

PROMOTING CRITICAL THINKING ABOUT SCIENTIFIC RESEARCH

Trent W. Maurer and Marlo Rouse-Arnett
Georgia Southern University

ABSTRACT. This paper explores the efficacy of 2 methods of teaching critical thinking about scientific research: the traditional research paper vs. critiquing published research. Thirty-two students enrolled in either of 2 sections of a research methods course completed a pretest/posttest truncated version of Lawson's Psychological Critical Thinking Exam. One section of the course required students to write a research paper and a basic critique of a published study; the other required students to summarize 3 studies and critique in detail 3 additional studies. Results indicated that students in neither section significantly improved in their critical thinking skills from pretest to posttest, though there was substantial variability within and between sections.

Higher-order critical thinking skills are essential to success in most careers in the social sciences, and arguably the most important of these skills is the ability to critically examine and understand published research in one's field (Zablotsky, 2001). However, most college students lack these skills (de Sanchez, 1995; Lawson, 1999) and need to be taught how to learn and apply them. Ironically, this deficiency in critical thinking skills among college students may actually be a product of the educational system itself, which stresses "memorization of isolated knowledge components, which are devoid of meaning, lack transferability, and are easily forgotten" (de Sanchez, 1995, p. 73). Further, although this focus on learning without critical thought is potentially cognitively easier for students, it is not as meaningful, and many students in the social sciences lament their lack of formal training and experience specifically with regard to critically reviewing and evaluating research (Dingfelder, 2005).

The awareness of this problem, and the dedication to addressing it, is pronounced in the social sciences, particularly in the field of psychology: (a) there are myriad books dedicated to the topic of critical thinking (e.g., Gilovich, 1991; Stanovich, 1998); (b) the journal *Teaching of*

KEY WORDS. Critical thinking, Blooms' taxonomy, Research paper, Critique

Correspondence concerning this article should be addressed to the first author: Trent W. Maurer, Department of Hospitality, Tourism, and Family and Consumer Sciences, Georgia Southern University, P.O. Box 8021, Statesboro, Georgia 30460 (E-mail: tmaurer@GeorgiaSouthern.edu).

Psychology has published an entire issue on the topic (Halpern & Nummendam, 1995) and continues to publish articles on it; (c) most recently, the 2005 Best Practices in Teaching Psychology conference was devoted entirely to methods and techniques for promoting critical thinking about psychology. The most common method for teaching critical thinking skills in the social sciences is the assignment of a research paper, though other alternatives have been proposed (Oliver, 2005). Below, we outline one theoretical framework from which to view the efficacy of both the research paper and an alternative, the research article critique, in teaching critical thinking skills.

Theoretical Framework

Bloom's (1956) Taxonomy provides a useful framework for examining and differentiating between different levels of learning and thought. In this taxonomy, there are six levels of reasoning. Each level is progressively more abstract, complex, and effortful, builds upon the prior levels in a hierarchical fashion, and requires and promotes more critical thought than the levels below it. Bloom classified the first three levels, knowledge, comprehension, and application, as "lower-order" thinking skills, and the second three levels, analysis, synthesis, and evaluation, as "higher-order" thinking skills.

The creation of a research paper requires thought at the synthesis level of the taxonomy. Synthesis-level thinking is a higher-order thinking skill and involves such abstraction and complexity as to create something "new" (i.e., a research paper). This skill requires mastery of the prior four levels of the taxonomy. In order to create and appropriately structure a research paper, an individual must first have: a basic knowledge of the topic and the disciplinary methods and formats for communicating about a topic, an understanding of the meaning and implications of prior research on that topic, the ability to apply basic research methodology and the findings

of prior research to new situations, and the ability to examine the relative importance of different components of prior research to the current understanding of the topic. Only after mastery of the prior four levels of the taxonomy will the individual be able to attempt synthesis, which requires them to appropriately and critically marshal evidence and structure and present it in the most convincing way. Synthesis assignments like research papers should therefore promote and require substantial critical thought.

However, evaluation skills are also a higher-order skill in Bloom's (1956) Taxonomy, and Bloom placed them above synthesis skills. According to Bloom, evaluation is even more abstract, complex, and effortful than synthesis and requires the highest level of critical thought, since evaluation occupies the highest level of the taxonomy. In this framework, for an individual to complete an evaluation-level assignment, such as critiquing published research, they would need to master the same skills required for synthesis, plus synthesis itself, before they could attempt such an assignment. Thus, Bloom's Taxonomy suggests that although the creation of an original research paper requires students to use many skills, it does not promote or require the most abstract and complex reasoning and critical thinking skills.

In contrast, Anderson and Krathwohl (2001) have argued for a revision of the taxonomy that would invert the placement of synthesis and evaluation, placing synthesis at the highest level (which suggests that a research paper would require the highest level of skills). Huitt (2005) has even argued that synthesis and evaluation actually occupy the same level, but that synthesis can more appropriately be viewed as "creative thinking" whereas evaluation can be viewed as "critical thinking," with both being essential to the problem-solving process. Given the disagreement about which type of thinking promotes and requires the highest level of critical thinking skills, it seems prudent to investigate whether synthesis-based assignments and

evaluation-based assignments are comparable in the extent to which they promote critical thinking, or if they even promote critical thinking at all. Although the research paper is a clear example of a synthesis assignment, examples of an evaluation assignment require further elaboration.

One possible evaluation assignment would require students to critique published research. Not only would doing so address the need of students to learn how to evaluate research and gain experience doing it (Dingfelder, 2005), but it would teach students to constantly question the information they consume. Researchers have argued that critical thinking includes, but is not limited to: (a) questioning information (King, 1995), (b) evaluating the credibility of sources (Carlson, 1995), and (c) examining assumptions (Yanchar & Slife, 2004). All three of these elements are essential to critical thinking, especially about research in the social sciences, and they are far more likely to be learned and practiced if one is being taught how to use them to critique the published work of others.

Teaching students how to critique published research is not a new idea (Henderson, 1995; Kipp & Lamson, 2005; Oliver, 2005; Zablotsky, 2001), but like the research paper, its efficacy in teaching critical thinking skills has not been tested. This is particularly of interest given the debate about whether synthesis or evaluation assignments require higher levels of critical thinking skills. What makes these two methods ideal for comparison (aside from the fact they both occupy high levels in Bloom's Taxonomy) is the fact that they both require substantial written elements. Written assessments provide significant advantages in promoting critical thinking over nonwritten methods, such as oral expression (e.g., presentations, discussions, etc.), including: (a) promoting greater self-reflection, (b) promoting greater perspective taking, (c) requiring equal participation from all students, thus promoting active learning, (d) promoting

dialectic reasoning by requiring students to articulate and evaluate multiple sides of an issue, and (e) promoting careful revision of ideas, positions, and articulations (see Wade, 1995, for a discussion).

Additionally, as Tomcho et al. (2005) note, there is a substantial gap in the literature with respect to empirically assessing students' abilities to weigh support for conclusions and determine how well reasons support conclusions. This investigation will address this gap and investigate the efficacy of using each of the two methods, a traditional research paper and critiques of published research, to teach students to think critically about scientific research.

Because both the research paper and the critique focus on the highest levels of thinking in Bloom's (1956) taxonomy, there is no a priori reason to expect that students who are taught one technique would outperform students who are taught the other on a measure of critical thinking skills. Indeed, the major goal of this investigation is to document whether either method does indeed promote gains in critical thinking skills. Thus, we examined how each section performed on a pretest and posttest measure of critical thinking skills, improvements in critical thinking skills from pretest to posttest, and variability within and between the groups.

Method

Participants

Participants were 32 students enrolled in either of two sections (Research Paper or Critique) of a research methods course required for all Child and Family Development majors. The two authors each taught one section. With the exception of one student in the Research Paper section, all students were female. We did not collect information on age, ethnicity, year in school, or prior or concurrent courses that may have stressed critical thinking.

Instruments

We used a truncated version of Lawson's (1999) Psychological Critical Thinking Exam (PCTE) to assess students' critical thinking. The PCTE is a 14-item measure that presents information about different scenarios along with an interpretation of the meaning of those scenarios with respect to a relationship between variables. For example, one item describes how a florist played a subliminal tape for her employees and after two days, her employees started making creative floral arrangements out of materials in a scrap box. The presented interpretation is that the music caused an increase in creativity. Participants are then asked to explain if there are any problems with that interpretation, and if so, to explain the problem (i.e., they are asked to methodologically critique the "experiment" in much the same way as a journal reviewer might).

The PCTE taps nine different components of psychological critical thinking across the 14 items, with many items assessing more than one component. Lawson's (1999) results suggest adequate validity of the measure: senior psychology students outperformed senior natural science students, and both outperformed introductory psychology students, as would be expected if the skills tested were taught and practiced progressively throughout the curriculum.

Unfortunately, the full PCTE can be a time consuming assessment (e.g., allowing an average of just 4 min per item to complete the exam would require more than one full 50 min class period). Because items not attempted (due to insufficient time) could be misclassified as critical thinking errors, it is essential that participants have sufficient time to complete the measure. For this investigation, we set an a priori time limit of 15 min, which reflected the maximum amount of class time we could spare for the assessment on each of two separate class periods. We thus truncated the exam to six items. Six items were the minimum number required to assess all nine components of the PCTE and the maximum number we thought participants could be reasonably expected to complete within 15 min. However, to minimize practice effects

and incorporate the largest possible number of items, only two of the six items appeared on both the pretest and posttest. The remaining four items differed from pretest to posttest, although they assessed the same components.

Procedure

Instructors in the two sections of the course changed lecture material in each section to teach the appropriate skills for the different assigned tasks, but in all other respects, the curriculum and assignments were similar. The instructor in the Research Paper section used a coursepacket and the instructor in the Critique section used the Brown, Cozby, Kee, and Worden (1999) textbook and a coursepacket. Neither set of reading materials focused explicitly on critical thinking skills. The instructors neither advertised nor denied the difference between the sections during the period when students registered for the course, and anecdotal evidence suggests that students were aware of the difference between the sections at the time of registration.

The Research Paper section required students to write a 2-3 page critique of a published article (approximately 3% of the final grade) in addition to writing a traditional research paper of 12-18 pages that they had to present orally to their peers in the class via a 10-15 minute presentation (33% of the final grade). The critique assignment required students to evaluate the organization of the article, the abstract, the introduction and literature review, the methods, and the conclusions, using criteria similar to those found in the APA Publication Manual (2001). Students evaluated the organization of the article in terms of the appropriateness of the title and the use of headings, subheadings, and section headings, as well as general flow. Students evaluated the abstract based on its brevity, clarity, and the inclusion of sufficient information on the study's purpose, subjects, procedures, and results. For the introduction and literature review,

students evaluated the presentation of the research problem, the prior research presented related to that problem, and the justification for further research on that problem. Additionally, students evaluated the ease with which the research variables could be identified and how they were operationally defined. For the methods section, students evaluated the clarity of presentation of the sections, the appropriateness of the instruments and procedures, the inclusion of information on reliability and validity, and any potential confounding variables or influences. Finally, students evaluated the practical significance and applications of the research, the extent to which conclusions were logically related to the hypotheses, and the suggestions for future research provided by the authors. Students were allowed to select their own article to critique, but it had to be a research article from a scholarly journal and the topic had to address an issue in the field of Child and Family Development.

The research paper required the completion of an original research project utilizing at least two Likert-scale questionnaires of at least four items each, in addition to demographic questions. This paper required a title page, abstract, 1-2 page introduction, 6-7 page literature review with at least six peer-reviewed journal articles cited, 1-2 page methods section, 1-2 page results section, 2 page discussion, 1-2 pages of references, and an appendix with the questionnaires. This format is consistent with the format for research manuscripts suggested by the APA Publication Manual (2001).

The Critique section required students to summarize three articles in one 3-6 page paper (9% of the final grade) and critique another three articles in three separate 3-5 page papers (21% of the final grade). The summary assignment in this section was similar to the literature review section of a research paper, and required students to identify the purpose of the study described in each of the three articles, the sample size and sample characteristics for each article, the

methods utilized in each article, and the main findings and conclusions of each article. The three articles students read and summarized were Meston and Frohlich (2003), Jackson and Tlauka (2004), and Seal (1997).

In the critique assignments, we required students to evaluate all five major sections of a research article: abstract, introduction/literature review, methods, results, and discussion. These evaluations used preset criteria to focus students' attention on specific issues, similar to those recommended by the APA Publication Manual (2001). For the abstract, students evaluated the degree to which it included sufficient information on the purpose, sample, and results of the study. Students evaluated the literature review in terms of identifying key prior studies and their results, questions raised or left unanswered by prior studies, how the current study proposed to address those questions, and the specific hypotheses proposed (e.g., are the hypotheses falsifiable?). Students evaluated the methods section in terms of sample size, sample characteristics, recruitment procedures, measures used, and appropriateness of the procedure and any confounding variables or influences. Students evaluated the results in terms of the statistical and analytical procedures used, the statistical significance of the results, and comparisons between any tables or figures and the text. For the discussion section, students evaluated how the authors called attention to the most significant results (or any results they neglected to discuss), how they related their findings back to their hypotheses, how well their findings address the problem the authors identified in the literature review, and what they proposed were the implications of their findings for service and future research. The three articles students read and critiqued were Lawson, Schwiers, Doellman, Grady, and Kelnhofer (2003), Teachman (2003), and Tucker, McHale, and Crouter (2003).

Students completed the truncated PCTE early in the semester and again at the end of the semester. Students in both sections had 15 min to complete the tests. Due to absences, not all students completed both tests. Of the 16 students in the Research Paper section, all completed at least one test, and 9 completed both the pretest and the posttest. Of the 16 students in the Critique section, 15 completed at least one test, and 13 completed both. The final sample of 22 students represented a response rate of 69%.

The first author blindly scored the tests. Students who correctly identified an error in reasoning in a test item received one point. Students who articulated part of the problem, but were unable to fully explain it, received one half point. Students who did not identify a problem or incorrectly identified the problem received zero points. Thus the score range was between zero and six points on each test. The second author rescored a randomly selected subset of 10% of the items to check reliability. The interrater reliability coefficient was $\alpha = .94$.

Results

Results for the two sections were computed by total test score as well as separately for the two shared items in the pretest and posttest and the four non-shared items in the pretest and posttest (see Table 1). Analysis of students' performance on the critical thinking test at both total pretest and total posttest revealed substantial variability both within sections and between sections. At pretest, the average total score for both sections was approximately two of six items answered correctly (33%) with a standard deviation of almost one full item (17%). At posttest, the average total score for the Critique section was still approximately two of six items correct, with a standard deviation of almost one item. However, the Research Paper section had an average total score of approximately one item correct, with a standard deviation of almost one item.

Pretest-posttest difference scores also revealed substantial variability both within and between sections. The Critique section averaged a difference score of approximately zero, but the range was between a decrease of over 40% and an increase of 33%. Of the 13 students in the Critique section, 1 scored the same on pretest and posttest, 6 scored lower on posttest than pretest, and 6 scored higher on posttest than pretest. The Research Paper section averaged a difference score of nearly -1.00, but the range was between a decrease of over 40% and an increase of nearly 10%. Of the 9 students in the Critique section, 2 scored the same on pretest and posttest, 6 scored lower on posttest than pretest, and 1 scored higher on posttest than pretest.

Discussion

The results of this investigation revealed substantial variability in students' critical thinking skills, both from pretest to posttest and within and between groups. On average, students in the Critique section demonstrated no change in critical thinking skills from pretest to posttest, but the average masks the bimodal distribution: approximately half of the students demonstrated an increase in their critical thinking skills and half demonstrated a decrease. In contrast, students in the Research Paper section demonstrated an average decline in critical thinking skills of nearly 15%, with only one student demonstrating any increase at posttest. However, in both sections at both times, nearly all the students incorrectly answered the majority of the questions. We believe there are at least two possible explanations for the observed differences between sections.

First, it is possible that these results are the product of changing four of the test items from pretest to posttest. Recall that to minimize practice effects and maximize the number of PCTE items included, we retained only two items and replaced the other four items with conceptually identical items presented in slightly different scenarios. Neither section differed

significantly from pretest to posttest on their scores on the two items that were on both tests. It is possible that students in the Research Paper section simply had more difficulty identifying the same problems in different scenarios, and that had the presentation of the questions been reversed, they would have showed significant improvement in critical thinking. It is important to note that we are not arguing that critical thinking skills themselves are highly context dependent, but only that assessing them using this instrument may require sensitivity to context because of unique elements of critical thinking tapped by each item.

Second, the findings may be attributable to structural differences between the sections. In the Research Paper section, student presentations of their projects occupied the last two weeks of class. The instructor had already completed delivery of all lecture material and students had already taken their last exam by the time the posttest was given. Although the instructor required students to question their peers during their presentations, student motivation for retaining or applying what they had learned may have been low. In the Critique section, the instructor administered the posttest at the same time, but in this section students had just completed their final critique and were still preparing for the last exam. Students were still using their critical thinking skills in this section at the time of the posttest. Further, other extraneous variables which may not have been evenly distributed between the sections may have contributed to the difference (e.g., students in the Critique section may have had more prior classes which required them to use critical thinking skills). Given the small sample size and the fact that assignment to the sections was self-selected, this is a distinct possibility. However, although the small sample size may be problematic, it is reflective of the smaller class size common for research methods courses and thus preserves mundane realism.

Limitations and Future Directions

It is important to note several limitations to this project. First, we did not collect demographic data on the participants, which makes it impossible to assess any potential demographic differences. In this case, the point was moot as we did not have sufficient power with such a small sample to use inferential statistical tests with multiple independent variables without greatly inflating the risk of Type I error (Cohen, 1988). Although it is difficult to make a case that certain demographic variables would influence critical thinking (e.g., ethnicity), others certainly could (e.g., age, particularly in reference to cognitive development and postformal thought). Future research with larger samples could address this issue, particularly in programs that offer multiple course sections each semester so that larger samples could be obtained without larger classes.

Second, our sample size, while accurately representing a common class size for the material, was very small and provided insufficient power for additional statistical analyses. Future research with larger samples, particularly through the use of multiple small sections, could better investigate these issues. Additionally, future research could consider differences between universities which may covary with class size (e.g., research-oriented vs. teaching oriented, public vs. private, use of graduate student teaching assistants, etc.).

Third, we did not require students to complete the test. Students who were absent for the administration of one of the exams were not required to complete the exam on their own time (or under alternative conditions). Ideally, all students in both sections should have been required to complete the full exam. This was especially a concern for us because missing data was not evenly distributed between the two sections. Perhaps future investigations could incorporate the full PCTE into existing testing or evaluation procedures to ensure the full participation of all students.

Fourth, we did not query students about prior courses they may have taken that promoted critical thinking. Although anecdotal evidence suggested that the students had not taken such courses (e.g., comments in class, poor performance on the pretest, etc.), we cannot conclusively rule out the possibility that differences in prior classwork may explain the observed variation in performance on the critical thinking measure.

Although additional research is clearly needed before a definitive conclusion about the pedagogical value of either method for teaching critical thinking can be reached, we believe that future research investigating both methods is warranted, and we encourage our colleagues to consider exploring this valuable line of research. We suggest two specific avenues of future research.

First, researchers could use the full PCTE, rather than a truncated version. This research could also explore the factor structure of the measure to determine if the nine components are individual factors or combinations thereof. Further, researchers could compare the PCTE with other critical thinking tests (new or existing) to explore various aspects of measure validity (i.e., concurrent, criterion, predictive) in an attempt to determine its usefulness as a measure of critical thinking. This is particularly important as the measure was developed specifically as a test of psychological critical thinking and may be content-dependent.

Second, researchers could explore additional variables potentially related to critical thinking, including: (a) different academic majors or disciplines, (b) different instructors for critical thinking-based courses, (c) different textbooks or class materials that explicitly encourage critical thinking (Gilovich, 1991; Suter, 2006), and (d) regional, national, or international differences in the most effective pedagogies.

Given the importance of critical thinking skills to success in everyday life, and their centrality to the educational mission of higher education, we are hopeful that future research on these issues will identify effective pedagogical tools and techniques that will enable instructors to best assist their students in learning and applying critical thinking skills.

References

- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: McKay.
- Brown, K. W., Cozby, P. C., Kee, D. W., & Worden, P. E. (1999). *Research methods in human development* (2nd ed.). Mountain View, CA: Mayfield.
- Carlson, E. R. (1995). Evaluating the credibility of sources: A missing link in the teaching of critical thinking. *Teaching of Psychology*, 22, 39-41.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- de Sanchez, M. A. (1995). Using critical-thinking principles as a guide to college-level instruction. *Teaching of Psychology*, 22, 72-74.
- Dingfelder, S. F. (2005). With red pen in hand. *gradPSYCH*, 3, 28-29.
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: The Free Press.
- Halpern, D. F., & Nummedal, S. G. (Eds.). (1995). Psychologists teach critical thinking [Special issue]. *Teaching of Psychology*, 22, 1-83.
- Henderson, B. B. (1995). Critical-thinking exercises for the history of psychology course. *Teaching of Psychology*, 22, 60-63.

- Huitt, W. (2004). *Bloom et al.'s taxonomy of the cognitive domain*. Educational Psychology Interactive. Valdosta, GA: Valdosta State University. Retrieved August 19, 2005, from <http://chiron.valdosta.edu/whuitt/col/cogsys/bloom.html>
- Jackson, C. S., & Tlauka, M. (2004). Route-learning and the Mozart effect. *Psychology of Music*, 32, 213-220.
- King, A. (1995). *Inquiring minds really do want to know: Using questioning to teach critical thinking*. *Teaching of Psychology*, 22, 13-17.
- Kipp, K., & Lamson, N. (2005, October). "Edited volumes" as alternative term papers to promote critical thinking. Poster session presented at the annual Best Practices in Teaching Psychology Conference, Atlanta, GA.
- Lawson, T. J. (1999). Assessing psychological critical thinking as a learning outcome for psychology majors. *Teaching of Psychology*, 26, 207-209.
- Lawson, T. J., Schwiers, M., Doellman, M., Grady, G., & Kelnhofer, R. (2003). Enhancing students' ability to use statistical reasoning with everyday problems. *Teaching of Psychology*, 30, 107-110.
- Meston, C. M., & Frohlich, P. F. (2003). Love at first fright: Partner salience moderates roller-coaster-induced excitation transfer. *Archives of Sexual Behavior*, 32, 537-544.
- Oliver, J. (2005, October). Critical thinking about research methods: Designing studies. Poster session presented at the annual Best Practices in Teaching Psychology Conference, Atlanta, GA.
- Seal, D. W. (1997). Interpartner concordance of self-reported sexual behavior among college dating couples. *Journal of Sex Research*, 34, 39-55.
- Stanovich, K. E. (1998). *How to think straight about psychology* (5th ed.). New York: Longman.

- Suter, W. N. (2006). *Introduction to educational research: A critical thinking approach*. Thousand Oaks, CA: Sage.
- Teachman, J. (2003). Premarital sex, premarital cohabitation, and the risk of subsequent marital dissolution among women. *Journal of Marriage and Family*, 65, 444-455.
- Tomcho, T., Foels, R., Rice, D., Nonokawa, M., Johnson, J., Wetherbee, R., et al. (2005, October). Critical thinking teaching activities: Systematic review, exemplary models, & design recommendations. Paper presented at the annual Best Practices in Teaching Psychology Conference, Atlanta, GA.
- Tucker, C. J., McHale, S. M., & Crouter, A. C. (2003). Dimensions of mothers' and fathers' differential treatment of siblings: Links with adolescents' sex-typed personal qualities. *Family Relations*, 52, 82-89.
- Wade, C. (1995). Using writing to develop and assess critical thinking. *Teaching of Psychology*, 22, 24-28.
- Yanchar, S. C., & Slife, B. D. (2004). Teaching critical thinking by examining assumptions. *Teaching of Psychology*, 31, 85-90.
- Zablotsky, D. (2001). Why do I have to learn this if I'm not going to graduate school? Teaching research methods in a social psychology of aging course. *Educational Gerontology*, 27, 609-622.

Table 1

Critical Thinking Descriptive Statistics by Group

Variable	Section					
	Critique			Research Paper		
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
1. Four-item pretest	1.46	1.03	.00-3.00	1.39	1.02	.00-2.50
	(36.54%)	(22.75%)	(.00-75.00%)	(34.72%)	(25.60%)	(.00-62.50%)
2. Two-item pretest	.58	.40	.00-1.50	.33	.43	.00-1.00
	(28.85%)	(20.02%)	(.00-75.00%)	(16.67%)	(21.65%)	(.00-50.00%)
3. Total pretest (1 + 2)	2.04	.95	.50-3.50	1.72	.91	.00-2.50
	(33.97%)	(15.76%)	(8.33-58.33%)	(28.70%)	(15.09%)	(.00-41.67%)
4. Four-item posttest	1.38	.71	.00-2.00	.61	.49	.00-1.00
	(34.62%)	(17.80%)	(.00-50.00%)	(15.28%)	(12.15%)	(.00-25.00%)

KEY WORDS. Critical thinking, Blooms' taxonomy, Research paper, Critique

Correspondence concerning this article should be addressed to the first author: Trent W. Maurer, Department of Hospitality, Tourism, and Family and Consumer Sciences, Georgia Southern University, P.O. Box 8021, Statesboro, Georgia 30460 (E-mail: tmaurer@GeorgiaSouthern.edu).

5. Two-item posttest	.54 (26.92%)	.43 (21.56%)	.00-1.50 (.00-75.00%)	.22 (11.11%)	.51 (25.34%)	.00-1.50 (.00-75.00%)
6. Total posttest (4 + 5)	1.92 (32.05%)	.79 (13.11%)	.50-3.00 (8.33-50.00%)	.83 (13.89%)	.71 (11.79%)	.00-2.00 (.00-33.33%)
7. Four-item difference score (4 – 1)	-.08 (-1.28%)	1.34 (22.27%)	-2.50-2.00 (-41.67-33.33%)	-.78 (-12.96%)	1.12 (18.69%)	-2.50-1.00 (-41.67-16.67%)
8. Two-item difference score (5 – 2)	-.04 (-.64%)	.52 (8.65%)	-1.50-.50 (-25.00-8.33%)	-.11 (-1.85%)	.70 (11.62%)	-1.00-1.50 (-16.67-25.00%)
9. Total difference score (6 – 3)	-.12 (-1.92%)	1.33 (22.09%)	-2.50-2.00 (-41.67-33.33%)	-.89 (-14.81%)	1.02 (17.07%)	-2.50-.50 (-41.67-8.33%)

Note. In each row, raw scores are presented on top and conversions to percentages are presented beneath in parentheses.