2013-2014 Psychology's Report of

Assessment of Scientific Thinking, Methods, and Skills

Consistent with the College's focus on the institutional level Program Learning Goal (PLG), *Critical Thinking*, the psychology department selected two parts of their PLG, *Scientific Thinking, Methods, and Skills*, that were congruent with Critical Thinking, to assess in the 2013-2014 academic year. The description of the institution's PLG states that "graduates of Westmont College will accurately evaluate the strength of evidence in support of a claim."

In support of this goal, the department was interested in how well its senior students in capstone courses accomplished two tasks related to critical thinking, more broadly defined. First, their abilities to "use disciplinary . . . databases . . . effectively and efficiently to refine research questions in psychology and identify extant answers within the literature . . . " were assessed. In this assessment, students in PSY 111, History and Systems of Psychology, were given a scenario and asked to search PsycINFO, the primary database indexing psychological, professional publications, in order to determine the availability of information that could be used to answer a research question related to the scenario.

The second assessment focused on how well senior psychology majors "demonstrate[ed the] ability to recognize good versus bad experimental designs, theories, and arguments in psychology . . ." In order to assess these abilities, the department reviewed the literature reviews of students who completed PSY 198, Capstone Senior Research in Psychology .

Assessment 1:

Using Databases Effectively

Method

Participants. Eighteen students from PSY 111 (3 males; 15 females), which is a required course for all psychology major seniors, attempted the PsycINFO assessment for up to 3% extra credit in the PSY 111 course (69% response rate).

Materials. A scenario was presented that described a young woman having difficulty controlling her eating behaviors when she was trying to control her reactions to an old boyfriend at the same time. The research suggests that people have a limited amount of energy to control their behaviors and if that energy is used to exert self-control in one area, there may not be enough energy left to successfully control one's

behavior in another area. This scenario was presented in Aronson et al.'s (2014) *Social Psychology* textbook (See Appendix A for the scenario and task) and was shortened a bit for this assessment.

Students were then asked:

Suppose you want to investigate the existence of this phenomenon in other cultures, but you don't want to waste your time replicating other people's research. We would like to know what you would do to research what experimental studies have already been done.

Procedure. Students in PSY 111 were emailed during the last week of the Spring, 2014, semester with the scenario and instructions for completing the task. They were told that they could only use PsycINFO, the primary electronic database for psychology; they had to complete the task before the end of the last day of classes (11:59 pm); and they could only spend 15 minutes on the task to help them be as efficient as possible and minimize levels of stress.

They were asked to turn in a list of the searches that they performed and, based on those searches, a list of articles that they thought were most relevant to understanding what effects self control has on the depletion of psychological resources for people in other cultures.

Rubric. A rubric was developed that allowed assessment of the number and quality of the searches that were performed; the number and quality of the articles that were chosen; and the ability of the students to follow the directions. The levels of achievement selected for the rubric were: No Attempt, Beginning, Developing, Accomplished, and Exemplary. Benchmarks for the 3 assessed dimensions were that 75% of students would perform at or above the Accomplished level (indicated by the red outline in Table 1, below).

Results

Of the 18 students who attempted the database search task, 61% completed it. Although students were told to only spend 15 minutes on the task, not everyone complied. The median time spent on the task was 15 minutes, but the amount of time for each student varied from 7 to 30 minutes (*SD* = .39).

Across all 18 students, the average number of articles identified was 8.8 (*SD* = 7.9), and the average, student-identified articles that actually pertained to cross cultural studies of self control and resource depletion was 56%.

In Table 1, below, one can see the numbers of students who performed at each level of the rubric. Considering the *Number and Quality of Searches Performed* dimension, all of the students performed at least

one search. It was expected that more than one search would be required, given that this particular topic doesn't have a standard set of search terms that would allow one to pull up nearly all relevant articles at once,

Table 1

Rubric for, and Student Performance on, a PsycINFO Search, Evaluated on Three Dimensions and by Level of Development

	No Attempt 0	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
Number & quality of searches performed	No attempt to perform a relevant search; Or performs one search using ineffective search terms	Performs only one search; some effective search terms are used **1**	Performs at most 2 searches using some effective search terms **1**	Performs more than 2 searches, using mostly effective search terms **5**	Performs multiple searches using the most effective search terms **11**
Quality & relevance of articles identified & selected	Articles are at best tangentially related to the topic but not to a cross-cultural perspective **No Articles: 7**	Articles are relevant to the general topic but not to a cross-cultural perspective **4**	Some articles are relevant to a cross-cultural perspective of the topic **3**	The majority of articles are directly relevant and some are critical for understanding research on cross-cultural perspectives for this topic **4**	The majority of articles are critical for understanding research on cross-cultural perspectives for this topic
Ability to follow instructions	Articles or search list is missing but student claims to have completed the task	Articles or search list is uploaded but student does not complete task by date deadline **8**	Articles & search list are uploaded but student does not complete task by date deadline **3**	Articles & search list are uploaded; student completes task by date deadline but takes more than 15 min **3**	Articles & search list are uploaded; student completes task by date deadline and takes no more than 15 min **4**

so it was gratifying to see that 11 of the 18 students performed multiple searches using the most effective search terms available. Another 5 students also performed multiple searches using mostly effective search terms. Only 2 students (11%) fell below the Accomplished level, which met our benchmark for this dimension.

Examining the second dimension of the assessment, *The Quality of the Articles Identified and Selected*, it became clear that 7 of the students (39%) did not finish the task. As we assessed the task itself, we realized that this particular topic is one for which identifying effective search terms is very difficult, and students may have spent all of their time trying to identify productive search terms but then didn't have time to select

relevant articles. In addition, as we evaluated the task, we realized that few to no articles directly address the idea of limited self-control in other cultures. Therefore, the Exemplary level was not really possible.

The performance of those students who did complete the task as best they could, is distributed evenly among the Beginning, Developing, and Accomplished levels. Seven of these students (39%) selected many articles that addressed the intersection of cross cultural research and self-control. Another 4 (22%) found articles on self-control but never searched for both cross-cultural studies and self-control. Thirty-six percent of students who completed the task performed at the Accomplished level.

Finally, we evaluated the *Ability to Follow Instructions* dimension, which also requires a level of critical thinking and discernment that we would want to see in our seniors. Again we saw the effects of the difficulty of the task: Eight students did not complete the task by the last day of classes. These students attempted the task but did not or could not complete it within the required time limit. An additional 3 students completed the task, but did not do so until after the last day of classes.

The other 7 students completed the task before the date deadline. Three took longer than 15 minutes; four, took no more than 15 minutes.

All 18 students did, however, upload a search list or a list of articles as instructed; that is, no one emailed their results to the instructors directly.

On this last dimension, then, the benchmark was difficult to apply, given the varied aspects of following instructions and the difficulty of the task.

Discussion

There are a number of findings in this assessment that are interesting and important. First, the majority of students in PSY 111 were willing to try to complete the task, indicating this was a task that senior students generally thought they could successfully complete. Second, the assessment was not completely effective, given the topic that was chosen and the time limit that was imposed. Third, the majority of students (89%) performed multiple searches, indicating that they persevered (another quality that is rewarding to see) in trying to find articles that satisfied the requirements of the task. For us, this is an important finding because it indicates that students are critically evaluating the results of their searches and continuing to look for articles, rather than stopping with one or two searches. Fourth, of the students who completed the task

(11), 7 (64%) were able to find some articles that pertained to the study of self control in other cultures. Fifth, seniors mostly used reasonable judgment in completing the task according to the instructions provided.

After discussing these results, the department agreed that the assessment results were helpful, given the limitations of the assessment instrument, and concluded that the seniors are accomplished in most aspects of this critical thinking assessment.

Assessment 2: Ability to Recognize Good versus Bad Experimental Designs, Theories, and Arguments

As was described above, the second assessment focused on senior psychology majors' literature reviews in PSY 198, Capstone Senior Research in Psychology, in order to examine their abilities to present and evaluate relevant theories, and link their literature reviews to their experimental studies.

Method

Participants. Seven enrollees (2 males; 5 females) in PSY 198, Capstone Senior Research in Psychology, which some seniors choose to complete as part of their capstone experience, wrote literature reviews for their research studies as part of the course requirements.

Materials. The literature review was described in the PSY 198 syllabus as "a review of the scientific literature (25-35 references) . . ." In addition, one of the instructors of the course described the characteristics of a scientific literature review in a presentation during one of the class periods.

Procedure. The two instructors of the course distributed the 7 literature reviews (with identifying information removed) to department members before meeting to evaluate them. The dimensions on the rubric were examined one at a time. With some discussion, department members achieved a consensus on each literature review for each dimension.

Rubric. A rubric was prepared that allowed department members to assess 1) whether and how well students presented and discussed relevant theories and 2) appropriately linked their literature reviews to their own experimental studies (See Table 2, below). Levels of achievement selected for the rubric were: Absent, Beginning, Developing, Accomplished, and Exemplary. Again, benchmarks were that 75% of students performed at or above the Accomplished level for each dimension.

Results

The results are presented in Table 2, below. As can be seen in the table, no students performed at the Exemplary level for either dimension. Three students (42%) performed at the Accomplished level in

Presenting and Discussing Relevant Theories dimension. But one student performed at the Beginning level and another student, at the Absent level.

Two students (29%) performed at the Accomplished level in the *Appropriately Link Their Literature Reviews to Their Own Experimental Studies* dimension. But again, two students showed no evidence of this ability, and one student performed at a Beginning level.

Table 2

Rubric for, and Student Performance on, a Senior Level Literature Review, Evaluated on Two Dimensions and by Level of Development

Students will demonstrate the ability to:	**Absent 0**	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
present and discuss relevant theories	No evidence of this ability **1**	Identifies theories, but does not discuss **1**	Identifies theories and describes reasons for choosing them but does not flesh them out **2**	Identifies and discusses theories, evaluating various alternatives but doesn't provide rationale for theory that is selected **3**	Identifies theories, comparing and contrasting where appropriate, and evaluating various alternatives; provides rationale for selected theory
appropriately link their literature reviews and their own experimental studies	No evidence of this ability **2**	States that there is a connection or it is clear that the student is assuming a connection but the link is not explicit **1**	States that there is a connection, describes it to some extent but not enough to be compelling **2**	States that there is a connection, describes it in enough detail to be somewhat convincing **2**	States that there is a connection, and describes it in enough detail to be compelling

Discussion

The results of this assessment were quite disappointing. The department expected that these students would perform more highly on both dimensions of this assessment. Some possible reasons for these results that were discussed included the possibility that this sample was small; that this group of students happened to be atypical in their performance; that perhaps seniors aren't getting the training and experience they need in writing literature reviews and linking them to their experimental studies in PSY 197, the first semester of Capstone Senior Research in Psychology; or that the students are not taking the literature review

task seriously, so they procrastinate in PSY 197, during the Fall semester and do not do the research that is required to produce excellent literature reviews.

The department focused on the last 2 possibilities and identified changes in PSY 197 that will be implemented Fall, 2014, in order to provide more training and guidance for this year's students. After applying these changes, the department will consider re-assessing this aspect of critical thinking.

First, the emphasis in the syllabus and conversations with students will be that they focus first on a research topic and question, and only after that, on a method. Many students seem to select a method (e.g., using EEG to record brain activity) and only after that, try to find a research question that allows that particular method to be used. It is expected that a focus on a research topic and question will result in broader reading in various research areas and then greater familiarity with the literature in one area as students narrow their focus.

Second, course meetings will shift from a discussion of weekly chapter readings in ethics and experimental design, to weekly reports of each student's progress in identifying a research topic or question. Students will answer questions about the chapters in ethics and experimental design using an LMS. In addition, students will be required to keep a "journal," identifying and summarizing the literature that they have read each week.

Third, each student will create a timeline in which he or she lays out what needs to be accomplished each week for his or her project. This task, if completed properly, is intended to provide more structure for students, compared to what they have experienced in the course in the past, helping them see more clearly what must be accomplished each week if they are to complete the semester's tasks. In addition, the course syllabus will make it clear when students must identify a research topic, and then a research question.

Adapted from Aronson et al. (2014):

Sarah has vowed to take the high road with her ex-boyfriend Jake, letting bygones be bygones and just forgetting about all the stupid things he did. "That's ancient history," she thinks. "Time to move on." One night she runs into Jake at a party and, wouldn't you know it, there he is with her friend Meghan, whom Jake swore he wasn't interested in. Sarah is sorely tempted to make a scene and tell them both off, but she grits her teeth, puts on her best smile, and acts as if she is the belle of the ball. She is proud of herself, but a little while later she finds herself devouring a bowl of potato chips, even though she had promised herself to eat a healthy diet and shed a few pounds. In this example Sarah is doing something familiar to us all--trying to exert self-control. She succeeds in one respect, by suppressing her desire to tell off her ex-boyfriend, but fails in another, by failing to keep to her healthy diet. What determines how successful we are at exerting self-control?

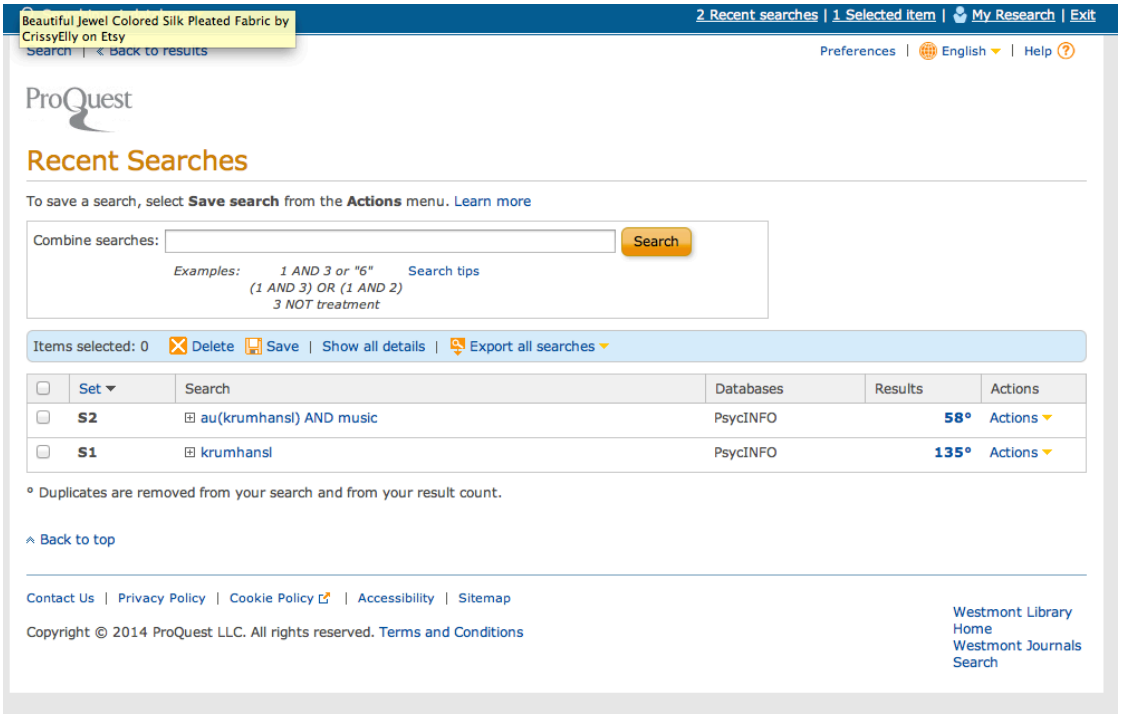
An important function of the self is to be the chief executive who makes choices about what to do in the present and plans for the future (Baumeister, Schmeichel, & Vohs, 2007; Carver & Scheier, 1998; Higgins, 1989, 2005; Vohs & Baumeister, 2011). . . . Regulating our behavior and choices in optimal ways, of course, can be easier said than done, as anyone who has been on a diet or tried to quit smoking knows.

When are we likely to succeed? The answer, according to the self-regulatory resource model, is to make sure that we have plenty of energy when we are trying to control our actions (Baumeister & Hetherington, 1996; Baumeister, Vohs , & Tice, 2007; Schmeichel & Baumeister, 2004). According to this approach, self-control requires energy, and spending this energy on one task limits the amount that can be spent on another task, just as going for a 5-mile run makes it difficult to immediately play a game of basketball afterward. To test this idea, Muraven, Tice, and Baumeister (1998) asked participants to exert self-control on one task, to see if this reduced their ability to exert control in a later and completely unrelated task. Although the tasks were quite different, Muraven et al. suggest that the first depleted the resource that people use to control their behaviors and feelings, making it difficult to engage in an act of self-control in the second task. This is why Sarah found it hard to avoid eating the potato chips--she had used up a lot of energy when checking her impulse to tell off her ex-boyfriend, making it difficult to exert self-control on something else.

Your task:
Suppose you want to investigate the existence of this phenomenon in other cultures, but you don't want to waste your time replicating other people's research. We would like to know what you would do to research what experimental studies have already been done.

The only source you have available to you for this task is PsycINFO. Please show us how you would use that data base to accomplish this. So, perform a search or searches in PsycINFO and report the following information:

1. The list of search(es) that you tried. See screenshot here for how to find this information:



Export this list as a pdf, create a screenshot of the search page, OR print and save as pdf. Upload this document using the email address: upload.PSY001_.5nlmerm4iz@u.box.com

2. A list of the most relevant articles. Upload this document using the email address: upload.PSY001_.5nlmerm4iz@u.box.com (Check the box next to each article in order to select for emailing and email from within PsycINFO. Enter your full name when prompted). You may send more than one email.

Upload the results of this assignment using this email address:
upload.PSY001_.5nlmerm4iz@u.box.com