


Training Scientific Thinking Skills: Evidence From an MCAT²⁰¹⁵-Aligned Classroom Module

Teaching of Psychology
2014, Vol. 41(2) 115-121
© The Author(s) 2014
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0098628314530341
top.sagepub.com


Courtney Stevens¹ and Melissa R. Witkow¹

Abstract

The present study reports on the development and evaluation of a classroom module to train scientific thinking skills. The module was implemented in two of four parallel sections of introductory psychology. To assess learning, a passage-based question set from the medical college admissions test (MCAT²⁰¹⁵) preview guide was included as extra credit on the final exam in all sections. This provided an outcome that was distinct in content from the module, while tapping the same underlying scientific thinking skills. Students in the experimental classrooms answered more questions correctly on targeted scientific thinking skills than students in the comparison classrooms. These data support the benefit of targeted activities for training scientific thinking skills in introductory psychology.

Keywords

critical thinking, pedagogy, teaching of psychology, graphical literacy

As a scientific discipline, an important component of psychology programs is training in the scientific method, including research design, data analysis, and study interpretation (American Psychological Association, 2007). Even at the introductory level, the American Psychological Association recommends that the course should “reflect the nature of psychology as a scientific discipline” (American Psychological Association, 2011, p. 13). The reports of several national panels echo the importance of incorporating the scientific process into first-year science courses (American Association for the Advancement of Science, 2009; National Research Council, 2011, 2012). Indeed, the ability to think scientifically is important not only in the science classroom but also in the real world where basic scientific literacy skills are critical for personal decision making and basic citizenship competence (National Research Council, 1996). As the National Science Board wrote in a report on the state of science education and public scientific literacy, “Appreciating the scientific process can be even more important than knowing scientific facts” (National Science Board, 2008, p. 17, Chapter 7).

Recent announcements about impending changes to the medical college admissions test (MCAT) have spurred additional discussion concerning the training of scientific thinking skills in introductory psychology (e.g., Frazer & Twohig, 2012). Beginning in 2015, the MCAT will include a new section on social and behavioral sciences, with an estimated 60% of questions linked to introductory psychology, a topic not previously included on the MCAT (Association of American Medical Colleges, 2011, 2012). This new section will be equal

in weighting to other sections on biological/biochemical processes and chemical/physical processes. As in the other MCAT sections, many of the psychology questions will take the form of “passage sets.” In passage set questions, students read multiparagraph descriptions of research studies and findings—including graphical representations of data—and then answer corresponding skill-based, multiple-choice questions about the passage. About half of these questions will require the application of the scientific method, including interpreting data, critiquing research designs, and evaluating conclusions drawn from hypothetical studies. The remaining questions will relate to the application and recognition of concepts, principles, and theories from psychology.

As described in the MCAT preview guide, each question will target one of four specific skills. The first skill assesses scientific concepts and principles (e.g., recognizing correct scientific principles, identifying examples of observations that illustrate scientific principles). The second skill assesses reasoning and evidence-based problem solving (e.g., using relevant theories to explain phenomena or make new predictions). The third skill assesses reasoning about the design and execution of research (e.g., identifying testable research

¹Department of Psychology, Willamette University, Salem, OR, USA

Corresponding Author:

Courtney Stevens, Department of Psychology, Willamette University, 900 State Street, Salem, OR 97301, USA.
Email: cstevens@willamette.edu

questions and hypotheses, distinguishing experimental and nonexperimental research studies, and critiquing conclusions that can be drawn from particular types of studies). Finally, the fourth skill assesses data-based and statistical reasoning (e.g., interpreting data presented in figures, graphs, or tables, applying descriptive statistics, and using data to draw conclusions).

Thus, similar to definitions of scientific literacy (e.g., National Research Council, 1996), which involve both content and process knowledge, the new MCAT is structured around the application of scientific content knowledge (MCAT Skills 1–2) as well as the application of the scientific method (MCAT Skills 3–4). Indeed, the latter two skills are both described in the MCAT preview guide as requiring students to “show that you can ‘do’ science” (Association of American Medical Colleges, 2012, pp. 16–17). As such, these questions require students to move beyond content toward the analysis, synthesis, and evaluation skills that lie at the mid to high range of Bloom’s Taxonomy (Anderson & Krathwohl, 2001; Krathwohl, 2002). Moreover, as Frazer and Twohig (2012) have cogently argued, these higher level MCAT skills are important not only to aspiring physicians but also to all students and citizens who must interpret and evaluate information in many contexts. Thus, the new MCAT²⁰¹⁵ underscores the need—already recognized by the psychology community—to incorporate training in the scientific method and scientific thinking skills into introductory courses.

Cast in a broader framework, these scientific thinking skills can be considered a subset of critical thinking skills. Although there are many definitions of critical thinking, Halpern described critical thinking as “purposeful, reasoned, and goal-directed” (1998, p. 450). Halpern goes on to say, “It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions” (1998, pp. 450–451). With respect to teaching critical thinking skills, a large literature suggests methods for promoting critical thinking across the psychology curriculum, including the introductory level (e.g., see Carroll, Keniston, & Peden, 2008; Dunn, Halonen, & Smith, 2008; Halonen, 1995; King, 1995). For example, some general psychology instructors ask students to conduct term-long research projects (Lamson & Kipp, 2008) or investigate whether common myths in psychology (e.g., eating carrots improves eyesight) are supported by the scientific literature (Blessing & Blessing, 2010; see also Wilkinson, Dik, & Tix, 2008). Other general psychology instructors have developed assignments that challenge students to read an original article cited in the textbook and then critique the text’s presentation of the article (Gareis, 1995). Still others ask students to respond to specific written prompts that target aspects of critical thinking (Wade, 1995). To improve specific aspects of data interpretation, some instructors develop activities that engage students in interpreting and/or constructing graphical representations of data (Holmes, 2008; Lutsky, 2006; Nolan & Heinzen, 2009). As these diverse examples illustrate, the existing literature provides a range of activities for introductory psychology that tap different elements of critical thinking. At the same time, very few of these reports include empirical assessment

to evaluate whether the activities actually impact the targeted skills (but note there are exceptions, e.g. Blessing & Blessing, 2010; Thieman, Clary, Olson, Dauner, & Ring, 2009). As well, it is unclear whether these types of activities will facilitate performance on questions similar to those on the new MCAT²⁰¹⁵, which require students to apply the scientific method across different psychology content domains. However, if it is granted that the types of skills tested on the MCAT²⁰¹⁵ are laudable goals for *all* psychology students, any activities that improve these skills will be valuable to psychology teaching community.

In this article, we report on the development and assessment of an MCAT²⁰¹⁵-aligned module targeting scientific thinking skills in research design and data-based reasoning for introductory psychology. This represents the first step in our long-term goal of developing a set of validated teaching resources that can enhance scientific thinking skills in introductory psychology across the range of content domains typically covered in the course. Drawing on the framework of scientific thinking skills provided by the development of the MCAT²⁰¹⁵, we designed a 45-min module for introductory psychology targeting research design and data-based reasoning skills (MCAT Skills 3–4; Association of American Medical Colleges, 2011, 2012). By including specific, targeted activities in introductory psychology that require and reinforce these research design and data-based skills critical to the scientific method, we aimed to embed scientific thinking training in psychology content with meaningful, real-world applications. We implemented the module in two of four parallel sections of introductory psychology at a small liberal arts college. Students in all sections completed an MCAT passage set question series as extra credit on the final exam. We hypothesized that students receiving the scientific thinking module would show improved performance on the questions tapping research design and data-based reasoning relative to students in classes not implementing the module, while no differences between groups were expected on the more content-focused questions. We further hypothesized, given the focal effect of the scientific thinking module on particular types of questions, that the relative pattern of performance across the content-focused versus research design and data-based reasoning questions would differ between groups. Specifically, we hypothesized that students in the comparison classrooms would show poorer performance on the research design and data-based reasoning questions compared to the content questions, whereas this difference would be smaller or nonexistent in students receiving the scientific thinking module.

Method

Participants

Ninety-three undergraduate students from four sections of introductory psychology at Willamette University participated in the present study. Two of the sections received a scientific thinking module ($n = 45$ students, one section taught by each author), with the other two serving as comparison sections

($n = 48$). Each section enrolled 22–26 students. A different instructor taught each section, but all instructors had comparable years of full-time teaching experience and all sections were taught in the same semester.

Across sections, the sample was 57% female, with students primarily in their first year of study (48%) and of undeclared major (38%). There were no differences across the four sections in any of these background characteristics. Specifically, there were no significant differences in student gender, $\chi^2(3, N = 93) = .956, p = .812$, or year in college (first year compared to all others), $\chi^2(3, N = 89) = 1.63, p = .653$. Similarly, there were no differences across the four sections in whether students had a declared versus an undeclared major, $\chi^2(3, N = 79) = 1.93, p = .587$, or for those with declared majors, across course major category (social science compared to all other declared majors), $\chi^2(3, N = 49) = 4.43, p = .219$. A parallel set of analyses confirmed that there were still no differences between groups on these background variables if sections were grouped based on those receiving the scientific thinking module and those serving as comparison classes. Specifically, comparing the treatment and comparison groups, there were no significant differences in student gender, $\chi^2(1, N = 93) = .322, p = .57$, or in year in school, $\chi^2(1, N = 89) = .546, p = .460$. Similarly, there were no differences across the treatment and comparison classes in whether students had a declared versus an undeclared major, $\chi^2(1, N = 79) = .07, p = .792$, or for those with a declared major, in course major category, $\chi^2(1, N = 49) = 2.66, p = .103$.

Measures

Scientific thinking module. Drawing on a set of four best practice recommendations for science teaching, the authors developed a modular in-class activity to target research design and data-based reasoning skills. First, the module presented scientific content in a real-world, relevant context (American Association for the Advancement of Science, 2009; American Psychological Association, 2011; Association of American Colleges & Universities, 2011; National Research Council, 2011). Second, the module engaged students as active learners (American Association for the Advancement of Science, 2009; American Psychological Association, 2011; National Research Council, 2011). Third, the module incorporated quantitative reasoning to domain practice (American Association for the Advancement of Science, 2009; American Psychological Association, 2011; Association of American Colleges & Universities, 2011). Fourth, the module supported students in the process of evidence-based decision making and generation and consideration of alternative explanations (American Psychological Association, 2011).

In the activity, students evaluated the design and findings of a published research study comparing the effects of saffron to fluoxetine (Prozac) for the treatment of depression (Noorbala, Akhondzadeh, Tahmacebi-Pour, & Jamshidi, 2005). The module mirrored the MCAT passage set questions by providing students with a typed handout containing a brief research synopsis

of the study design including a graphical representation of the study findings. A set of discussion questions engaged students in data-based reasoning and critical evaluation of research design. The student discussion questions required interpreting patterns of data from the graphs, drawing appropriate conclusions based on aspects of the study design (e.g., lack of a no-treatment control group) and proposing follow-up experiments to test issues unresolved by the initial study. These questions were first discussed by students in small groups. Follow-up discussion involving the whole class provided an opportunity for instructor feedback as well as class discussion on the pros and cons of different follow-up studies suggested by student groups. An accompanying two-page instructor guide provided a framework for leading students through the activity and highlighting key instructional points through discussion.

Assessment. To assess learning outcomes, in all four sections of introductory psychology, a passage set question series was drawn verbatim from the first edition of the MCAT preview guide (Association of American Medical Colleges, 2011). The passage set questions provided a learning outcome that was distinct in content from the module implemented (treatment for depression vs. susceptibility to the cold virus), while still tapping the same underlying research design and data-based reasoning skills targeted in the module. For example, the sample module required students to interpret data presented in the form of bar graphs and tables, and the MCAT questions used for assessment also required interpreting data from a bar graph and drawing conclusions about the relationships among variables (data-based and statistical reasoning). Similarly, the sample module engaged students in small group work designing and proposing follow-up studies, and the MCAT questions used for assessment required students to determine which of several proposed follow-up studies would allow causal conclusions to be drawn about the relationship between several variables (reasoning about the design and execution of research). However, although the classroom module was purely discussion based, the MCAT questions were in a multiple-choice format.

Specifically, the assessment passage included a research description of a study examining susceptibility to the cold virus as a function of perceived and experienced stress. The description included a bar graph presenting results from the study. This passage was selected as it included one question tapping each of the four skills assessed on the MCAT (the original Question 4 was omitted from analysis as it provided double coverage of Skill 3). The research design and data-based reasoning questions (MCAT Skills 3 and 4) involved drawing conclusions and interpreting data from the graph. In contrast, the content-based questions did not require students to interpret the data presented in the graph per se but instead asked about the nervous system as it related to the study findings and to extend the study to the concept of learned helplessness (MCAT Skills 1 and 2). All students attempted to answer all questions (i.e., all students at least offered their best guess for each multiple-choice item, with no questions left blank by any student). None of the four sections of introductory psychology specifically covered stress

and health, but all sections included coverage of concepts in social psychology as well as specifically on the sympathetic and parasympathetic nervous system as part of units on biopsychology.

Procedure

During the final 2 weeks of the semester, instructors in two sections of introductory psychology implemented the scientific thinking module. The module took 30–45 min of class time and was held during a portion of a regular class period in the psychopathology unit, a unit covered by all four instructors. The module replaced time that would traditionally have been devoted to lecture on psychopathology treatments. Otherwise, no changes were made to any of the class sessions during the semester to specifically target the MCAT skills described. Two other sections of introductory psychology participated but did not include the module, although they also covered psychopathology as a standard unit. These other sections served as comparison classrooms. To assess learning outcomes, in all four sections of introductory psychology, extra credit multiple-choice questions were included “cold” on the final exam.

Data Analysis Strategy

Two composite scores for each student were calculated, representing performance on the two questions tapping research design and data-based reasoning, and the two questions tapping more content-based skills not specifically targeted in the module. Data were collapsed into targeted and nontargeted questions because there were no a priori predictions concerning differential performance on the individual skills. This also reduced the number of comparisons made and increased the precision of point estimates for targeted skills (i.e., by averaging over two questions per category, rather than using a single item to tap each skill).

All analyses were conducted using both nonparametric and parametric statistics. Given our data set, with outcomes limited to 0, 1, or 2, the primary analyses are reported using the Mann–Whitney U test, followed by parametric analysis using independent samples t -tests, which are more directly interpretable. The between-subject factor was instruction type (experimental class vs. comparison class). Experimental classes received the scientific thinking module, whereas comparison classes did not. The within-subject factor was question type (targeted vs. nontargeted), with targeted questions referring to those items tapping research design and data-based reasoning and nontargeted questions referring to the more content-oriented questions not specifically targeted in the science literacy module. Nontargeted questions were included to establish similarity of content knowledge across sections and to provide a strong test of specificity of gains for the experimental group. Additional analyses comparing performance on the two types of questions within instruction type were conducted using related-samples Wilcoxon signed-rank tests, followed by paired samples t -tests.

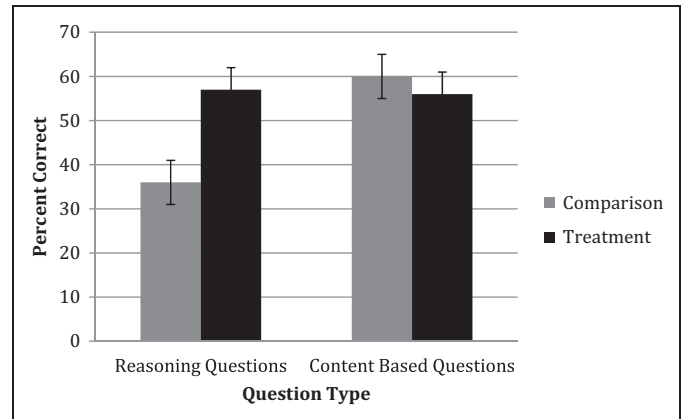


Figure 1. Percentage of students correctly answering medical college admissions test (MCAT) questions in final exam. Students in experimental classes including the scientific thinking module answered more questions tapping research design and data-based reasoning correctly than students in comparison classes not implementing the module. No differences were observed between groups on the content-based questions. Error bars represent standard error of the mean.

Results

Figure 1 presents the results of the MCAT assessment, separately for the Experimental and Comparison classes. Responses to questions tapping research design and data-based reasoning (targeted questions) versus more content-based questions are presented separately. Subsequently, the results of parametric and nonparametric analyses of these data are presented. As detailed subsequently, in all cases the results of the parametric and nonparametric analyses converged on the same conclusions.

First, we tested the hypothesis that students receiving the scientific thinking module would show improved performance on the questions tapping research design and data-based reasoning relative to students in classes not implementing the module, while these differences would not be observed between groups on the more content-focused questions. Consistent with this hypothesis, a Mann–Whitney U test indicated a significant difference between groups in performance on the research design and data-based reasoning questions, $U = 1,398$, $p = .009$, $r = .27$, but no significant difference between groups in content-based questions, $U = 1,000$, $p = .504$, $r = .07$. Independent samples t -tests showed the same pattern of results, with students in classes receiving the experimental module answering more research design and data-based reasoning questions correctly than students in the comparison classes did, $t(91) = -2.7$, $p < .01$, 57% versus 36% correct. This represented a large effect size: Cohen’s $d = +.59$. In contrast, there was not a significant difference between classes in response to the more content-based questions representing MCAT Skills 1 and 2, which were not targeted in the module, $t(91) = .66$, not significant, 60% versus 56% correct, $d = +.11$.

Another way of examining the data is to examine the relative pattern of performance across the more content-oriented

(nontargeted) questions versus the research design and data-based reasoning (targeted) questions. Thus, a second set of analyses tested the hypothesis that, given the focal effect of the scientific thinking module on particular types of questions, students in the comparison classrooms would show poorer performance on the research design and data-based reasoning questions than the performance on content questions, whereas this difference would be smaller or nonexistent for students receiving the scientific thinking module. Consistent with this hypothesis, Wilcoxon signed-rank tests indicated that only students in the comparison classes showed weaker performance on the research design and data-based reasoning questions relative to the more content-oriented questions (treatment classes, $Z = .22$, $p = .824$, $r = .03$; comparison classes, $Z = 3.25$, $p = .001$, $r = .47$). Paired t -tests showed the same pattern of results. Students in the comparison sections showed poorer performance on questions requiring research design and data-based reasoning relative to the more content-focused questions, 36% correct versus 60% correct, $t(47) = -3.6$, $p = .001$, $d = -.68$. In contrast, the treatment classes receiving the scientific thinking module did not show significant differences between research design and data-based reasoning questions relative to the more content-focused questions, 57% correct versus 56% correct, $t(44) = .17$, $p = .86$, $d = +.03$.

Discussion

The present study demonstrates that the type of scientific thinking required by passage-based MCAT questions can be improved through targeted classroom activities. Relative to students in traditional introductory psychology courses, students in classrooms that included a single in-class activity that targeted research design and data-based reasoning showed improved performance on exam questions tapping these same skills. Moreover, the effect of this training was large in magnitude, leading to over a half standard deviation advantage in performance for students receiving the module. As the exam questions were on a different research study from that used in the training activity, it appears that the underlying skills in reasoning about the design and interpretation of research and data, rather than rote memorization, were improved by the activity. Further, there was no difference between classes in performance on questions tapping general content, suggesting that higher scores on scientific thinking skills were not simply due to stronger overall performance or content knowledge in experimental classes. Together, these findings support the hypothesis that students receiving a scientific thinking module exhibit improved performance, specifically on the questions tapping research design and data-based reasoning relative to students in comparison classes.

Interestingly, students in comparison classes performed worse on the research design and data-based reasoning questions relative to the content questions, whereas this difference was not observed in the classes receiving the scientific thinking module. This finding supported the hypothesis that students in comparison classrooms would perform more poorly on research design and data-based reasoning questions relative to more

content-based questions, whereas this effect would be smaller or nonexistent among students in the classes receiving the scientific thinking module. Indeed, this suggests that students in comparison classes were disproportionately underprepared to answer questions requiring research design and data-based reasoning, although these are arguably the very skills most important to the real-world use and application of psychology.

One implication of this finding for instructors and textbook authors is the importance of considering their goals in determining the relative balance of the dissemination of content information versus the training of scientific thinking skills. Training scientific thinking skills may require additional, deliberate activities that provide students practice using these skills. Indeed, the scientific thinking skills trained in the present study represent a subset of skills that fall under the larger umbrella of critical thinking (Bensley & Murtagh, 2012). As such, this study adds to a growing body of literature indicating that aspects of critical thinking can be trained and importantly at different levels within the psychology curriculum (e.g., Bensley, Crowe, Bernhardt, Buckner, & Allman, 2010; Blessing & Blessing, 2010; Carroll et al., 2008; Dunn et al., 2008; Gareis, 1995; Halonen, 1995; Holmes, 2008; King, 1995; Lamson & Kipp, 2008; Lutsky, 2006; McLean & Miller, 2010; Nolan & Heinzen, 2009; Wade, 1995; Wilkinson et al., 2008).

Limitations

It is important to note the limitations of the present study. First, as is common in educational research, it was not possible to randomly assign participants to conditions. Thus, we cannot conclusively attribute the strong performance of students in the experimental sections to inclusion of the treatment module. This is a particularly important issue, given that we did not collect pretest data on MCAT style or scientific thinking questions in the four sections. Thus, we cannot demonstrate equivalency across sections at pretest on the specific skills targeted in the modules. However, students did not differ across sections in any background variable available to us (gender, class year, or major), nor did they differ on performance in the more content-oriented questions, suggesting some similarity in mastery of course content. As well, all classes were taught at the same university, during the same academic term, and by instructors of comparable experience. Of course, it is possible that differences could have emerged if other background information about students, instructors, or courses was collected. For example, it is possible that the instructors differed in terms of how much they emphasized critical thinking or scientific reasoning in their courses, apart from the specific module included in treatment classes. Identifying differences between instructors is particularly relevant in this study, given that the authors themselves were the instructors of the experimental sections. A stronger test than the one used in the current study would be to examine change in performance from a pretest in two sections (one experimental and one comparison) taught by the same instructor. Our results must be considered preliminary until a more rigorous design such as this is used.

Second, it is important to note that the scientific thinking modules contained a number of elements, any one or combination of which may be important for supporting students' scientific thinking skills. For example, it is possible that improvements in scientific thinking could have been realized as a simple consequence of gaining practice and experience reading graphs and interpreting passages, rather than something unique about the discussion questions students engaged during class. That is, students in the experimental classes had experience with a similar question format (albeit in a discussion-based, rather than multiple-choice, format), which may have helped them perform better during the assessment. Although this explanation may be less interesting to some, we believe that, if accurate, it would still point to the importance of providing students increased experience with the type of graph reading activities encompassed in the current modules. As noted by others (e.g., Lutsky, 2006), there is increasing emphasis on the importance of incorporating quantitative literacy and graphical literacy skills into psychology courses so that students are able to evaluate the quantitative components of arguments and interpret newly encountered data.

Finally, although short-term benefits in scientific thinking skills were observed in the experimental classrooms, we do not know how long these effects will last. This question is important, as the skills trained are ones that we hope students will be able to use after leaving introductory psychology, both in their daily lives and, for students taking the MCAT, on the high-stakes exams that are often not encountered until several years after leaving introductory psychology.

Future Directions and Conclusion

In future research, we plan to adapt the module and other activities for use in larger or online class formats. We also plan to develop and assess additional modules. The broad range of topics covered in psychology courses provides a natural opportunity to embed practice in scientific thinking skills in multiple distinct content domains (e.g., the effects of media violence on children's aggression or sleep the effects of deprivation on cognitive performance). As such, introductory psychology offers a unique platform for providing the repeated exposure to key skill-based concepts in multiple real-world contexts that is recognized to be critical to effective training of higher order and scientific reasoning skills (Halpern, 1993, 1999; Willingham, 2007). That is, if students are to become better scientific thinkers, they need the opportunity both to use these skills repeatedly and to see their relevance to real-world issues. The activity developed here, and the evidence for its effectiveness, represents a first step toward this goal. This finding is promising for efforts to meet the goal of moving beyond content in introductory courses to ensure students leave with not only the knowledge but also the skill set that will be useful after leaving the classroom.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: A portion of this work was supported by a generous grant from the W.M. Keck Foundation in support of Willamette University's *iScience* Initiative.

References

- American Association for the Advancement of Science. (2009). *Vision and change in education biology education: A call to action*. Washington, DC: Author.
- American Psychological Association. (2007). *APA guidelines for the undergraduate psychology major*. Washington, DC: Author. Retrieved from <http://www.apa.org/ed/resources.html>.
- American Psychological Association. (2011). *APA principles for quality undergraduate education in psychology*. Washington, DC: Author.
- Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives: Complete edition*. New York, NY: Longman.
- Association of American Colleges & Universities. (2011). *What works in facilitating interdisciplinary learning in science and mathematics*. Washington, DC: Author.
- Association of American Medical Colleges. (2011). *MCAT 2015: A better test for tomorrow's doctors. Preview guide for MCAT 2015*. Washington, DC: Author.
- Association of American Medical Colleges. (2012). *MCAT2015: A better test for tomorrow's doctors. Preview guide for the MCAT2015 exam* (2nd ed.). Washington, DC: Author.
- Bensley, D. A., Crowe, D. S., Bernhardt, P., Buckner, C., & Allman, A. (2010). Teaching and assessing critical thinking skills for argument analysis in psychology. *Teaching of Psychology, 37*, 91–96.
- Bensley, D. A., & Murtagh, M. P. (2012). Guidelines for a scientific approach to critical thinking assessment. *Teaching of Psychology, 39*, 5–16.
- Blessing, S., & Blessing, J. (2010). PsychBusters: A means of fostering critical thinking in the introductory course. *Teaching of Psychology, 37*, 178–182. doi:10.1080/00986283.2010.488540
- Carroll, D. W., Keniston, A. H., & Peden, B. F. (2008). Integrating critical thinking with course content. In D. S. Dunn, J. S. Halonen, & R. A. Smith (Eds.), *Teaching critical thinking in psychology* (pp. 101–115). Malden, MA: Wiley-Blackwell.
- Dunn, D. S., Halonen, J. S., & Smith, R. A. (Eds.). (2008). *Teaching critical thinking in psychology*. Malden, MA: Wiley-Blackwell.
- Frazer, N. B., & Twohig, M. P. (2012). The new social and behavioral requirements: Inspired innovation, missed opportunity, or both? *Peer Review, 14*, 14–17.
- Gareis, K. C. (1995). Critiquing articles cited in the introductory textbook: A writing assignment. *Teaching of Psychology, 22*, 233–235.
- Halonen, J. S. (1995). Demystifying critical thinking. *Teaching of Psychology, 22*, 75–81.
- Halpern, D. (1993). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education, 50*, 270–286.
- Halpern, D. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning, 80*, 69–74.

- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and meta-cognitive monitoring. *American Psychologist, 53*, 449–455.
- Holmes, J. D. (2008). An introductory exercise for promoting critical thinking about psychological measurement. In D. S. Dunn, J. S. Halonen, & R. A. Smith (Eds.), *Teaching critical thinking in psychology* (pp. 277–280). Malden, MA: Wiley-Blackwell.
- King, A. K. (1995). Designing the instructional process to enhance critical thinking across the curriculum. *Teaching of Psychology, 22*, 13–17.
- Krathwohl, D. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*, 212–218.
- Lamson, N., & Kipp, K. (2008). A module-based research project: Modeling critical thinking in psychology. In D. S. Dunn, J. S. Halonen, & R. A. Smith (Eds.), *Teaching critical thinking in psychology* (pp. 247–250). Malden, MA: Wiley-Blackwell.
- Lutsky, N. (2006). Teaching quantitative reasoning: How to make psychology statistically significant. *American Psychological Society Observer, 19*, 35–38.
- McLean, C. P., & Miller, N. A. (2010). Changes in critical thinking skills following a course on science and pseudoscience: A quasi-experimental study. *Teaching of Psychology, 37*, 85–90.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Research Council. (2011). *Promising practices in undergraduate science, technology, engineering, and mathematics education. Summary of two workshop*. Washington, DC: Author.
- National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.
- National Science Board. (2008). *Science and engineering indicators 2008*. Arlington, VA: National Science Foundation (volume 1, NSB 08-01; volume 2, NSB 08-01A).
- Nolan, S. A., & Heinzen, T. E. (2009). Graphing literacy in the psychology major: Florence Nightingale and the creation of a beautiful display of data. *American Psychological Society Observer, 22*, 33–36.
- Noorbala, A., Akhondzadeh, S., Tahmacebi-Pour, N., & Jamshidi, A. (2005). Hydro-alcoholic extract of *Crocus sativus L.* versus fluoxetine in the treatment of mild to moderate depression: A double-blind, randomized pilot trial. *Journal of Ethnopharmacology, 97*, 281–284.
- Thieman, T. J., Clary, E. G., Olson, A. M., Dauner, R. C., & Ring, E. E. (2009). Introducing students to psychological research: General psychology as a laboratory course. *Teaching of Psychology, 36*, 160–168.
- Wade, C. (1995). Using writing to develop and assess critical thinking. *Teaching of Psychology, 22*, 24–28.
- Wilkinson, T., J, Dik, B. J., & Tix, A. P. (2008). The critical thinking lab: Developing students skills through practical application. In D. S. Dunn, J. S. Halonen, & R. A. Smith (Eds.), *Teaching critical thinking in psychology* (pp. 263–265). Malden, MA: Wiley-Blackwell.
- Willingham, D. (2007). Critical thinking: Why is it so hard to teach? *American Educator, 31*, 9–19.