2017-2018 Philosophy Assessment Report

by David Richardson, Unit Assessment Liaison
(with support from Fernando Miranda-Mendoza, Data Analyst)
Harold Washington College
Humanities Department
May 1, 2018

Assessment History & Purpose

In fall 2016 an email poll of philosophy faculty members identified "student critical reading abilities and reading/learning beliefs" as a consensus choice for assessment across all HWC Philosophy courses.

The direct assessment portion was modeled after and adapted from a critical reading assessment tool commonly used in college and graduate school readiness tests and associated test-prep books. It featured a direct assessment of reading skills through student engagement with an argumentative passage from a journal article and questions related to our specific learning goals in philosophy classes. That was followed by a survey of student learning behaviors while reading and a survey of student learning and reading beliefs, all rooted in the research literature related to literacy and reading instruction.

This measure was piloted toward the end of the semester in the fall of 2016 across five philosophy classes taught by full-timers and adjuncts. Data from the pilot showed "robust" reliability and validity in the measure (the assessment set a new high-score for reliability as based on a Cronbach alpha test). Two questions were adjusted based on Point Biserial scores, and I collected feedback from instructors leading to a suggested 30-minute time frame for the administration of the measure. In Week 13 of fall 2017, the assessment was distributed to all of our philosophy faculty along with instructions for them and for students, including an informed consent statement that made clear that both instructor and student participation was entirely voluntary.

Sample Details

Four philosophy instructors (out of five) offered the assessment to students, resulting in 143 returned assessments from various classes across our curriculum. Of those 134 included valid student ID numbers, allowing for analysis. Initial enrollment for the semester was 594 registrations for seats in our 19 sections of face-to-face classes (the unduplicated student numbers were not available to me), meaning that our sample was at least 22% of the initial philosophy enrollment (and likely higher). The assessment was also digitized and offered by at least two (of our five) online instructors in some of our 11 online sections, however, no online students chose to complete the assessment. Nonetheless, the size of the response pool and the variability of the sources suggest reason to have good confidence that the sample is a representative one for the face-to-face classes, at least.

Course history analysis of the sample showed that of the 134 valid samples:

- 75.4% of the students (n=101) had NOT taken or successfully completed a philosophy class previously;
- 17.9% (n=24) had previously, successfully completed ONE philosophy class;
- 6.7% (n=9) had previously, successfully completed TWO or more philosophy classes; and

Post-term analysis showed that of the 134 valid samples

- 82.1% (n=110) successfully completed (C or better) the class in the weeks following the assessment;
- 17.9% (n=24) did NOT successfully complete the class in which they took the assessment.

Students were also divided into those who were taking Philosophy 105 and those who were in other, reading/text-based philosophy classes where these elements are more likely to be directly taught, but I do not have those percentages and we did not learn anything helpful from having made the distinction.

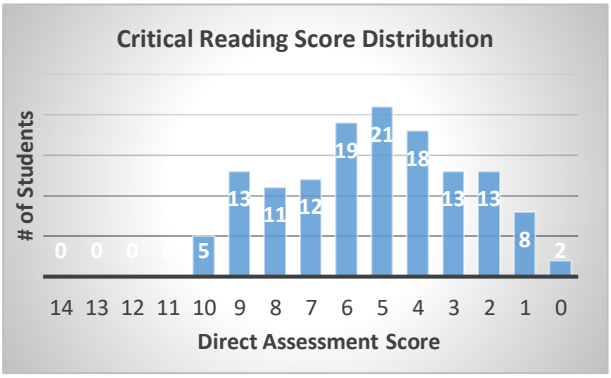
Generally, the numbers for the main groups in each of the two aforementioned analyses tracked the numbers for the full sample without significant differences, and differences in the small groups were likely the result of statistical noise. Any statistically significant differences between the groups will be noted in the analysis that follows, but most of the discussion will focus on the sample as a whole.

Data Analysis

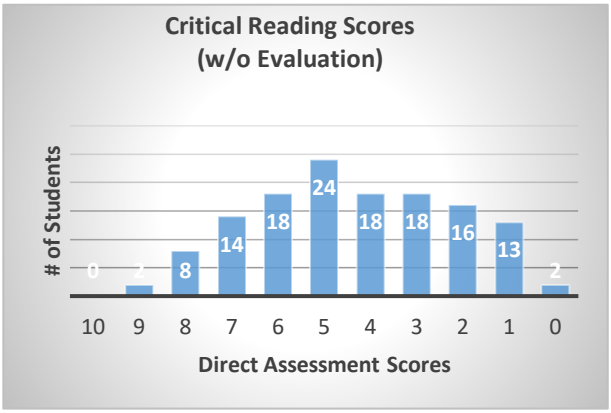
The tri-partite structure of the assessment created three distinct buckets of data related to: 1) Direct measurement of students' critical reading skills; 2) Indirect measurement of student reading behaviors; and 3) Indirect measurement of student beliefs about reading and learning.

1) Critical Reading Skills: The critical reading questions included questions testing Argument Comprehension, Analysis, Inference, and Evaluation.

Student scores on the 14 direct assessment questions included in the survey were distributed as shown on the graph to the right. The mean score was 5.3 (or 37.9% correct), the median was 5 (or 35.7% correct), and the mode was also 5.



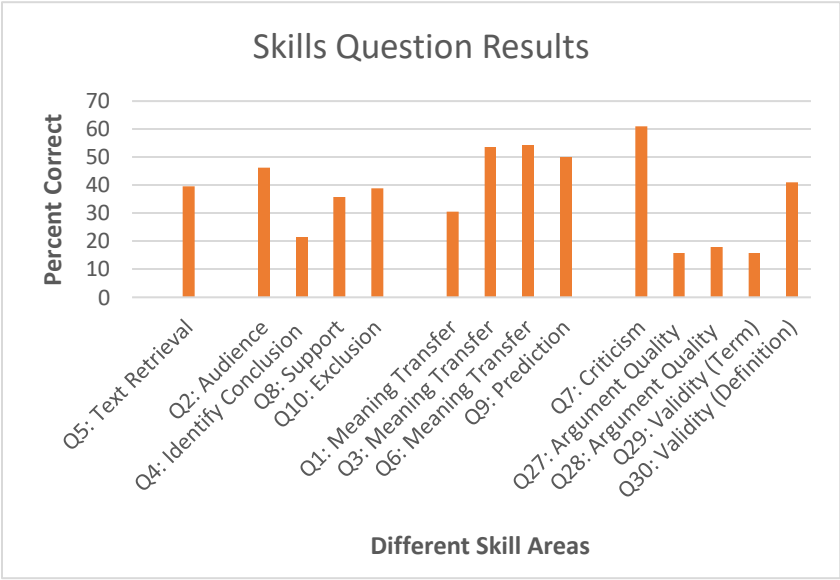
The scores improve, however, in terms of percentage correct, and are more normally distributed, when the evaluation questions related the argument's strength and validity are removed from consideration as shown in the second graph. The mean score rises to 4.4 (or 44% correct), the median to 4 (or 40% correct) and the mode is also higher, in terms of percent correct.



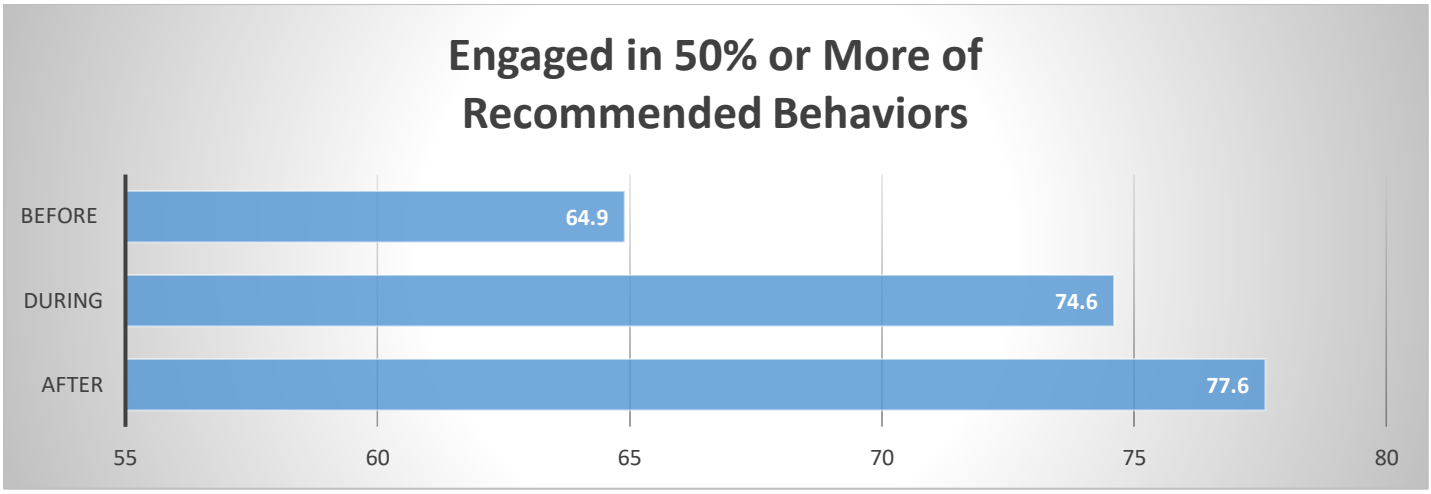
Since only about 1 in 5 students correctly identified the conclusion of the argument featured in the excerpt, and less than that evaluated it correctly as well as other reasons (see DISCUSSION below) While those numbers are rather low, it is important to keep in mind that the text and questions were adapted from an LSAT question set, and so it's not shocking to find out that first and second year students had trouble with it.

There were two statistically significant differences among groups on specific questions: students with two or more previous philosophy classes were significantly MORE likely to correctly evaluate the argument quality and students with one previous class were significantly LESS likely to correctly identify the conclusion (p-values unavailable).

Critical Reading Skill	% Correct
Comprehension	**39.6%**
Q5: Text Retrieval	39.6%
Analysis	**35.6%**
Q2: Audience	46.3%
Q4: Identify Conclusion	21.6%
Q8: Support	35.8%
Q10: Exclusion	38.8%
Inference	**47.2%**
Q1: Meaning Transfer	30.6%
Q3: Meaning Transfer	53.7%
Q6: Meaning Transfer	54.4%
Q9: Prediction	50.0%
Evaluation	**30.3%**
Q7: Criticism	61.1%
Q27: Argument Quality	15.7%
Q28: Argument Quality	17.9%
Q29: Validity (Term)	15.7%
Q30: Validity (Definition)	41.0%



2) Student Reading Behaviors (self-reported): Research-based useful reading behaviors (for comprehension and recall) are distinguished as things readers (should) do before, during, and after engaging in reading tasks. Literacy research suggests that students do best when they learn and practice reading as a staged, recursive process and are taught what to do, when to do it, and how to do it.



Before" Behaviors include:

- Activating prior knowledge
- Previewing Text
- Defining Reading Purpose
- Using Text Protocol

"During" Behaviors include:

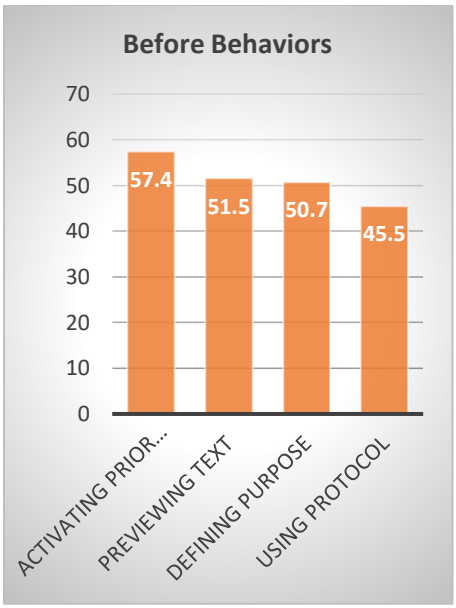
- Annotating/Writing
- Visualizing
- Questioning
- Predicting
- Selectively identifying Key Terms
- Comprehension Monitoring (i.e. Metacognition)
- Connecting (e.g., text-to-text, text-to-self, text-to-world)

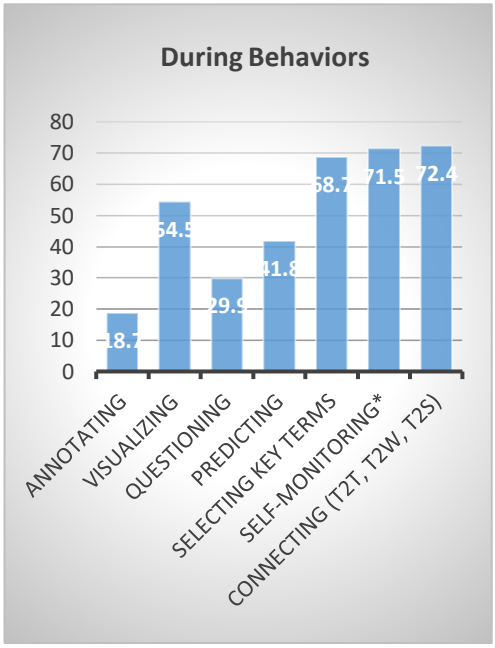
"After" Behaviors include:

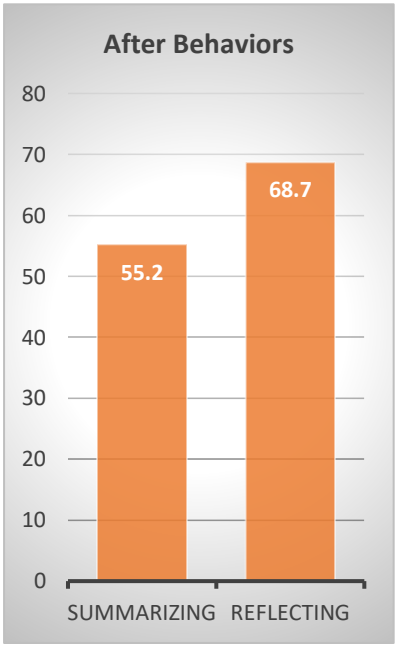
- Summarizing
- Reflecting/Comprehensive Metacognition
- Discussing*
- Elaborating/Extending*
- Writing*

*Not included in assessment

In terms of specific behaviors reported, there were no significant differences among groups. Percentages of students who reported engaging in the specific behaviors listed are displayed below:





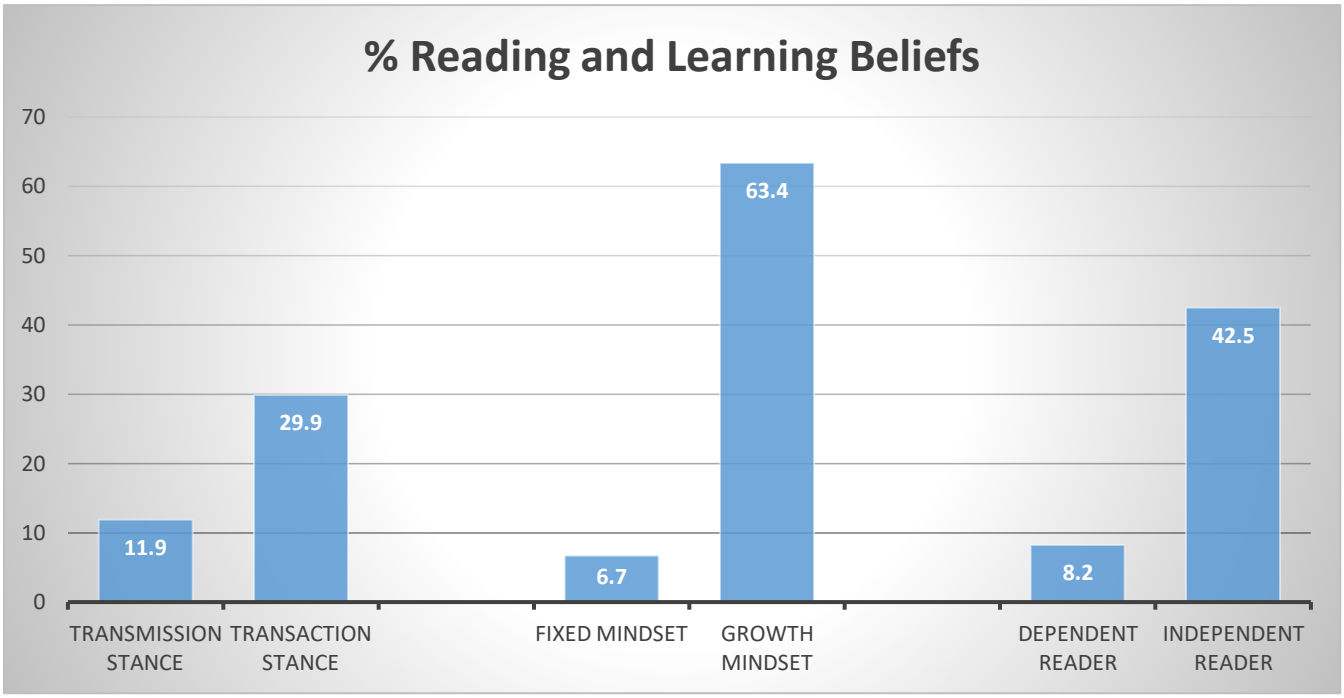


3) Student Beliefs about Reading and Learning (self-reported): Student beliefs about reading and learning have been seen to have powerful impacts on their motivation, execution, and persistence with reading tasks and so affect the outcome. Students were identified as exhibiting beliefs of a specific category if they answered 50% or more of the questions associated with the conceptual area in a way consistent with beliefs of a specific category in that area AND agreed with *less than* 50% of the opposing category. The chart below shows the percentage of students who qualify for the label according to that formula. The three areas of interest each with a pair of categories, explored through this assessment are:

- Reader Stance: This concept relates to reader expectations regarding the production of meaning, captured by the description *Transmission Stance,* indicating the reader expects text to 'transmit' the meaning and, so, must be 'found,' and *Transaction Stance,* which describes the expectations of a reader who believes that meaning is created through a mutual exchange of ideas and is not found in neither the reader nor text alone, but in their interaction).

- Mindset: Made famous by Carol Dweck, this set of beliefs relates to readers' relationship with and interpretation of difficulty or failure and their subsequent likely response. The theory describes these beliefs as creating a specific kind of disposition or "mindset" in which capabilities are understood as *Fixed* or subject to change through work and practice, which are beliefs that indicate a *Growth* mindset.

- Reader Responsibility: Questions related to this concept aim at evaluating the degree of responsibility for making sense of and understanding the text that the student puts on themselves when engaged with a reading task. Readers are classified as either *Independent,* in the case of a student who sets their own goals, reads for their own reasons, develops their own interpretation, and, upon encountering confusion or difficulty, works to develop their own understanding, versus *Dependent* readers who engage with reading tasks passively, give up easily, and are happy to wait for someone else to tell them what the text means.



60% don't fall neatly into either reading stance category, which might mean they are uncertain about how to approach reading or might mean that they know that different stances are appropriate for different reading tasks and skilled readers can do both depending on their needs. Statistically significant differences were found in comparisons of Philosophy 105: Logic students to the rest of the sample with respect to reading stance, with 22.7% showing agreement with taking a "Transmission Stance" in comparison to just 6.67% of the rest (p-value=.007091). Logic students were also more likely to show a Fixed Mindset (13.64%) than the rest (3.33%) (p-value=.02525).

Recommendations

Based on the findings, I recommend:

1. Developing additional text and question pairings with similar aims to seek data that might help clarify if some of the direct assessment difficulties, such as correctly identifying the conclusion, were text-specific problems or consistently challenging for students.

2. It is clear from the results for Q29 and Q30, that students interpret the word “valid” in a non-technical sense, and so instructors across our classes should make a concerted effort to teach and emphasize the technical, argumentative sense of the term and its use, especially given its frequent use in philosophical reading.

3. The data shows that while a majority of students report engaging in productive pre-reading behaviors, there are far fewer than those who are active during and after reading texts. Teaching students that they should set a conscious reading goal and teaching them what to read for in terms of purpose and outcomes, such through the use of an explicit reading goal or text protocol, as well as teaching students how to preview a text for structure and content, is an area we could improve on.

4. In regard to “During” Behaviors, it is striking how few students write, question, and predict while reading (especially while reading a text that they know will answer questions about). Student reticence to write in textbooks they may try to sell is one thing, but encouraging students (through practice and explicit instruction) to write on class texts, especially examinations, is an obvious thing to encourage and improve.