



Using Empirical Article Analyses to Assess Students Learning of Psychology Research Methods

Sarah Richardson, Michael Schiel, Kaetlyn Graham, & Allen Keniston
Department of Psychology, University of Wisconsin – Eau Claire



Introduction

Bachiochi et al (2011) developed an article analysis activity to assess psychology students' mastery of American Psychological Association student learning outcomes (SLOs) for research methods (Goal 2; APA 2007). Students read a research report and answer factual and analysis questions about it. Bachiochi et al used the activity to assess their department's success at teaching research methods over a two course sequence, finding that it both documented students' learning and identified ways in which they could improve their program.

We applied Bachiochi et al's approach to assess student attainment of the UWEC psychology department's SLOs for research methods. Using a cross-sectional design, we expected that:

1. students advanced in our curriculum would score higher on questions about a research report .
2. Improvements would be greater for analysis versus fact questions.
3. Degree of improvement would vary depending on the SLO each question addressed.
4. The variation in improvement per question would suggest possible improvements to our curriculum
5. Undergraduate researchers would be effective participants in the assessment process.

Method

Participants - 109 Male and 247 female undergraduates (n=356) in first through fourth year psychology courses. Professors volunteered students in their classes. Except introductory psychology students, all students were majors or minors. We identified five participant levels, students who had 0 previous psychology courses, (138); 1-2 courses, (42); 2-4 courses, (41); 5-7 courses, (59); 8 or more courses: (75).

Materials – Students analyzed an abbreviated report by Powell and Drucker (1997), presented in Lomond (2002), and used by Bachiochi et al. They answered questions shown in Table 1. Five questions addressed factual information in the article and five addressed report analysis. Questions at the end of the worksheet collected demographic information.

Procedure – In the first four weeks of the semester, participating psychology instructors volunteered 30 minutes of a regular class period for data collection. Instructors introduced two members of the research team who explained the study, took questions, and gathered informed consent forms before distributing the research report and the questions. Participants took as much time as they needed, but all completed the worksheets within 30 minutes. They could refer to and annotate the research report both as they read it and as they answered questions.

Prior to data collection, three undergraduates and one professor trained on and practiced question scoring using pilot data collected in the fall 2011 semester. Difficulties with applying Bachiochi et al's scoring key were addressed and led to an improved key. To estimate scoring reliability we obtained intraclass correlation coefficients for four pairs of raters' individual item scores for 138 participants. Intraclass correlations ranged from .78 to .83. Scores for analyses reported here are percentages of total points for all questions, percentages of "fact" scores and "analysis" scores, and percentages for individual item scores.

Results

A mixed 2 (within, type of test) by 5 (between, level) ANOVA compared mean percentages scores . There were main effects for type of test ($F(1, 350) = 512.83, p < .001$) and course completion level ($F(4, 350) = 37.10, p < .001$). Mean percentage total scores on "fact" items was .73 (SE = .01) and .45 (SE = .01). Figure 1 shows mean percentage total scores across levels; the significant contrast occurred between Level 1 and Level 2 (means are .51, SE = .02 vs. .62, SE = .02). Type of test interacted with level, ($F(4, 350) = 3.64, p = .006$) Overall, fact scores increased more over levels than did analysis scores.

A mixed 10 (within, item) by 5 (between, level) ANOVA assessed individual item effects versus level. There were main effects for item ($F(9, 3150) = 155.71, p < .001$) and level ($F(4, 350) = 42.64, p < .001$). The interaction term was also significant ($F(36, 3150) = 5.74, p < .001$). The levels effect is the same result reported above and illustrated in Figure 1. The item effect reflects diverse degrees of difficulty among items; The interaction effect results from varying degrees of improvement in item scores across level. Figure 2 illustrates this variability. For example, there was a large increase in scores for Item 4 between level 0 and level 4, but much smaller increases for items 9 and 10. The overall pattern of these differences reflects the interaction between type of test and level.

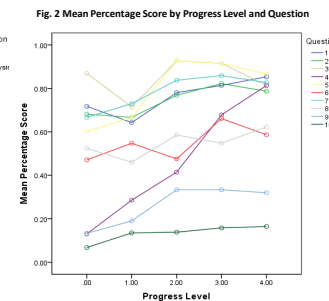
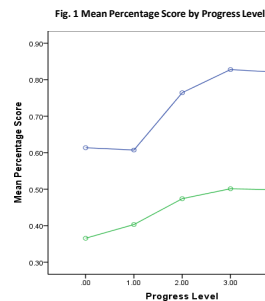


Table 1.
Questions, Related APA Guideline Learning Outcome(s), and Scoring

Question	APA Guideline	Points
Did the driver have a fake beer in his hand in all four conditions?	NA	1
Identify the independent variables in this experiment.	2.1	2
How many of the 10 participants refused to enter the car in the condition where the confederate refused to enter the car?	2.3	1
At what probability level was the relationship between the IV and the DV significant?	2.3	1
According to the researchers, what would be the problem with making the driver's degree of intoxication more extreme in future studies?	2.4	1
The researchers used undergraduates as participants in this study. Does this limit the generalizability of the results? Explain.	2.6	2
Was deception used in this study? Explain.	2.5	3
The researchers recommend gender as an interesting variable for future study. How might gender influence the results?	2.4, 2.6	3
What was wrong with how concern was measured in this study?	2.4	3
What type of research design was used in this study? Support your answer.	2.2	3

APA SLO	Description
2.1	Describe the basic characteristics of the science of psychology
2.2	Explain different research methods used by psychologists
2.3	Evaluate conclusions derived from research
2.4	Design and conduct basic studies using appropriate research methods
2.5	Follow the APA ethics code pertaining to research and participant treatment ethics
2.6	Generalize research conclusions appropriately based on the parameters of particular research methods

Discussion and Conclusions

We believe that Bachiochi et al's article analysis approach effectively assesses our students' learning of research methods in psychology. Students' answers to fact and analysis questions about an article they have just read improved across levels defined by courses completed. Importantly, the method sketched details that provide implications for program improvement.

First, overall, significant improvement occurred from the second to the third level of course completion, corresponding to completion of statistics and methodology courses. No change occurred later. Second, unexpectedly fact scores increased faster than analysis scores. We thought training in research methods would differentially enhance student's analytic ability because that ability would be relatively lower at the outset but would be the target of training.

We also identified variable improvement with respect to specific APA Goal 2 objectives. For example, dramatic improvement occurred in students' identification of a finding's statistical significance (a fact). But ability to identify a study's research method and to critique a measure (analysis) improved less. Our results suggest that we need to find better ways to extend students' understanding research types and their ability to evaluate aspects of method (Goals 2.2, 2.3, and 2.4).

Concerns about some of the article analysis questions qualify these findings. For example Question 9 asked, "What was wrong with how concern was measured in this study?" The key for this question stipulates three points to address. Our students never provided more than one, which seems reasonable given how the question was asked. Question 10 asked, "What type research design was used in this study?" Advanced students typically gave a structural answer (e.g., a "2 x 3 design"). While incorrect, such answers reflect greater sophistication in understanding research reports and derive from what students learn the question is asking.

We have concerns about the validity of our findings, but not about the quality of undergraduates' participation in obtaining them. Intraclass correlations of assistants' scoring of participants' responses were very strong, comparable to those of Bachiochi et al.

While our work is certainly preliminary, and limited (uneven participant numbers across levels, participants not being representative of all psychology majors; measurement problems just noted), our results are interesting and compelling. They support further testing of an article analysis method to assess our department's teaching of research methods and possibly other aspects of the psychology curriculum.

References and Thanks

American Psychological Association. (2007). *APA guidelines for the undergraduate psychology major*. Washington, DC: Author. Retrieved from www.apa.org/ed/resources.html

Bachiochi, P., Everton, W., Evans, M., Fugere, M., Escoto, C., Letterman, M., & Leszczynski, J. (2011). Using empirical article analysis to assess research methods courses. *Teaching of Psychology, 33*, 5-9.

Lomond, T. C. (2002). Cross section of journal articles for discussion and evaluation: Social science research (3rd ed.). Los Angeles: Pyrczak Publishing

Powell, J. L., Drucker, A. D. (1997) The role of peer conformity in the decision to ride with an intoxicated driver. *Journal of Alcohol and Drug Education, 43*, 1-7.

Special thanks to ORSP for the funding of this poster