To: Learning Outcomes Assessment Committee (LOAC)

From: Jean Kirnan and Madeline Weinland

Date: October 1, 2014

Re: Psychology Department Assessment Report for Fiscal Year 2013-14

Our goal for this year was the investigation of four learning outcomes reflected in two of the required courses of our methodological core (PSY121, Methods and Tools in Psychology; and PSY203 Design and Statistical Analysis). These learning outcomes include **Psychology as a Science, Critical Thinking**, **Ethics** (PSY121 only), and **Information Competence**. See attached definition of learning outcomes for additional detail.

The most recent versions of both instruments can be viewed on Qualtrics:
PSY121, Methods and Tools in Psychology: https://tcnj.qualtrics.com//SE/?SID=SV_2rDwl61clelyqxf
PSY203, Design and Statistical Analysis: https://tcnj.qualtrics.com//SE/?SID=SV_6fGcfogXf4GU7TD

PSY121, Methods and Tools in Psychology

Development of Assessment Tool

We involved faculty and students in various stages of the assessment process. Unlike other researchers noted in the published literature, the development of the instrument was the responsibility of one faculty member who did not teach the course. We felt this was beneficial as the final measurement was more broadly based; not biased toward a specific faculty member's pedagogy. Full time and adjunct faculty were solicited for input on course topics to include in developing the measure, during data collection, and to "close the loop". Student input was obtained from students participating in the TAPLab (Testing and Assessment in Psychology) and those enrolled in PSY386, Psychological Testing, both of whom assisted with administration, item analysis, and survey changes.

Formal course descriptions and syllabi from the previous academic year listed the course learning objectives and were used to derive 55 subtopics relevant for Methods and Tools (i.e., independent vs. dependent variables, sample vs population, research methodologies, internal and external validity, threats to internal validity, reliability and validity, level of measurement, ethics, APA style). To confirm that all the relevant topics had been identified, an on-line survey was developed where the current and recent faculty of Methods and Tools (both full time and adjunct) rated each of these 55 topics as "essential," "optional," or "unessential" for inclusion in the course. Faculty respondents were also invited to submit additional topics that they viewed as essential. Questions were then developed to measure the subtopics rated as "essential" by at least 50% of the respondents thus ensuring content validity. Test items were developed over the summer and converted to an on-line survey format. The instrument was piloted in fall 2010 and revised for use in a full pre/post methodology the following year.

Students registered in all six sections of Methods and Tools in the fall 2011 semester were invited to take the assessment on-line and those who did participate received "participant pool" credit. The instrument was administered as a pre-test on the first day of class and again near the end of the semester as a post-test. Of the 108 participants who took both the pre-test and

post-test, a total of 43% failed a validity check inserted into a series of attitudinal items; thus the analysis was performed on the remaining 67 participants. Subsequent data were collected the following fall, 2012, again using all students in PSY121 (n = 128) and in spring 2014 (n = 57) using students recruited via our human participants pool as part of a related research project. Both subsequent administrations employed a post-test only design

Statistical Findings

Throughout the pilot phase we conducted item analyses focusing on item difficulty, item differentiation, item validity, and item reliability. Individual items were retained, edited or removed based on these findings. Once finalizing the instrument our analyses focused on demonstrating the psychometric value of the new assessment tool through acceptable reliability and validity and then also demonstrating growth over the course of the semester.

We demonstrated the validity of the assessment measure by correlating total post test scores with final course grades. A statistically significant relationship was found between post-test scores and final grades, $r(65) = .522$, $p < .001$; demonstrating criterion validity. Content validity had already been established in the development of the instrument through the use of faculty agreement who are considered to be Subject Matter Experts (SMEs) of the topics included.

Acceptable reliability was demonstrated with a Cronbach's coefficient alpha value of $\alpha = .75$. While on the more moderate side in terms of strength of relationship for a cognitive skills measure, several authors have noted that measures of achievement will have lower reliabilities due to restriction in range (relative to measures of aptitude used in selection or placement).

Finally, "value-added" learning was demonstrated across the semester by examining the difference between pre-test and post-test average scores using a paired samples *t* test to calculate the mean difference across the semester. Pre-test scores ($M = 29.49$, $SD = 5.59$) were significantly lower than post-test scores ($M = 40.58$, $SD = 5.64$), $t(66) = -17.50$, $p < .001$, $d = -1.98$, $CI_{.95} = -12.35, -9.82$ indicating that learning of these topics had occurred.

Validity and reliability of the Methods and Tools instrument were further confirmed the following year, 2012, in a post-test only study with a revised version of the instrument. This analysis revealed consistent reliability with a Cronbach's coefficient alpha of $\alpha = .74$, as well as validity with a significant positive correlation with final course grades, $r(128) = .491$, $p < .001$. Additional data became available as part of a separate research study conducted in spring 2014. This analysis revealed consistent reliability yielding a Cronbach's coefficient alpha of $\alpha = .71$, as well as validity with a significant positive correlation with final course grades, $r(57) = .496$, $p < .001$.

The first two analyses involved the administration of the PSY121 survey at the end of the semester in which the course was taken; however, the final data collection in Spring 2014 was conducted via our human participant pool with the sole requirement of completion of PSY121. Thus, none of the participants were currently enrolled in the class, 30% had completed PSY121 the prior semester, 33% completed two semesters ago, and the remaining 37% completed the course anywhere from 3 to more than 8 semesters prior. Considering this delay in assessment, the consistent reliability and validity obtained in this final analysis lend support to student retention of the course material.

Specific Findings of Learning Outcomes and Action Steps

An analysis of item clusters was conducted to understand which of the specific learning outcomes were being adequately addressed. Most of the topic areas that showed low growth appeared to have been influenced by weak items and would hopefully benefit from item revisions or movement of the assessment to a different part of the curriculum. Several changes were made to address these concerns.

Items designed to measure **Psychology as a Science** and **Critical Thinking** tended to be strong and many were improved through revisions suggested by the item analysis. However, there were a few items that yielded low growth and or low mastery that could not be explained by high pre-test scores or item weaknesses. These items included (with their item number and subtopic): 3 (research question vs. hypothesis), 7 (subject variable), 8 (dependent variable), 19 and 20 (operational definition vs construct), 26 (quota sampling), 43 (improving a research study), and 44 and 45 (level of measurement). Follow-up conversations with faculty who taught Methods and Tools indicated that all these concepts were taught, however some faculty used different terminology than what was used in the assessment tool. Thus, modifications to the items were made to determine if this might account for their weakness.

In larger conversations during a department meeting, several faculty noted that the main points of confusion on these very topics (hypotheses vs research questions, levels of measurement, independent vs. dependent variables) persisted in higher level courses of the curriculum such as Research Seminar, 300-level topics seminar courses, and faculty led research labs. Upon further discussion in the department there was agreement that these topics, as well as concepts related to experimental methodology are often "learned" for a particular experiment or article reading, but students have difficulty applying the concepts to novel situations. We re-affirmed the need for repetition of these topics across the three methodological courses Methods and Tools, Design and Statistical Analysis, and Research Seminar.

Both items designed to measure knowledge of **Ethical** guidelines were weak and we discussed in detail in the department, revisiting the original learning objective which specifically states students will both demonstrate and understand ethical standards; however, knowledge and behavior will likely need to be measured in different venues. To ensure knowledge and understanding of ethics, we revised the items and piloted at a later date and added relevant ethical items to a measure of learning in our Design and Statistical Analysis course. These changes reveal good understanding of informed consent and participant rights, but only a moderate level of understanding the differences between confidentiality and anonymity in data collection. We also mandated an on-line human research participant certification for all majors as part of the third course in the methodological core, Research Seminar. Between the new items and the certificate, we hope to be able to demonstrate ethical knowledge. We will attempt to measure ethical practice through an evaluation of feedback surveys completed by research participants (three items in an existing participant survey will be used as they ask about procedures, researcher professionalism, and participant rights). Additionally, student research proposals will be monitored for their adherence to ethical standards. See Psychology Annual Assessment Report for 2012-13 for details on these initiatives.

Information Competence was largely measured through the use and understanding of American Psychological Association (APA) research, writing, and journal formatting guidelines. Items on database searching and journal formatting showed strong growth. However, the four items measuring APA mechanics were mixed; while two items appeared acceptable, the other

two either showed a loss in knowledge or a low percent correct in the post-test. These questions used a "yes/no" response format and thus had a high probability of accuracy through guessing. Discussions with senior level students revealed that APA writing style is not memorized; rather students access the APA manual and on-line resources when writing a paper. Thus, the decision was made to remove these from the Method and Tools survey and measure this objective via a rubric being designed for the end of semester, APA style paper required in Research Methods and Senior Capstone courses.

In our spring department retreat of 2014 our Chair asked that we revisit our methodological core courses assigning this task to an ad hoc committee. The instrument designed here will add to those conversations; similarly, those conversation and any changes to PSY121 will need to be reflected in this instrument going forward.

PSY203, Design and Statistical Analysis

Development of Assessment Tool

Following the exact same process for the PSY121 measure we reviewed formal course descriptions, recent syllabi, and consulted with faculty teaching the course to develop a blueprint for test development. In the fall 2012 semester, an online survey was administered to all students enrolled in PSY 203 utilizing a pre-test/post-test format. Students completed the survey in the beginning of the semester and at the end of the semester. Concerned with the loss of sample size in the PSY121 study due to failed validity checks, we incorporated a different strategy for the PSY203 measure. In addition to serving as "detection" of valid responding, the validity checks also served the purpose of "prevention". Thus, four validity checks were embedded within the survey to ensure that the participants carefully read and responded to each question. These checks appear early as a "warning" to respondents, hopefully preventing inaccurate response sets, and later as a detection of inaccurate responding. Only the data for the students who passed all the validity checks were used in the analysis. Of the 74 students, 62 (84%) passed all the validity checks.

Statistical Findings

Throughout the pilot phase we conducted item analyses focusing on item difficulty, item differentiation, item validity, and item reliability. Individual items were retained, edited or removed based on these findings. Once finalizing the instrument our analyses focused on demonstrating the psychometric value of the new assessment tool through acceptable reliability and validity and then also demonstrating growth over the course of the semester.

We demonstrated the validity of the assessment measure by correlating total post test scores with final course grades. A statistically significant relationship was found between post-test scores and final grades, $r(62) = .51$, $p < .001$; demonstrating criterion validity. Content validity had already been established in the development of the instrument through the use of faculty who are considered to be Subject Matter Experts, SMEs, of the topics included.

Reliability was demonstrated with Cronbach's coefficient alpha yielding a value of $\alpha = .60$ which is weak for a cognitive measure. We re-ran this statistic eliminating the five **Ethics** questions at the end which are pedagogically distinct and the reliability improved to $\alpha = .64$. We hypothesize that the lower reliability reflects the distinct nature of the various topics

in PSY203 and would be better understood following additional statistical analysis such as a factor analysis.

Finally, "value-added" learning was demonstrated across the semester by examining the difference between pre-test and post-test average scores using a paired samples *t* test to calculate the mean difference across the semester. Pre-test scores (M = 18.02, SD = 3.17) were significantly lower than post-test scores (M = 26.61, SD = 4.17), $t(61) = -13.00$, $p < .001$, $CI_{.95}$ = -9.92, -7.27 indicating that learning of these topics had occurred.

Specific Findings of Learning Outcomes and Action Steps

Using a pre/post methodology we determine that the curricular goals of PSY203 are largely being met. We compared pre and post course performance on a direct knowledge test and showed improvement in most areas. Some items have been re-written to improve wording clarity but overall the measure is working well and students show learning. A few topics show weaknesses that cannot be explained by erroneous items and need to be reviewed with faculty teaching PSY203:

Psychology as a Science and **Critical Thinking** were reflected in many items where students had to apply specific statistical knowledge or technique to a problem. Students were particularly adept at the difficult task of deciding what statistical test to use when given a series of short research scenarios. There was a minor difficulty with one item that was more conceptual than applied. Students showed growth but not mastery of a conceptual topic on the differences between parametric and non-parametric tests.

Several items involved reading data from SPSS output, graphs, or tables. Overall students performed very well on these items demonstrating **Information Competence.** Only one item presented difficulty where students had to interpret main effects and interaction in a simple 2 x 2 table of means. Items referencing this table showed little growth from pre to post but did strongly discriminant between high and low scorers on the test. Upon review, it would appear that this skill in table reading is more appropriate for PSY121 curriculum which supports the lack of growth but higher level of this skill in high performing students. A different table showing SPSS output for cross-tabulation showed strong growth for three items referencing the table, but students had difficulty choosing the graph format that would best represent the data. As noted earlier, our department has begun a discussion of the course topics in our methodological core and graphing will be part of that conversation.

While the **Ethics** learning objective is not formally a part of PSY203 we had added a few items to this survey given problems with measurement in the PSY121 assessment. The five items presenting ethical dilemmas in research showed mixed learning. Three of the items showed high pre-test scores suggesting the topics may have been covered in PSY121 (a prerequisite for PSY203) or other coursework. The other two items showed negative growth. One dealt with parental consent and had some ambiguity in wording and has been corrected. The other item was concerned with a researcher changing hypotheses to fit the data. This will be shared with faculty teaching the course.

Additional Analyses

In academic year 2013-14 we correlated scores on both measures with grades in subsequent courses within the methodological core. Table 1 presents those results which

additionally demonstrate that both measures are able to predict success in the next course in the sequence, but not necessarily later courses.

The sample size shrinks over time as these subsequent courses are only open to majors and even among majors, students do not always complete the courses in sequence. The score on the PSY121 measure of knowledge correlates r (41) = .490, p <.001 with grade for PSY203, Design & Statistical Analysis; and r (38) =.165, p=.322 for grade achieved in PSY299, Research Seminar. From a learning outcomes perspective it is disappointing to show no relationship with Research Seminar as both Methods & Tools and Research Seminar have similar learning objectives. However, the type of student work that constitutes the final grades in these courses is vastly different with Methods & Tools traditionally consisting of objective tests, homework, and a final written project; while the Research Seminar grade is largely composed of a student project that incorporates writing, research design, data collection, data analysis. The final product in the course is a lengthy APA style paper submitted in stages and undergoing many revisions following faculty feedback.

Interestingly, although constituting different course topics, the assessment for PSY203, Design and Statistical Analysis, correlates strongly with grades for the subsequent methodological course, PSY299, Research Seminar, with r (63) = .353, p <.001. This relationship likely reflects the skills and knowledge learned in PSY203 that are selectively applied in the PSY299 course.

With the collection of data in Spring 2014, we had the opportunity to review changes made to the PSY121 **Ethics.** The original items were true/false in nature and were modified to a multiple choice format. New items in the most recent version of PSY121 instrument show student strength in the areas of informed consent and participant rights. There is still some confusion on the difference between confidentiality and anonymity; a finding that will be shared with faculty.

Conclusion

Through a repetitive process that included faculty and students we developed reliable and valid measures of PSY121, Methods and Tools in Psychology and PSY203, Design and Statistical Analysis. Both measures demonstrated that overall the learning outcomes are being met. As part of this process we identified more appropriate places in the curriculum to measure some aspects of the LOs and recommitted ourselves to the repetition of certain core concepts throughout our course sequence. Additionally, Psychology has constituted an ad hoc committee to re-evaluate PSY121 as a course but also more broadly as part of the required Methodological Core for the major. The information gained here will be quite helpful in that conversation.

Psychology

Learning Objectives

1. **Psychology as a science**. Students will develop skills to participate in the creation of and the constructive critique of psychology as a science.
   Students will be able to apply the scientific method to test research hypotheses and answer questions in psychology.

2. **Psychological knowledge**. Students will become familiar with major concepts and empirical findings in various areas of the discipline, will appreciate the variety of perspectives in psychology, and will learn to explore the discipline independently.
   Students will be able to identify and differentiate the most important theories and empirical findings in the major areas of psychology

3. **Critical thinking in psychology.** Students will respect and use critical and creative thinking, skeptical and open-minded inquiry, and the scientific approach to analyze issues related to behavioral and mental processes.
   Students will be able to evaluate the validity of information and findings of others as well as develop creative, yet valid modes of inquiry to address issues related to behavioral and mental processes.

4. **Diversity.** Students will recognize, understand, and respect the complexity and impact of diverse sociocultural, historical, and international systems on psychological phenomena.
   Students will be able to recognize and understand the complexity of sociocultural and historical diversity and develop a respect for those different than themselves.

5. **Ethical standards and conduct in psychology.** Students will demonstrate an understanding of and a commitment to the ethical standards of psychology when engaged in such activities in the discipline as describing, explaining, interpreting, evaluating, applying, and/or practicing.
   Students will be able to apply knowledge of ethical standards of psychology to distinct activities.

6. **Applying psychology.** Students will apply psychological principles to personal (self and other), interpersonal, organizational, community, and cultural issues.
   Students will be able to use psychological principles in multiple contexts (i.e., personal, interpersonal, organizational, community) in an effort to inform, intervene or change

7. **Career preparation.** Students will pursue realistic ideas about how to implement psychological knowledge, skills, interests, and values in occupational pursuits in a variety of settings and in socially responsible ways.
   Students will be able to articulate a variety of career options for the psychology major and demonstrate knowledge of the steps needed to progress towards their own career goals.

8. **Effective communication.** Students will communicate effectively in a variety of formats (i.e., writing, oral, aural, quantitative, interpersonal, and collaborative).
   Students will be able to communicate effectively in a variety of mediums (i.e., oral communication and professional writing).

9. **Information competence.** Students will demonstrate information competence and the ability to use computers and other technology for many purposes, including scientific inquiry, application, and communication.
   Students will be able to use computers and other technology to identify, access, and provide meaning to information

Table 1

Correlation of PSY121 and PSY203 Assessments with Measures of Academic Success

	Course Grades		
	PSY121	PSY203	PSY299
Course Assessment			
PSY121	.526	.490	.165
	($p < .001$)	($p = .001$)	($p = .322$)
	$N = 67$	$N = 41$	$N = 38$
PSY203	----	.444	.353
		($p < .001$)	($p < .001$)
		$N = 68$	$N = 64$