What Constitutes Valid Evidence? An Investigation of Criteria Used by Undergraduate Students to Evaluate the Validity of Research

Steven W Brown,
Rockhurst University

Author's Contact Information

Steven W Brown, Ph.D.
Professor of Psychology
Rockhurst University
1100 Rockhurst Rd., Kansas City, MO, USA 64110
Phone: 816-501-4039
email: steven.brown@rockhurst.edu

Abstract:

   The purpose of the present investigation is twofold. First, I wanted to determine the characteristics undergraduates employ in determining the validity of evidence and then evaluate the effectiveness of an intervention to improve students' ability to consider methodological issues when evaluating the validity of research. To address the first question, students evaluated the validity of evidence presented in research scenarios that varied in terms of the specificity of methodology and analysis information discussed in the scenario. Comparing student evaluations of the scenarios to faculty evaluations indicated students consider many of the same criteria used by faculty to evaluate the validity of research, but weaknesses in these criteria have less of an impact on their validity ratings compared to faculty. In a second assignment, students evaluated issues discussed in a review article by identifying the author's argument, discussing evidence used to support the argument and evaluating the validity of the author's conclusions based on the evidence presented. Results from this assignment indicated many students do not consider methodological issues when evaluating the validity of research results presented in a journal review article. I then designed an intervention to promote critical, methodologically based evaluation of evidence. Students were then given a third assignment to assess potential changes in evaluation performance due to my intervention. Results indicated significant improvement in students' ability to incorporate methodological considerations into their evaluation as well as a reduction in confirmation bias when evaluating evidence. These results indicate that modest instruction on the role methodological considerations play in evaluating the validity of evidence can have a significant impact on students' critical evaluation skills.

Key Words:

critical thinking, evaluating evidence, scientific literacy, reasoning.

Introduction

Promoting critical thinking in students has been an explicitly recognized goal for higher education since the early 1980's (Paul, 1997). Evidence for this emphasis can be found in the requirement of least 3 credit hours in critical thinking coursework adopted by the California State University and Colleges system in 1980. Similar requirements have been adopted by institutions nationwide. The emphasis on teaching critical thinking rests on the belief that the ability to think critically is essential for decision making in a world where access to information from a wide range of sources is growing exponentially. Current students need to develop skills that enable them to make informed judgments in the context of complex and frequently contradictory information. There is growing evidence to suggest that students can acquire transferable critical thinking skills if these skills are taught explicitly and adequate opportunities for rehearsal are provided (Halpern, 2003; Marin & Halpern, 2011). However, there are still methodological issues with a considerable proportion of research investigating critical thinking that need to be addressed before educators can be confident that gains in students' ability to think critically are due to explicit instruction (see review by Behar-Horenstein & Niu, 2011).

Although there is no universally endorsed definition of critical thinking (Sternberg, 1986; Facione, 1990), there is enough overlap in definitions to identify specific skills good critical thinkers should exhibit. Halpern (2001) has offered an example of a thinking skills list that identifies nineteen skills most people would consider components of critical thinking. One of the skills identified on her list is "evaluating evidence." The ability to evaluate the validity of research evidence is one of the basic components of critical thinking educators need to emphasize across a variety of academic programs. Regardless of discipline, students need to be able to evaluate the quality of evidence presented to them in their upper division coursework as they grow from passive recipients of information to active evaluators able to make professional judgments concerning the validity of arguments and the evidence used to support arguments. Components of evaluating evidence should include the ability to evaluate the source of the information, the methodology used to gather and analyze the data, and the ability to evaluate the conclusions of the author in the context of data presented with a particular emphasis on generating alternative explanations for a set of results.

Guidelines designed to assist professionals in evaluating the quality of research evidence typically focus on evaluating information provided in full journal articles (Meltzoff, 1998; Pyrczak, 2003). However, much of the research evidence presented to undergraduate students comes from secondary source materials such as textbooks, research abstracts or literature reviews. These summaries frequently provide few methodological details on how data were gathered and analyzed. Thus, I was interested in investigating the criteria students use to evaluate the validity of research findings presented in a psychological review article where the results and conclusions from previous research is emphasized but information concerning methodological details of how the data were collected is often missing or incomplete.

The purpose of my investigation is twofold. First, I wanted to determine the characteristics my undergraduate students use to determine the validity of evidence presented in research scenarios and a literature review article. After determining criteria

students use to evaluate evidence, I then designed and evaluated the effectiveness of an intervention to improve critical evaluation skills.

Method

The First Literature Review Assignment

Twenty-eight junior- and senior-level undergraduate students in my Psychology of Motivation course participated in my first assessment. The students were asked to evaluate two issues discussed in a review article by identifying the author's argument relative to the issue, discussing the evidence used to support the argument and evaluating the validity of the author's conclusions based on the evidence presented. I identified the first issue they were required to address, and I allowed them to choose the second issue they wanted to evaluate. I provided students with a specific outline structure to help them organize their responses and explained that they should organize their final paper utilizing this outline. I stressed that the purpose of the assignment was to assess their ability to recognize and evaluate arguments, not write an integrated research paper as required in the majority of their upper division classes. I also supplied them with the grading rubric I would use to evaluate their performance on various components of the assignment. The rubric identified a variety of dimensions of the paper I wanted to assess, but here I will only focus on the dimensions related to the evaluation of evidence. The categories were Issue Identification and Description, The Author's Claim, The Evidence, The Student's Evaluation of the Evidence, and The Author's Conclusion and Student Response. These categories were rated on the three-point scale of Excellent, Needs Work, and Not Acceptable. A description of performance that would earn a particular rating on the scale was provided for each category.

Evaluation of the First Literature Review Assignment

After completing my evaluation of the first paper, fifty-four percent of the twenty-eight students in the class had earned an overall grade of A on their paper. The other forty-six percent scored a B or below, but there were actually very few B scores, indicating a high level of variability in student performance on the assignment. I next examined which components of the paper were resulting in the greatest loss of points by the students. Since some components of the grading rubric were worth more total points than other components, I converted the points earned for each component of the rubric into a percentage of points earned to evaluate where the greatest percentage of points were being lost on the assignment. This analysis indicated that the Student's Evaluation of the Evidence component accounted for the majority of points that separated the A and non-A students on this assignment.

Description of my Intervention

Based on my assessment of the first literature review assignment, it was clear nearly half of the students would potentially benefit from additional instruction, practice, and formative feedback to improve their ability to evaluate research evidence. Consequently, I designed an assignment and additional learning opportunities to assist students in developing the knowledge and skills needed to more effectively evaluate the validity of research.

Practice with Research Scenarios

Students were given a homework assignment where they were asked to evaluate the validity of evidence presented in nine one-paragraph research scenarios. The content of the scenarios varied systematically in terms of sample information, type of research design, procedural details, specificity of the results, and the number/recency of references cited. By manipulating these different aspects of the report across the different scenarios, I could potentially isolate how each category of information influenced the validity rating of the scenario. The rating scale employed ranged from "no evidence" to "weak evidence," "moderate evidence," or "strong evidence." Additionally, participants were asked to identify the characteristics of the evidence they used to determine their ratings. Although this assignment was not formally graded, students were given course credit for completing the assignment and we spent 20 minutes discussing their responses in class. Our discussion focused primarily on the scenarios where students disagreed on their ratings.

My Presentation on What Constitutes "Good" Evidence

Following the Practice with Research Scenarios assignment, I presented one fifty-minute lecture on the role research methodology plays in a researcher's ability to evaluate the quality of evidence. The content of this lecture was focused on the role that sample information, research design, procedural information and specificity of results play in evaluating research evidence.

Rewrite Opportunity

After reviewing the role research methodology plays in evaluating the validity of evidence, students who scored less than an A on the first Literature Review assignment were given an opportunity to rewrite the assignment and earn back points deducted from their first draft. All of the non-A students turned in a rewrite of their first attempt on the paper.

The Second Literature Review Assignment

The second assignment was identical to the first Literature Review assignment except that a different target article was given to the students for evaluation. The same rubric was employed for evaluation.

Analysis of Data Collected from the Intervention

Practice with Research Scenarios

Since the first goal of my investigation was to determine the criteria students use to evaluate the validity of research evidence, I compared the characteristics used by students to evaluate the validity of research evidence in the research scenarios to the characteristics used by psychology faculty that regularly teach the research methods courses. I felt as if identifying how the evaluation criteria between students and faculty differ might help me design learning experiences that would help students develop a more sophisticated understanding of the characteristics of "good" evidence as presented in secondary sources.

Analysis of participant responses compared the ratings given by undergraduate students that had scored 90% or greater on a previous assignment where they were asked to evaluate evidence (A students) to students that had scored 89% or below on the assignment (non-A students) and five departmental faculty that regularly teach research methods courses. Analysis of Variance comparing the ratings of these three groups indicated reliable differences among the A students, non-A students and faculty on four of the nine scenarios. Post hoc analysis designed to determine how the three group mean ratings differed from one another was conducted using the Tukey HSD procedure for each scenario.

Scenario 1

The first scenario contrasted student learning and behavior in a constructivist learning environment to a more traditional didactic learning environment. Research findings were presented as general outcome statements accompanied by multiple citations of the articles from which the information was drawn. A one-way ANOVA comparing the means of A, non-A and faculty ratings indicated a reliable difference between groups $F_{(2,30)} = 4.92$, $p = .014$). Post hoc analysis of the "quality of evidence" rating scale indicated that faculty ratings were significantly lower than non-A students on this scenario. However, there was no reliable difference between faculty and A students or A students and non-A students. Faculty rated this scenario as "no evidence" while A students rated the scenario as "weak evidence." Non-A students rated the scenario as "moderate evidence." Examination of the comments provided by faculty and students to justify their ratings indicated that the student ratings were more influenced by the multiple references cited in the scenario compared to faculty.

Scenario 2

The second scenario evaluated the effectiveness of different instructional media on student learning and retention. The scenario described two different types of instructional media and presented results as the percent recall of information produced by the different types of media. General results and conclusions were also provided. A one-way ANOVA comparing the means of A, non-A and faculty ratings indicated a reliable difference between groups $F_{(2,30)} = 4.27$, $p = .023$. Post hoc analysis indicated faculty ratings were significantly lower than non-A students on this scenario. However, there was no reliable difference between faculty and A students or A students and non-A students. Faculty rated this scenario as weak evidence while A students rated the scenario as moderate evidence. Non-A students ratings fell between moderate and strong evidence. Examination of the comments provided by faculty and students to justify their ratings indicated that both faculty and students believed the scenario needed more methodological information to evaluate the validity of the percentages reported. The students indicated that the inclusion of conclusions had a positive impact on their ratings while faculty thought the conclusions were over-stated given the data presented, and this negatively influenced their ratings.

Scenario 3

The third scenario described a survey administered to college students. Information concerning sample size and basic demographics of the sample were provided. The types of statistical procedures used to analyze the data were identified, but no actual statistics were presented. General results and conclusions from the survey were also

presented. A one-way ANOVA comparing the means of A, non-A and faculty ratings indicated a reliable difference between groups $F_{(2,30)} = 6.54$, $p = .004$. Post hoc analysis indicated faculty ratings were significantly lower than both the A and non-A students on this scenario. There was no reliable difference between the A and non A students. Faculty rated this scenario as between weak to moderate evidence while the students rated the scenario as strong evidence. Overall, faculty comments focused primarily on the lack of operational definitions for the constructs being measured in the survey (e.g. motivational intensity). Students were impressed with the details concerning the sample as well as the fact that results and conclusions were provided. The primary weakness of the evidence cited by students was the lack of statistical results to support the author's claims.

Scenario 4

The fourth scenario presented the construct of achievement motivation within the context of athletic competition. Achievement motivation was defined and several examples were provided illustrating how concepts derived from achievement motivation could be applied to sports. General findings from previous studies were presented along with references. A one-way ANOVA comparing the means of A, non-A and faculty ratings indicated a reliable difference between groups $F_{(2,30)} = 4.49$, $p = .02$. Post hoc analysis indicated faculty ratings were significantly lower than non-A students on this scenario. However, there was no reliable difference between faculty and A students or A students and non-A students. Faculty rated this scenario as weak evidence while non-A students rated the scenario as moderate evidence. The ratings of A students fell between the weak and moderate categories. Both A and non-A students were impressed with the number of sources cited to support the authors claims, apparently not realizing that only two different sources were cited repeatedly. The number of examples provided also positively influenced the ratings of students. A students and faculty were more frequently concerned about the lack of methodological information about the studies used to support the claims compared to non-A students.

When considering the overall pattern of numerical ratings from the four scenarios, it is clear that faculty were more critical concerning the quality of evidence presented than non-A students. Depending on the particular scenario, A student ratings were more similar to faculty in some scenarios and more similar to non-A students in others. Examination of the criteria used by respondents to determine their ratings indicate both students and faculty value more details concerning methodology and analysis than was presented in the scenarios. However, students weighed this lack of detail less strongly than faculty when assigning their numerical ratings of quality of evidence. Students were also more likely to rate scenarios that provided general conclusions related to the analysis presented more highly than faculty. Faculty ratings were negatively influenced by the lack of data provided to support specific conclusions whereas students appeared to be more accepting of conclusions even though necessary details needed to support the conclusions were not provided. Finally, multiple citations to articles not actually discussed in the scenario tended to produce higher validity ratings for the students compared to faculty.

Content Analysis of the Literature Review Assignments

After identifying some of the characteristics students use when evaluating research evidence using the research scenario assignment, I began to investigate my second question concerning the extent to which differences in research design and methodology impact student judgments about the validity of research. For example, data gathered from a representative sample of a population would generally be considered to have more generality (i.e., external validity) than data collected from a convenience sample of individuals. Similarly, research results generated using a true experimental design would generally be better able to determine causal relationships between variables (i.e., internal validity) than data collected using a quasi-experimental design where random assignment of subjects to groups is not possible. To assess the degree students used methodological considerations in evaluating the evidence presented in the target articles, I performed a content analysis on Assignment 1, the rewrite of Assignment 1, and Assignment 2. I classified student evaluations of evidence as either based on methodological considerations or non-methodological considerations. The categories of methodological considerations included reference to sample size, demographics, research design, methodology/procedural information, discussion of results, presentation of statistics/data, and presentation of conclusions. My categories on non-methodological considerations included results/conclusions agree with existing beliefs, reference from a reliable authority, multiple reference sources cited, and presentation of evidence well written. These categories were created based on the classification of actual student comments derived from the papers.
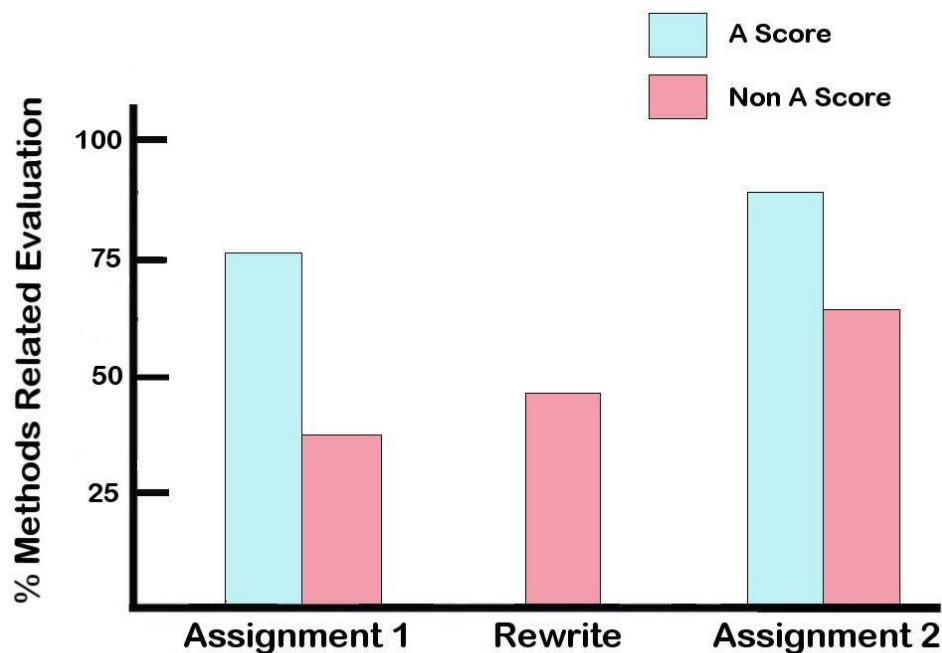

Figure 1: Comparison of Scores

Figure 1 presents the percentage of methods related evaluation of evidence for A and Non-A students across the three assignments. Of course, only Non-A students

performed a rewrite of Assignment 1 so only one group is presented. This figure illustrates that A students cited considerably more methods-related issues in evaluating the validity of evidence in the target articles compared to Non-A students. However, both groups of students improved their ability to evaluate evidence from a methodological perspective following my lecture and assignments designed to enhance this ability. I should also note the very modest increase in the Non A-students' ability to incorporate methods based evaluation in their rewrite of Assignment 1. Examination of rewrite scores on the grading rubric indicated that many students did not attempt to revise the "Students Evaluation of the Evidence" component of the assignment and opted to earn back points in other segments of the paper where less effort was required to improve their grade. Apparently, several Non-A students were satisfied with a lower grade on the assignment and felt that the time and effort required to revise this section exceeded the value of the points they could potentially earn.

My final analysis of the two student writing assignments involved a closer examination of the specific categories of both methodological and non-methodological categories of student evaluations. I identified the three most frequently occurring methodological and non-methodological categories based on my content analysis and compared A and Non-A students on both assignments. The following table provides a summary of how frequently the different categories were referenced by the students in their evaluation of the evidence presented in the articles.

Methods Evaluation Categories	A Scores		Non A Scores	
	Assign 1	Assign 2	Assign 1	Assign 2
Demographics/ Generality	30	23	3	11
Research Design/ Variables	11	19	1	17
Methodological/ Procedural	8	26	5	17

Non Methods Evaluation Categories	A Scores		Non A Scores	
	Assign 1	Assign 2	Assign 1	Assign 2
Conclusions Agree with Beliefs	9	1	21	9
Multiple Sources of Information	8	5	2	4
References/ Authority	1	6	0	8

Table 1: Comparison of Evaluation Categories

Comparing A and Non-A students on the three most frequently occurring Methods Evaluation categories indicates the extent to which A students recognize the influence of the demographic characteristics of the sample, research design, and methodological issues, on evaluating the validity of research findings. It is also evident that following my lecture and research scenarios practice assignment both A and Non-A students paid more attention to research design and methodological issues when evaluating the validity of research presented in the second assignment.

From my point of view, the most troubling result of this analysis is in the Non-Methods Evaluation table. On assignment 1, Non-A students seem to be relying on their existing belief system to evaluate the validity of new information to a greater extent than A students. This phenomenon is known as confirmation bias in the psychological literature (see Nickerson, 1998 for a review). What is troublesome about this result is that individuals that only accept evidence that confirms their existing beliefs are less likely to recognize a wrong belief they may hold and correct their previous faulty thinking. A hallmark of scientific thinking is the willingness of scientists to be wrong. To quote Albert Einstein, "No amount of experimentation can ever prove me right; however, a single experiment can prove me wrong." I believe this openness to being wrong in light of valid empirical evidence is a crucial component of critical thinking that must be developed in our students. While my lecture and research scenarios practice reduced the confirmation bias effect in both A and Non-A students, the bias was still occurring at what I consider an unacceptable level in Non-A students on the second assignment.

Conclusions

Results indicate that many students do not consider factors that contribute to the methodological quality of research when evaluating the validity of research findings. Interestingly, even circuit court judges seem to exhibit difficulty applying this skill set. Kovera and McAuliff (2000) provided judges with descriptions of an expert's research in which peer review status and internal validity were manipulated. They found that the methodological quality of the psychological research presented to the judges did not influence whether the judges would admit the research into evidence. However, they also found that judges given specific training in in the evaluation of research evidence rated valid evidence more positively than judges not provided with the training. So, it appears that many of my undergraduate students exhibit a lack of methodological sensitivity similar to the judges even though they have recently completed a research methods course in which scientific methodology and threats to internal validity are explicitly presented and rehearsed. The assumption that students would transfer skills learned in the research methods course to their upper division coursework was not true for a substantial number of students.

The finding that my students exhibited confirmation bias when evaluating the quality of research evidence presented to them is not surprising given it is probably the best known inferential error in human reasoning (Evans, 1989). Nickerson (1998) suggests that a student's tendency to seek evidence consistent with their existing beliefs and discount information that is not consistent with their beliefs is frequently reinforced by assignments used in higher education. He argues that assignments that urge students to present evidence for their opinions leads students to seek confirming evidence for their ideas and

disregard disconfirming evidence thus reinforcing the natural tendency of confirmation bias. He suggests educators clearly distinguish between assignments designed to foster "case-building" and assignments that foster "evidence-weighing" and choose the approach that best serves their course goals.

The Follow-up Study

Based on what I had learned from the previous semester, I modified my Psychology of Motivation course the following semester to explicitly teach the skills I expected students to utilize on the literature review assignment before they began working on the assignment. Basically, I introduced the lecture reviewing methodological concepts to consider when evaluating the validity of research findings along with the research scenario exercise prior to assigning the article review project. The lecture required one 50-minute class period. The research scenario exercise was given as homework and we spent 30 minutes discussing the students' responses to the research scenarios the following class period. I then gave students the same article review assignment used the previous semester. Comparing the "before intervention" A students from the previous semester to the "after intervention" A students in the current semester should provide additional evidence whether the improvement in students' ability to use methodologically based evaluation of information presented in a review article benefits from my instruction/exercise intervention. There were thirty junior and senior level undergraduate students in this section of the course.

Content Analysis of the Literature Review Assignment

After grading the review article, it was clear that the overall performance on the paper had improved dramatically. Whereas only 54 percent of the students made an A on the paper the previous semester, 84 percent of the students earned an A on the identical assignment this semester. However, I was most interested in determining whether my lecture/exercise intervention enabled students to use more methodological considerations when evaluating research evidence presented in the review article. To investigate this question, I again performed a content analysis on the student's evaluation of the evidence component of the article review assignment. I then decided to compare the performance of the A students in the current semester to the A students in the previous semester since I only had five students in the current class fail to make an A on the article review assignment. Given different numbers of students were to be compared across semesters (15 verses 25) I needed a way to standardize the comparison that would take different sample sizes into account.

Given that students were asked to evaluate three research studies used as evidence for two different author arguments, there was a maximum of six possible instances that any particular category in my content analysis could potentially be mentioned. I then multiplied the number of potential mentions by a single student (6) times the total number of students in the group to determine the total possible potential mentions for any particular category in my content analysis. What is reported in the table below is the proportion of actual mentions of each category (actual mentions/total possible mentions) for students in the previous and current semesters of the course.

15

Methods Evaluation Categories

	Previous Semester	Current Semester
Demographics/ Generality	.33	.30
Research Design/ Variables	.12	.38
Methodological/ Procedural	.09	.32

Non-Methods Evaluation Categories

	Previous Semester	Current Semester
Conclusions Agree with Beliefs	.1	.05
Multiple Sources of Information	.09	.11
References/ Authority	.01	.19

Table 2: Comparison of Evaluation Categories by Semester

By looking at the methods evaluation categories, it can be seen that students improved their ability to use methods based evaluation of research after experiencing my lecture/exercise before writing their article review assignment. While students in both semesters of the course frequently cited sample demographics as a limiting factor in generalizing research findings, A students in the current semester were much more likely to point out limitations due to research design and methodological considerations compared to students that earned an A on the assignment the previous semester.

Examining the Non-Methods Evaluation categories indicates generally low incidence of using non-methods based information when evaluating the validity of evidence for A student across semesters. However, the current semester students did exhibit a tendency to mention multiple citations and the impression that the author of the evidence appeared to be an authority on the topic as a rational for their validity assessment. This finding is consistent with the result obtained from the research scenario evaluation where students tended to rate summaries with multiple citations as more valid compared to faculty. The infrequent mention of statements such as "I believe this research is valid because the findings are consistent with my personal experience"

in the current semester students suggests I was successful in making the vast majority of students aware of the flaws in this type of reasoning.

Discussion

My results indicate I was successful in encouraging students to consider methodologically based criteria when evaluating the validity of research findings presented in a literature review article. When comparing the same students on a first and second review article assignment, I think introducing the lecture and research scenario exercise prior to assigning the second article review project had a significant impact on the students' ability to perform well on the project. The comparison of students across two semesters of the course on the same assignment indicated a substantial increase in the number of A grades earned on the article review project by students in the second semester. This increase in grades was due primarily to the increased use of methodologically based information when evaluating the validity of research presented in the review article. For the very modest time investment of one class period to review concepts taught in the research methods course and an additional 30 minutes to discuss the research scenario homework in class, student performance benefitted tremendously.

The success I experienced with my modest intervention has motivated me to be more structured in presenting guidelines for evaluating research evidence. I have now developed a handout that summarizes criteria that can be used to evaluate research evidence as well as discussion of various flaws in human reasoning such as confirmation bias. I am also considering developing a computer based tutorial specifically designed to teach and develop specific evaluation skills needed to evaluate psychological research. It is my goal not only to make students more aware of the importance of methodologically based criteria when evaluating research, but also to provide more experience with "evidence-weighing" as described by Nickerson (1998).

References

Behar-Horenstein, L. S. & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning*, *8(2)*, 25-41.

Evans, J. St., B. P. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Erlbaum.

Facione, P. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Retrieved from the Web March 9, 2011. http://www.eric.ed.gov/PDFS/ED315423.pdf

Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education*, *50(4)*, 270-286.

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Mahwah, N.J.: Lawrence Erlbaum Associates.

Kovera, M. B. & McAuliff, B. D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *Journal of Applied Psychology*, *85(4)*, 574-586.

Marin, L. M. & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*, 1-13.

Meltzoff, J. (1998). Critical thinking about research: Psychology and other fields. American Psychological Association. Washington, DC.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2(2)*, 175-220.

Paul, R. (2009). The Critical Thinking Movement: 1970-1997: Putting the 1997 Conference into Historical Perspective. Retrieved from the Web March 3, 2011. http://www.criticalthinking.org/articles/documenting-history.cfm

Pyrczak, F. (2003). *Evaluating research in academic journals: a practical guide to realistic evaluation*. Pyrczak Publishing: Los Angeles.

Sternberg, R.J. (1986). *Handbook of Intelligence*. New York, Cambridge University Press.