked to revise their pre-test essay) only once the unit on rhetoric concluded, so significant growth is to be expected. Per discussion among raters, one salient challenge was that some instructors failed to distinguish between literary and rhetorical analysis, which made assessing the essays that offered excellent literary but little rhetorical analysis difficult. Questions also arose about how much rhetorical knowledge was sufficient for such an assignment. That said, the combined correlation coefficients is strong, the entire scale is used, and we see significant improvement in pre and post-test means. The pre and post-test format appears a good way to assess growth in the course, and assigning a cohesive essay written in a familiar genre likely benefited the assessment overall.

3. Information Literacy Essay

Like the Rhetorical Analysis, the Information Literacy Essay is designed to be administered twice, first as a pre-test and then as a post-test. The total sample (N=58) is broken into two equal groups, the pre-test (n=29) and post-test (n=29). The pre-test sample averaged a score of 10.5 ± 2.43 out of 18 on a 16 point scale (3-18), right at the scale's true mean. The post-test sample averaged a score of 11.74 ± 2.76 on the same scale, just above the true mean. The difference between the two means is statistically significant at the .05 level, $t(28)=2.69$, $p=.012$, $d=.48$ (see fig. 6) with a moderate effect size. The Pearson's correlation for ratings was strong ($r=.83$) for the pre-test sample as well as the post-test sample ($r=.94$), for an average correlation coefficient of $r=.88$. The exact-adjacent ratio was 16/29 (55%) for the pre-test and 14/29 (48%) for the post-test. The pre-test sample is approximately normally distributed around the true mean, while the post-test sample distribution is more spread out and negatively skewed (see figs. 7 and 8). The entire scale is used. The rubric used was the CLA+ sample rubric for their test.

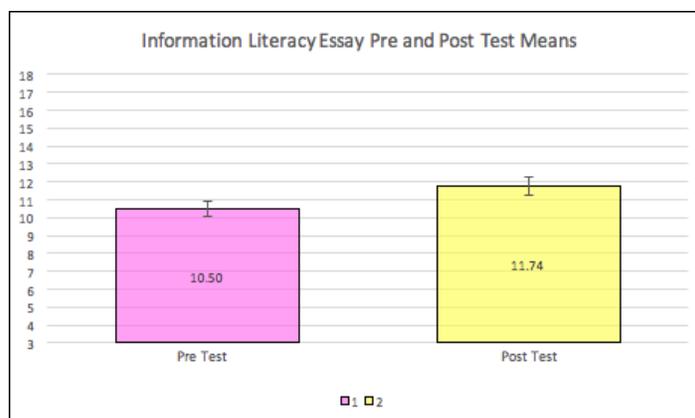


Fig. 6

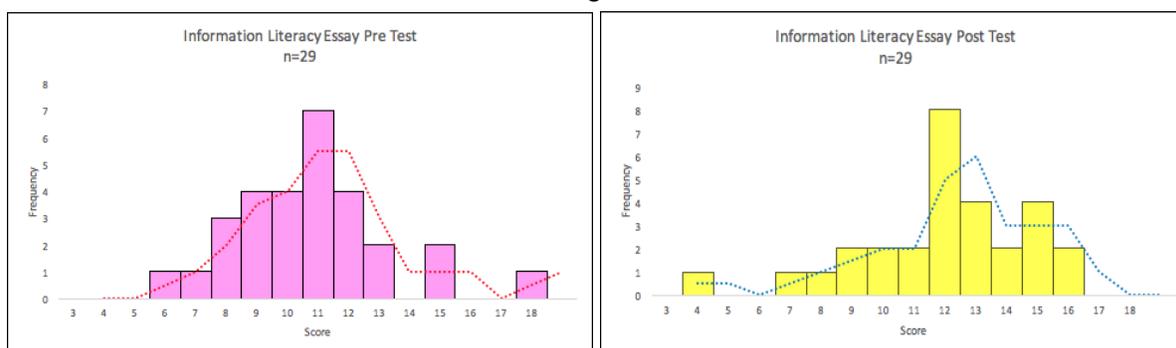


Fig. 7

Fig. 8

Discussion

The Information Literacy Essay was designed to serve as the “control” against which other assignments could be compared. The assignment is adapted from an item on the CLA+, a standardized test of college learning, and by administering it once at the beginning of the semester and again at the end, the assignment offers insight into how much student writing as traditionally defined and understood by outside stakeholders grows during the entire semester. The effect size ($d=.48$) is moderate, but the statistically significant growth in mean scores is promising. The assignment is also easy to administer and not very disruptive, as it is a timed, in-class essay that takes up only two class periods during the semester. However, it should be noted that timed writing isn’t always the best instrument for capturing writing ability, and some question the authenticity of such tasks. Per the rater discussion, challenges included determining how much student role-playing, as directed by the assignment’s prompt, influences writing quality, as well as how much essay length should factor into scores because of the time constraints.

4. Research Lit Review

The Research Lit Review sample ($N=30$) averaged a score of 12.82 ± 2.44 out of 18 on a 13 point scale (6-18), slightly above the scale’s true mean of 12. The Pearson’s correlation for ratings was strong ($r=.85$). The exact-adjacent ratio was 19/30 (63%), which means 11/30 (37%) differed by more than two scale points and required a third rating. The distribution is

negatively skewed and most of the 13 point scale was used (fig. 9). The rubric used was adapted from an instructor's.

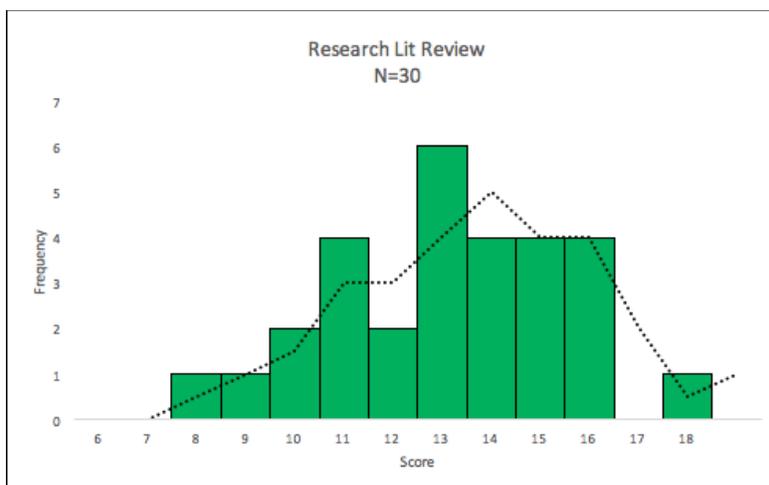


Fig. 9

Discussion

The Research Lit Review is a good assignment for its versatility and practicality, as well as its applicability to disciplines across campus. While the correlation coefficient is strong, the low exact-adjacent ratio (63%) suggests there was more inconsistency than might appear. Per discussion among raters, the major challenge for this assignment was trying to assess multiple parts/sections of the assignment (abstract, synthesis, summary, etc.) with one rubric. In particular, some raters expressed difficulty with not treating the rubric and assignment structure as simply a checklist, which can impact ratings. The ICaP team has discussed the benefits of requiring this assignment to be structured as more of a singular, cohesive essay.

5. Reading Annotations

The Reading Annotations sample (N=30) averaged a score of 8.3 ± 1.66 out of 12 on a 10 point scale, almost one scale point above the true mean of 7.5. The Pearson's correlation for ratings was moderately strong ($r = .70$). Further interrater reliability measures report 22/30 (73%) of essays differed by two or fewer scale points, while 8/30 (27%) differed by greater than two scale points and required a third rating. The distribution is negatively skewed and somewhat constrained to the middle of the scale (fig. 10). The rubric used was an ICaP outcomes-built rubric.

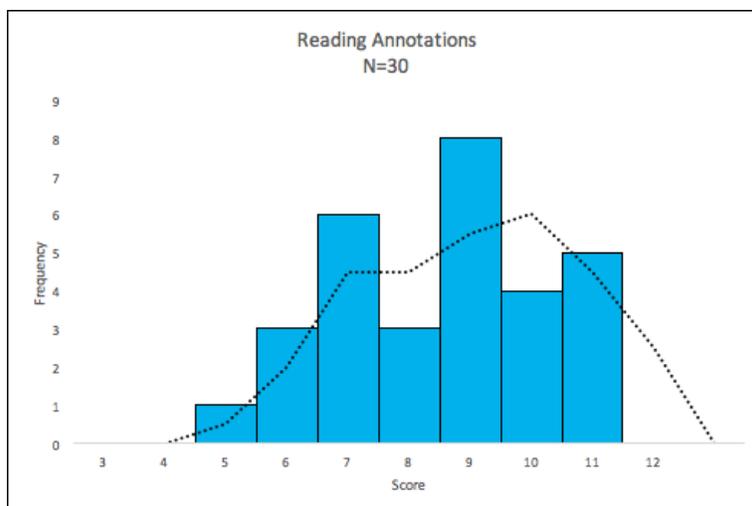


Fig. 10

Discussion

The Reading Annotations assignment is an interesting counterpart to the Research Lit Review. Both assignments attempt to measure students' research skills, critical reading and thinking performance, and synthesis abilities, and so it is also a versatile, practical, and valuable assignment across disciplines. Like the Research Lit Review, this assignment also had several distinct sections/mini assignments within it (summary, quote-pull, synthesis) that posed problems for raters who had only one rubric with which to rate it. Some raters had difficulty not using the rubric like a checklist. The distribution is somewhat compressed, suggesting perhaps a different rubric or assignment design could better represent the traits in question.

6. Portfolio

The Portfolio sample (N=20) averaged a score of 12.38 ± 1.65 on a 13 point scale (6-18), slightly above the scale's true mean of 12. The Pearson's correlation was moderate ($r = .62$). Further interrater reliability measures report 17/30 (85%) essays differed by two or fewer points, while 3/20 (15%) differed by greater than 2 points and required a third rating. The distribution is negatively skewed and constrained to the middle of the scale (fig. 11). The rubric used was an ICaP outcomes-built rubric.

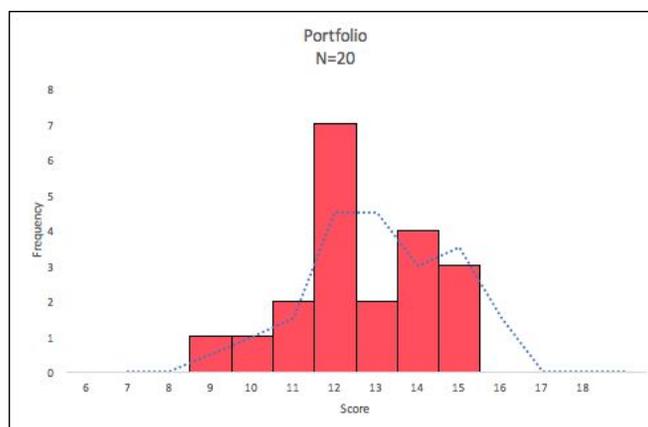


Fig. 11

Discussion

The benefits of the Portfolio assignment are that it is not very disruptive to the rest of the course and is a fairly common feature of FYC courses. The Portfolio also emphasizes revision and reflection, which offer a way for reviewers to assess student growth throughout the semester. The distribution is fairly constrained to the middle of the scale, suggesting the rubric may need revision. Per discussion among raters, there was at times difficulty using one rubric to assess the many moving parts of a portfolio featuring all major assignments throughout the semester. In addition, part of the rubric aimed at assessing revision and reflection requires evidence of revision from rough to revised drafts of writing, and participating instructors did not require submission of rough drafts for the Portfolio. Future versions of this assignment should require the submission of rough and revised drafts of writing.

Recommendations and Next Steps

After piloting the common assignments for one semester and assessing the student writing produced from each assignment, we have already made a number of changes for the future, and have further recommendations for next steps with the assessment project.

Common Assignment Structure for Fall 2018

By the end of the Spring 2018 semester, only two of the common assignment pilots had been assessed. Yet we wanted to announce the Fall 2018 requirements before instructors dispersed for the summer. Rather than make decisions about which common assignments to move forward without first seeing results from assessing the pilots, we decided to conduct a second generation of pilots for Fall 2018 using the same basic assignments as Spring 2018. This has allowed us to complete assessments of each pilot during the summer, then make small revisions to assignment requirements for fall instructors.

In Fall 2018, all instructors will be required to participate in the second generation of common assignment pilots. Instructors must choose one common assignment pilot to integrate into their syllabus and submit their students' writing to ICaP for assessment.

Common Assignment Pilot Changes

Based on our Summer 2018 assessment and feedback from Spring 2018 pilot instructors, we have made the following changes to each assignment:

1. Changes for All Assignments
 - a. Use outcomes-based rubrics for all assignments to simplify and strengthen the norming process during assessment.
 - b. Make assignment requirements more firm to decrease variance by instructor.

2. Professional Email

- a. Made requirements more firm to limit discrepancies in what “professionalism” means for emails. For instance, humorous approaches to the assignment are not allowed, length and audience expectations are clearer, etc.
- b. Moved from a flexible points-based rubric to an outcomes rubric. This should shrink the rating scale to improve norming during assessment, as well as clarify expectations for instructor grading.
- c. Clarified distinct audiences for the assignment’s two email scenarios. Where before one email was supposed to be “formal” and one “informal” to a professor, it proved difficult to distinguish the two during rating. We are now requiring one email to a professor and one to address sensitive group work issues, which should make the two emails more easily distinguishable and integrate a valuable professionalism skill to the assignment.
- d. Decided not to shift to a pre- and post- test format. This idea was raised as a possibility to make the professional email assignment more robust, but we determined that the time and length of the assignment does not make the added pre-test worthwhile.

3. Rhetorical Analysis

- a. Required that instructors use the common outcomes rubric. Instead of providing typical rubric examples that are disconnected from the outcomes-based rubric used for assessment, we have integrated outcome rubric requirements into classroom rubric formats and strengthened requirements to adhere to rubric details.
- b. Maintained the pre and post test sequence, allowing for analysis of two different texts or for the post test to be a revision of the pre test. We found that these differences did not affect assessment, so it is worthwhile to maintain instructor autonomy.
- c. Clarified expectations in instructor guide that the analysis must be rhetorical, not literary. This should strengthen rating scores for rhetorical analyses working with literary texts.
- d. Maintained instructor independence on choice of texts for analysis. Though there were issues with literary essays inadequately utilizing rhetorical tools, rather than eliminate a very popular assignment focus, we want to provide more resources for instructors to teach analysis with literature while meeting ICaP outcomes.
- e. Provided more instructor support for how to teach rhetorical analysis at Convocation.

4. Information Literacy Essay

- a. Made no significant changes. Because this assignment is adapted from a prompt featured on the CLA+, a standardized test of college learning, it serves as a useful “control” against which the other assignments might be compared. It is thoroughly validated by the testing company.

- b. Made one logistical adjustment: perhaps instructors might choose which of the two prompts they designate as the “pre-test” and the “post-test.” For the pilot we required the prompt about shampoo serve as the pre-test and the prompt about roller skates the post-test, but as long as the instructors report to ICaP which one they assigned as which the order shouldn’t matter.
5. Research Lit Review, now Research Essay
 - a. Required the assignment to function as a single, cohesive research essay, instead of just a literature review. To be renamed “Research Essay.”
 - b. Made requirements in both the instructor guide and the rubric more “loose” so that the assignment remains flexible for different kinds of syllabus approaches, disciplines, and research projects.
 - c. Required the use of an outcomes-based rubric, built from outcome 5.
 - d. Note that annotated bibliographies do not meet the criteria for this assignment model, as they are not integrated or cohesive essays.
6. Reading Annotations
 - a. After reviewing all six assignments, ICaP administrators decided to eliminate this common assignment from those offered. The assignment attempts to measure much of what the research essay measures, and the assignment’s multiple moving parts made it difficult for raters to assess. The assignment seems to work well in the classroom; however, it is difficult for ICaP to usefully assess at a program-wide level.
7. Portfolio
 - a. Made the inclusion of rough drafts required, since the assessment of both rough and final drafts shows evidence of revision and reflection.
 - b. Maintained the inclusion of a reflection essay assigned at the end of the semester in which students reflect on how the class helped them meet ICaP outcomes. The essay should be an integrated, cohesive essay and not a list of responses.
 - c. Required the submission of any digital or multimedia texts produced during the course.
 - d. Required use of an outcomes-based rubric, built from all six of the ICaP outcomes.

Rating Session Recommendations

Achieving reliable ratings and properly validating instruments is vital to obtaining useful and accurate assessment data. However, with limited funds and sources of labor, it remains difficult to conduct ideal norming sessions. Trial running different rating configurations proved instructive. Below are some observations and recommendations for future assessment projects involving norming and rating given limited resources:

1. The Assessment Research Coordinator, in conjunction with the Director and Assistant Director of ICaP, as well as any other interested parties, should continually refine any rubrics used for common assignment assessment on an ongoing basis. This report determined that building rubrics out of the ICaP outcomes is the best starting place, but as assessment continues, these rubrics should be updated and refined in accordance with feedback gained during rating sessions and any changes made to assignments.
2. The ICaP Assessment Committee should be further developed and made official with permanent members. During the pilots, raters participated on a volunteer basis so that each rating session featured different raters. However, several volunteers participated more than once, and the continued participation made norming easier. Continued practice at rating generally results in more reliable ratings.
3. If time and resources permit, the raters should spend more time norming on a bigger sample of practice essays before moving to the actual sample. For the pilots, we spent approximately 1-1.5 hours norming and practicing on 4-5 essays, but ideally raters would spend 2-3 hours at minimum practicing on approximately 10 essays.
4. Once program-wide participation in common assignments is mandatory, sample size should increase and greater randomness be strived for. Since these pilots were voluntary, there were significant disparities in participation among assignments. For example, we had 16 instructors participate in the Professional Email assignment, while only 4 participated in the Portfolio assignment. A more random sample is one in which we could randomly select 6-7 participating instructors, from a field of say 10, which would yield us a sample size of 120-140 student essays from 6-7 instructors, as opposed to 60 student essays from the only 3 instructors participating in another pilot.
5. Quantitative data should be collected whenever possible, particularly via rating sessions, but should also be interpreted with appropriate caution and recognition of potential threats to validity. Qualitative data should also be collected using a variety of methods--interviews, focus groups, etc.--and used to help interpret the quantitative data responsibly.

General Recommendations

Continued assessment initiatives at Purdue, via common assignments or otherwise, should begin by clarifying their primary goal(s) at the outset. The idea behind this iteration of common assignments was to provide a degree of consistency across ICaP, per the CWPA recommendation, while also preserving a good amount of instructor autonomy, a feature which has long defined the program here. But such a form of assessment might not be the best method for assessing student ability or growth, for example. Establishing one or two primary goals at the outset is absolutely essential.

Going forward with five of these common assignments, the Assessment Research Coordinator has some general suggestions. First, the rubrics should be continually revised, updated, and debated. After experimenting with different rubric types and scales during the pilots, we decided

that for the next generation of common assignments, all instructors should use the same base rubric that the ICaP assessment team will use for rating, but they will be allowed to add onto that base rubric. Additionally, while the rubrics used will be relatively uniform (depending on how much instructors customize them) the assignment guidelines will remain flexible, so long as instructors stay within the basic genre, so that instructors can teach, for example, the research essay in a way consistent with their unique expertise. The common rubric will help ICaP see the strengths and weaknesses of different syllabus approaches.

Second, while eventual culling may occur, offering multiple common assignment options (currently five) seems like a good way to preserve some instructor autonomy. Undoubtedly, some assignments will be more popular than others, but with approximately 80-90 instructors participating, each of the five assignments should get enough participation to yield sufficient sample sizes for programmatic assessment. In addition, the diversity of assignment styles should be maintained. Currently, two of the five assignments are structured as pre- and post-tests, which can help to demonstrate valuable statistical growth in particular areas of student learning and writing ability. For assignments designed as one-offs, subscores should be investigated for a more thorough quantitative picture and to see which outcomes from which the rubrics are built students excel or struggle in.

Thank You

Thank you to everyone involved, in particular all the ICaP instructors who volunteered to teach the common assignments in their courses. Thanks also to PIC and instructors involved in developing and leading the common assignment pilots, as well as all the raters and readers.

Appendix A

Professional Email Rubric, Instructor Submitted; 16-point scale (5-20)

Criteria	Points: 1 - 4
Genre Considerations: Emails include clear, specific subject lines; email signature is present; closing and sign-off is present and appropriate to the rhetorical situation of each email.	
Purpose: Each email includes a distinct “ask”; emails avoid vague language; emails anticipate information recipients may need and topics are covered in sufficient depth.	
Audience Awareness: Each email demonstrates a clear understanding of its intended audience; forms of address and tone may change between emails as appropriate.	
Structure/Organization: Emails are of an appropriate length, with clear organization around the situation; emails have a clear beginning, middle, and end.	
Line-level Considerations: Emails are devoid of typos or errors that impact readability.	
Overall Rating	/20

Appendix B

Rhetorical Analysis Rubric, Built from ICaP Outcomes; 10-point scale (3-12)

	1 (Poor)	2 (Fair)	3 (Good)	4 (Excellent)
Rhetorical Knowledge & Concepts	The writer does not identify or discuss rhetorical concepts.	The writer identifies a few key rhetorical concepts, including but not limited to: purpose, audience, context/setting, constraints, logos, ethos, pathos, kairos.	The writer identifies and discusses many key rhetorical concepts, including but not limited to: purpose, audience, context/setting, constraints, logos, ethos, pathos, kairos.	The writer identifies and discusses in detail a variety of key rhetorical concepts, including but not limited to: purpose, audience, context/setting, constraints, logos, ethos, pathos, kairos.
Cultural Context & Situation	The writer does not mention or analyze the text's rhetorical situation and context and fails to demonstrate how cultural factors affect both production and reception of ideas.	The writer mentions the text's rhetorical situation and context and attempts to describe how cultural factors affect both production and reception of ideas.	The writer adequately analyzes the text's rhetorical situation and context and demonstrates an understanding of how cultural factors affect both production and reception of ideas.	The writer insightfully analyzes the text's rhetorical situation and context and demonstrates a deep understanding of how cultural factors affect both production and reception of ideas.
Critical Thinking and Analysis	The writer does not describe or synthesize their analysis of rhetorical choices, ideas, information, design, and other components featured in the text. The writer fails to connect these components to the text's intended audience(s).	The writer describes their analysis of rhetorical choices, ideas, information, design, and other components featured in the text. The writer describes some connections between these components and the text's intended audience(s).	The writer synthesizes their analysis of rhetorical choices, ideas, information, design, and other components featured in the text by connecting these components to the text's intended audience(s).	The writer cohesively synthesizes their analysis of rhetorical choices, ideas, information, design, and other components featured in the text by insightfully connecting these components to the text's intended audience(s).

Appendix C

Information Literacy Rubric, from the CLA+; 16-point scale (3-18)

Score	1	2	3	4	5	6
<p>Analysis and Problem Solving</p> <p>Making a logical decision or conclusion (or taking a position) and supporting it by utilizing appropriate information (facts, ideas, computed values, or salient features) from the Document Library</p>	<p>May state or imply a decision, conclusion, position</p> <p>Provides minimal analysis as support (e.g., briefly addresses only one idea from one document) or analysis is entirely inaccurate, illogical, unreliable, or unconnected to the decision, conclusion, position</p>	<p>States or implies a decision, conclusion, position</p> <p>Provides analysis that addresses a few ideas as support, some of which is inaccurate, illogical, unreliable, or unconnected to the decision, conclusion, position</p>	<p>States or implies a decision, conclusion, position</p> <p>Provides some valid support, but omits or misrepresents critical information, suggesting only superficial analysis and partial comprehension of the documents</p> <p>May not account for contradictory information (if applicable)</p>	<p>States an explicit decision, conclusion, position</p> <p>Provides valid support that addresses multiple pieces of relevant and credible information in a manner that demonstrates adequate analysis and comprehension of the documents; some information is omitted</p> <p>May attempt to address contradictory information or alternative decisions/ conclusions/positions (if applicable)</p>	<p>States an explicit decision, conclusion, position</p> <p>Provides strong support that addresses much of the relevant and credible information, in a manner that demonstrates very good analysis and comprehension of the documents</p> <p>Refutes contradictory information or alternative decisions/conclusion s/positions (if applicable)</p>	<p>States an explicit decision, conclusion, position</p> <p>Provides comprehensive support, including nearly all of the relevant and credible information, in a manner that demonstrates outstanding analysis and comprehension of the documents</p> <p>Thoroughly refutes contradictory evidence or alternative decisions/conclusion s/positions (if applicable)</p>
<p>Writing Effectiveness</p> <p>Constructing organized and logically cohesive arguments. Strengthening the writer's position by providing elaboration on facts or ideas (e.g., explaining how evidence bears on the problem, providing examples, and emphasizing especially convincing evidence)</p>	<p>Does not develop convincing arguments; writing may be disorganized and confusing</p> <p>Does not provide elaboration on facts or ideas</p>	<p>Provides limited, invalid, overstated, or very unclear arguments; may present information in a disorganized fashion or undermine own points</p> <p>Any elaboration on facts or ideas tends to be vague, irrelevant, inaccurate, or unreliable (e.g., based entirely on writer's opinion); sources of information are often unclear</p>	<p>Provides limited or somewhat unclear arguments. Presents relevant information in each response, but that information is not woven into arguments</p> <p>Provides elaboration on facts or ideas a few times, some of which is valid; sources of information are sometimes unclear</p>	<p>Organizes response in a way that makes the writer's arguments and logic of those arguments apparent but not obvious</p> <p>Provides valid elaboration on facts or ideas several times and cites sources of information</p>	<p>Organizes response in a logically cohesive way that makes it fairly easy to follow the writer's arguments</p> <p>Provides valid elaboration on facts or ideas related to each argument and cites sources of information</p>	<p>Organizes response in a logically cohesive way that makes it very easy to follow the writer's arguments</p> <p>Provides valid and comprehensive elaboration on facts or ideas related to each argument and clearly cites sources of information</p>
<p>Writing Mechanics</p> <p>Demonstrating facility with the conventions of standard written English (agreement, tense, capitalization, punctuation, and spelling) and control of the English language, including syntax (sentence structure) and diction (word choice and usage)</p>	<p>Demonstrates minimal control of grammatical conventions with many errors that make the response difficult to read or provides insufficient evidence to judge</p> <p>Writes sentences that are repetitive or incomplete, and some are difficult to understand</p> <p>Uses simple vocabulary, and some vocabulary is used inaccurately or in a way that makes meaning unclear</p>	<p>Demonstrates poor control of grammatical conventions with frequent minor errors and some severe errors</p> <p>Consistently writes sentences with similar structure and length, and some may be difficult to understand</p> <p>Uses simple vocabulary, and some vocabulary may be used inaccurately or in a way that makes meaning unclear</p>	<p>Demonstrates fair control of grammatical conventions with frequent minor errors</p> <p>Writes sentences that read naturally but tend to have similar structure and length</p> <p>Uses vocabulary that communicates ideas adequately but lacks variety</p>	<p>Demonstrates good control of grammatical conventions with few errors</p> <p>Writes well-constructed sentences with some varied structure and length</p> <p>Uses vocabulary that clearly communicates ideas but lacks variety</p>	<p>Demonstrates very good control of grammatical conventions</p> <p>Consistently writes well-constructed sentences with varied structure and length</p> <p>Uses varied and sometimes advanced vocabulary that effectively communicates ideas</p>	<p>Demonstrates outstanding control of grammatical conventions</p> <p>Consistently writes well-constructed complex sentences with varied structure and length</p> <p>Displays adept use of vocabulary that is precise, advanced, and varied</p>

Appendix D

Research Lit Review Rubric, Instructor Submitted; 13-point scale (6-18)

Category	Developing (1)	Acceptable (2)	Exceeds (3)
<p>Thesis and Focus</p> <p>Does the literature review essay DEMONSTRATE A SCOPE OF LITERATURE? <input type="checkbox"/> Review summarizes and analyzes 6 sources in detail including author, title, year, and key arguments made in each of the essay.</p>			
<p>Organization</p> <p>Is the essay organized so the reader can easily sense a pattern of thought? <input type="checkbox"/> Essay has a standard introduction, body, and conclusion <input type="checkbox"/> Essay uses transitions between ideas/sections to keep the essay moving <input type="checkbox"/> Essay uses contrasting transitions to denote when essays have points of disagreement <input type="checkbox"/> Similar issues are grouped together, not scattered randomly through the essay</p>			
<p>Development/Research</p> <p>Are scholarly sources used effectively to support a specific topic? Are sources interpreted? <input type="checkbox"/> At least two paraphrases or two quotes from each of the essays being analyzed are used effectively to support/develop a main claim about the essay <input type="checkbox"/> The first time a new source is used, some introduction is given (the first and last name of the author, article title, and year). <input type="checkbox"/> Writer puts the six essays "in conversation" by highlighting key trends, disagreements, and controversies over the topic</p>			
<p>Voice</p> <p>Do most works have a cohesive voice? Is language clear? <input type="checkbox"/> Essay makes distinctions between writer's voice and the author he or she is analyzing in the Revised Analysis assignment <input type="checkbox"/> Pronouns have a clear referent (no unclear "this" or "he" or "it")</p>			
<p>Mechanics</p> <p>Does the essay have...? <input type="checkbox"/> Standard English usage <input type="checkbox"/> Correct Spelling, punctuation, capitalization, and documentation <input type="checkbox"/> Few or no run on sentences or fragments</p>			
<p>Documentation</p> <p>Does the essay have...? <input type="checkbox"/> A correct Works Cited page in APA format <input type="checkbox"/> Correct APA format for the whole paper, including headings and page numbers and spacing between paragraphs <input type="checkbox"/> An abstract <input type="checkbox"/> Anything that should be cited is cited (no plagiarism) <input type="checkbox"/> Correct APA in-text citations</p>			

Appendix E

Reading Annotations Rubric, Built from ICaP Outcomes; 10-point scale (3-12)

	1 (Poor)	2 (Fair)	3 (Good)	4 (Excellent)
Integrating ideas	The writer does not integrate sources with their own ideas.	The writer occasionally interprets, synthesizes, responds to, or critiques sources, but efforts are inconsistent or inadequate to integrate with the writer's own ideas.	The writer regularly interprets, synthesizes, responds to, and/or critiques sources to integrate with the writer's own ideas, though connections may be at the surface level.	The writer comprehensively interprets, synthesizes, responds to, and/or critiques sources to deeply integrate with the writer's own ideas.
Evaluating sources	The writer does not evaluate sources or use them to support claims.	The writer incompletely evaluates sources for credibility, sufficiency, accuracy, timeliness, and bias, or uses them inappropriately to support claims.	The writer adequately evaluates sources for credibility, sufficiency, accuracy, timeliness, and bias, and uses them to support claims.	The writer comprehensively evaluates sources for credibility, sufficiency, accuracy, timeliness, and bias, and uses them to robustly support claims.
Citing sources	The writer does not apply citation conventions for sources.	The writer inconsistently and incorrectly applies citation conventions for sources.	The writer correctly applies citation conventions with some mistakes.	The writer successfully and consistently applies citation conventions for sources.

Appendix F

Portfolio Rubric, Built from ICaP Outcomes; 13-point scale (6-18)

Criteria	Fails to Meet (1)	Meets (2)	Exceeds (3)
Demonstrate rhetorical awareness of diverse audiences, situations, and contexts.			
Compose a variety of texts in a range of forms, equaling at least 7,500-11,500 (25-39 pages) words of polished writing			
Critically think about writing and rhetoric through reading, analysis, and reflection.			
Provide constructive feedback to others and incorporate feedback into their writing			
Perform research and evaluate sources to support claims.			
Engage multiple digital technologies to compose for different purposes.			

Appendix G

Rhetorical Analysis Practice Sample Individual Rater Correlations

Raters	A	B	C	D	E	F	G	H	I
A	1.00								
B	0.96	1.00							
C	-0.42	-0.38	1.00						
D	0.97	0.99	-0.51	1.00					
E	0.97	0.89	-0.30	0.89	1.00				
F	-0.25	0.00	0.45	-0.12	-0.38	1.00			
G	0.98	0.96	-0.25	0.95	0.97	-0.15	1.00		
H	0.88	0.76	-0.69	0.84	0.88	-0.63	0.80	1.00	
I	0.95	0.87	-0.62	0.93	0.93	-0.48	0.89	0.98	1.00

Rhetorical Analysis Practice Sample Intraclass Correlation

Raters	A	B	C	D	E	F	G	H	I
B009-89 53	7	7	6	7	6	7	6	5	6
B009-37 57	9	8	4	9	8	5	7	10	8
B032-75 88	4	4	5	4	3	6	3	4	5
B032-66 27	6	5	6	5	6	5	5	6	6
								ICCorr	0.72