Teaching undergraduate students to read empirical articles: An evaluation and revision of the QALMRI method

Nicholaus P. Brosowsky[1], Olga Parshina[1], Anthony Locicero[2], & Matthew J.C. Crump[2]

[1] Department of Psychology and Neuroscience, Duke University
[2] Department of Psychology, The Graduate Center of the City University of New York
[3] Department of Psychology, Brooklyn College of the City University of New York

Preprint submitted for publication.

Reading and comprehending empirical articles are important skills for students to develop, yet many students struggle to identify and connect the essential information from empirical articles. Here we describe and evaluate a scaffolded approach for teaching undergradauate students to read empirical articles called the QALMRI method. The QALMRI is an generalizable instructional tool for teaching students to identify the key conceptual information necessary for the comprehension and critical evaluation of empirical articles. We had students in a first-year, introductory course and students in a third-year research methods course read empirical articles and complete QALMRI outlines throughout the semester. We found that students very quickly learned to use the QALMRI outline in both upper and lower course levels, with performance corresponding to traditional written summaries. However, we also found that students consistently performed poorly on some items, prompting an update and revision to the QALMRI method to address these limitations.

Keywords:

One important goal for any undergraduate psychology program is to teach students how to be critical consumers of research. The American Psychological Associatin (APA) guidelines, for instance, lists scientific literacy as its second major learning goal (Halonen et al., 2013) and, when surveyed, life sciences faculty rated skills like interpreting data, writing reports, and critically analyzing research articles among the most important for undergraduate students to learn (Coil, Wenderoth, Cunningham, & Dirks, 2010). As such, it is unsurprising that many psychology instructors incorporate primary source readings into their curriculum. By one estimate, over 70% of liberal arts college instructors assigned primary source readings into their courses. And although more prevalent in upper level courses, 46% of introductory psychology instructors still assigned primary source readings (Oldenburg, 2005).

Often, however, primary source readings are not introduced into the curriculum to teach students how to read primary sources. Instead, primary source readings, like empirical articles, are incorporated into the curriculum as a tool to help students learn related course content by way of demonstration and/or elaboration. For example, primary source reading has been used to demonstrate basic statistical and methodological principles (Christopher & Walter, 2006; Pennington, 1992; Ware, Badura, & Davis, 2002), demonstrate writing strategies (Price, 1990; Ware et al., 2002), to spark classroom discussions about course content (Suter & Frank, 1986), promote critical thinking (Chamberlain & Burrough, 1985), help motivate student engagement with course content (Carkenord, 1994), and as a tool to assess student learning and performance (Bachiochi et al., 2011). Although these learning goals do not explicitly include reading and comprehension, it is of course necessary and begs the question: how do students learn to read empirical articles in the first place?

If empirical articles are to be a useful pedagogical tool for learning other course content, students must have the basic skills required to read those articles. Moreover, becoming a critical consumer of the scientific literature is an important step in transitioning from post-secondary education into the scientific community–in and of itself an important goal. Yet there has been considerably less focus on teaching students how to read and critically evaluate empirical articles (e.g., Kershaw, Lippman, & Fugate, 2018; Sego & Stuart, 2016). In the current study, we evaluated a scaffolded approach for

Correspondence concerning this article should be addressed to Nicholaus P. Brosowsky, Duke University, 417 Chapel Dr, Durham, NC 27708, USA. E-mail: nicholaus.brosowsky@duke.edu

Table 1: An outline of the QALMRI method

Questions	What was the broad question being asked by this research project? What was the specific question being asked by this research project?
Alternatives	What was the author's hypothesis? What were the alternative hypotheses?
Logic	What was the logic of the hypothesis? i.e., if the hypothesis was true, what should we expect to happen?
Methods	What were the methods?
Results	What were the important results?
Inferences	What inferences about the questions and hypotheses can be made base on the results?

Adapted from Kosslyn and Rosenberg (2003); See also Brosowsky and Parshina (2017) for a more detailed description and additional instructional materials

reading and summarizing empirical articles known as the "QALMRI" method in both lower- (first-year introductory course; study 1) and upper-level (third-year methods course; study 2) psychology courses.

Reading empirical articles

As compared to reading materials commonly encountered by undergraduate students, empirical articles can be particularly challenging. Whereas textbooks are written primarily with students in mind, empirical articles are written for other experts. Textbooks, for example, present ideas and concepts as a series of facts to be learned and memorized, whereas empirical articles contain persuasive arguments (Gillen, 2006; Suppe, 1998), using data and references to convince the reader of some claim (Latour, 1987; Van Lacum, Ossevoort, & Goedhart, 2014). Whereas textbooks will often provide a great deal of guidance, exploiting text structures and signalling techniques to highlight and point readers to important information (Kardash & Noel, 2000; Sanchez, Lorch, & Lorch, 2001; Varnhagen & Digdon, 2002), empirical articles contain explanatory and argumentative text structures inherently challenging for non-experts (e.g., Britt, Richter, & Rouet, 2014).

Because students are engaging with materials intended for an expert audience, they often find common assignments like summarizing an empirical article (e.g., Anisfeld, 1987; Gillen, 2006; Karcher, 2001; Levine, 2001; Suter & Frank, 1986), to be quite challenging (e.g., Day, 1986; Taylor, 1983). Students struggle identifying and understanding key components of empirical articles. For instance, they can have difficulty understanding the motivation and argument structure presented in an Introduction (Newell, Beach, Smith, & VanDerHeide, 2011; Van Lacum et al., 2014; Van Lacum, Ossevoort, Buikema, & Goedhart, 2012), the concepts of experimental design in the methods (Zieffler & Garfield, 2009), and the statistical concepts found in the results (Dasgupta, Anderson, & Pelaez, 2014).

In short, reading empirical articles poses a unique challenge for undergraduate students. Two aspects relevant to the current study are worth highlighting: First, students must be able to find and identify important conceptual information that may or may not be explicitly stated. For example, the author may or may not explicitly state their hypotheses, yet the reader should be able to discern the author's hypothesis from the introduction of the study. Second, students must be able to understand the logical connections between important ideas, both within the context of the individual research article and in the broader context of experimental design and the scientific literature. Students should, for example, be able to identify the author's research question, the logic of how the experimental design can address that question, and how the author's conclusions relate to both.

Scaffolded learning, drawing on the metaphor of scaffolding in building construction, describes how students can perform more complex tasks than they would otherwise be capable with the help and guidance of someone more knowledgeable (Wood, Bruner, & Ross, 1976). Scaffolded approaches can make learning more tractable for students by changing the task to be more accessible and better aligned with the students current abilities (Quintana et al., 2004; Rogoff, 1990; Vygotsky, 1978). One important feature is that it provides support and instruction about both how to do the task well and also why it should be done that way (Hmelo-Silver, 2006; Hmelo-Silver, Duncan, & Chinn, 2007). In the context of reading empirical articles, scaffolded approaches should aim to reduce the complexity of the specific task given at the time (e.g., reading a specific empirical article to learn the course content) while, at the same time, helping students develop more general reading, comprehension, and critical evaluation skills.

Two scaffolding approaches for reading primary sources have typically been emphasized (e.g., Kershaw, Lippman, Fugate, 2018). The first approach is to adapt the text itself to make it more appropriate and approachable for students (e.g., Yarden, 2009). For instance, one could simplify an article by removing technical jargon and unnecessary information or add text structures and signalling techniques more common in textbooks like headers, highlighting, ancillary sidenotes, etc. The basic idea is to adapt the article to be more easily understood by the student while at the same time maintaining the article's basic argument structure and content. This approach has primarily been used in biology and the life sciences and has been shown to provide some benefit for critical reading (Baram-Tsabari & Yarden, 2005). A possible downside to this approach is that instructional materials are highly specific, each tailored to a specific article. This could make it quite labor-intensive for instructors to create materials–particularly if different articles require different adaptation strategies. Moreover, although it is clear adapted materials should improve comprehension of the adapated articles, it is not clear if adapted materials help student develop more general strategies for tackling empirical articles.

The second approach is to design instructional materials to scaffold reading comprehension without altering the original article (Reiser, 2004; Varnhagen & Digdon, 2002). Supplemental materials can give students the opportunity to engage with complex materials that would otherwise be beyond their current abilities. For example, supplementing readings with organizational signaling materials like overviews (Lorch and Lorch, 1996), outlines or graphical organizers (Easterday, Aleven, & Scheines, 2007; Nussbaum & Schraw, 2007; see, also Newell et al., 2011; Scheuer, Loll, Pinkwart, & McLaren, 2010) have been shown to improve comprehension. Instructional materials can include prompts, directing students' attention toward key features of a text (Hmelo-Silver et al., 2007) or organizational outlines to teach students explicitly about the structure in the text.

This approach is flexible, in that it vary in terms of its specificity. Some instructors have created more general instructional tools, like Van Lacum et al. (2014), who provide an outline for identifying the rhetorical moves used to present an argument in the Introduction of a scientific article. Likewise, Sego and Stuart (2016) created a set of 22 general, open-ended questions that could apply to a number of empirical articles (e.g., "What was/were the independent variables?"). Bachiochi et al. (2011), in contrast, provide a more targeted approach tailoring a set of questions that apply only to the given article. Using these kinds of materials students tend to improve in their ability to dissect empirical articles (Bachiochi et al., 2011; Sego & Stuart, 2016)).

In the current study we evaluated one such scaffolded approach known as the "QALMRI" method. The QALMRI is an instructional tool for teaching students to identify the key conceptual information necessary for the comprehension and critical evaluation of empirical articles (Brosowsky & Parshina, 2017; Kosslyn & Rosenberg, 2003). "QALMRI" is an acronym that stands for Question, Alternatives, Method, Results, and Inferences (Table 1; see Brosowsky and Parshina, 2017 for longer description and student instructional materials). It is a loose framework to help students identify and draw connections between the research questions being asked, how the researcher tried to answer them, and the implications of the answer. The QALMRI methods differ from some of the earlier described approaches in that it takes a broader perspective, asking students to examine the bigger picture. That is, rather than asking students to list facts about an article (e.g., "who was the author?", "how many participants were in the study?"), it asks students to think about the empirical article in terms of its broader goals (i.e., asking and answering questions). Additionally, the QALMRI method is meant to be a simple outline that is easily remembered and can be applied more broadly to any empirical article.

The aim of the current study was to evaluate the use of the QALMRI method in undergraduate psychology courses. In study 1, the QALMRI method was introduced and used in an introductory psychology course. In study 2, the QALMRI method was introduced and used in two third-year experimental psychology methods courses.

Study 1

Methods

Participants

Participants were 28 students enrolled in the Introductory Psychology course at the College of Staten Island of the City University of New York. Participants completed the QALMRI assignments as part of their course requirements.

Materials

Participants completed QALMRI and written summaries of five empirical articles throughout the semester. Articles were chosen on the basis of their topics, such that the article content was relevant to the current course and textbook material. To align the technical level of the articles with the reading level of the first-year students, articles were selected from the Psi Chi Journal of Undergraduate Research (e.g., Sego and Stuart, 2016).

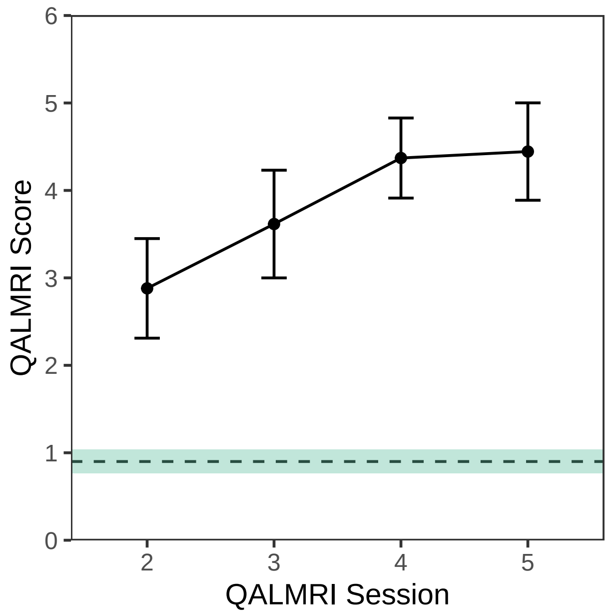

Figure 1. Results from Study 1 showing participant performance on the QALMRI summaries. The dashed line shows performance on the first QALMRI session (prior to the instruction intervention) and error bars represent 95% confidence intervals around the mean.

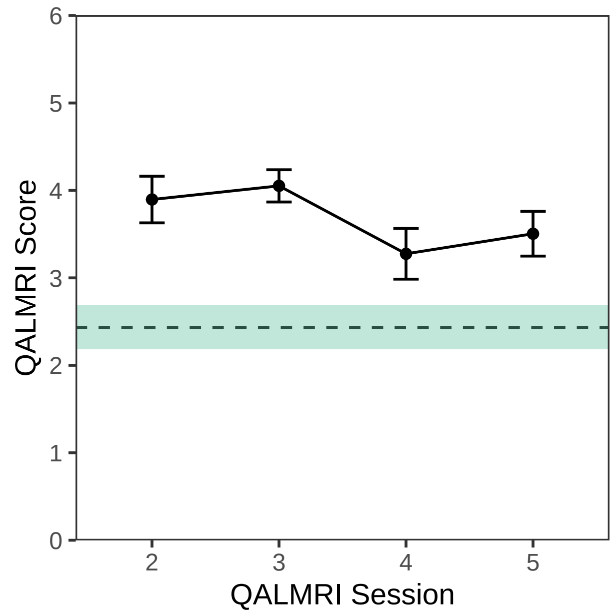

Figure 2. Results from Study 2 showing participant performance on the QALMRI summaries. The dashed line shows performance on the first QALMRI session (prior to the instruction intervention) and error bars represent 95% confidence intervals around the mean.

Procedure

Throughout the semester, students completed five QALMRI sessions. During each session students would read an empirical article and complete a QALMRI summary in-class. The instructor would then review the article and QALMRI content providing additional instruction and demonstration on how to find the relevant information in the article. Crtically, students did not receive any prior instructions on how to use the QALMRI before the first session. They were simply given the article and QALMRI outline and told to read the article and answer the questions. The first session therefore serves as a baseline (pre-intervention) to compare subsequent performance.

QALMRI responses were evaluated per section on a pass/fail basis. That is, students received a pass (for the "Alternatives" section, for instance) if they provided a response that resembled answer key (see Appendix B, for an example of an answer key) and received a score out of a possible 6. QALMRI responses were scored by two independent raters and any conflict reconciled between the raters. Given students could provide a range possible (and possibly correct) responses to the "broad question", we only assess responses for the "specific question".

Results and Discussion

We were primarily interested understanding how well students in an introductory course would learn to use the QALMRI as a method for summarizing empirical articles. To that end, we compared performance on the QALMRIs prior to any instruction (pre-intervention) to the subsequent four QALMRIs. Participants who failed to complete at least 3 out of the 5 QALMRIs were excluded, which, in this case, did not exclude any participants. We used a linear mixed effects model with QALMRI session as a fixed effect and participant as the random effect (linear mixed models use maximum liklihood estimation and is capable of handling missing data). The resulting analyses showed that participants scored significantly higher on the QALMRI after receiving the instruction intervention on each of the following sessions (see Figure 1): Session 2, $t(100.59) = 6.27$, $p < .001$; Session 3, $t(100.97) = 8.73$, $p < .001$; Session 4, $t(100.56) = 11.28$, $p < .001$; and Session 5, $t(100.56) = 11.51$, $p < .001$.

Study 2

Methods

Participants

Participants were students enrolled in third-year experimental psychology methods courses at the College of Staten Island of the City University of New York (N = 38) and Brooklyn

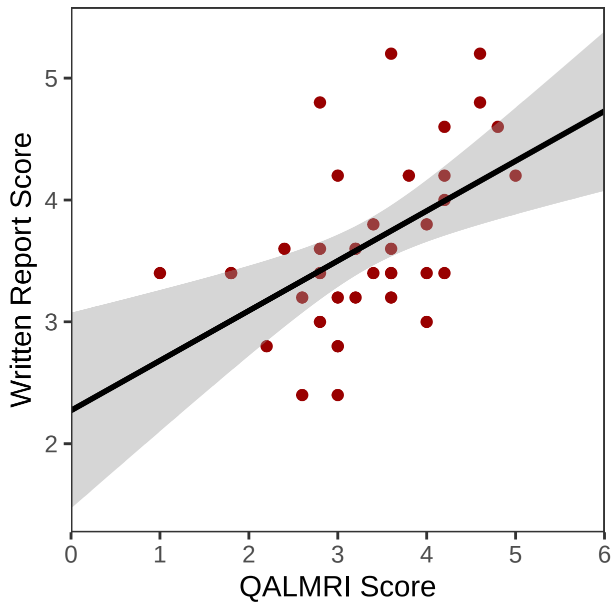

Figure 3. Results from Study 2 showing participant performance on the QALMRI summaries plotted agains written summary scores, r = .51, 95% CI [.23, .71], $t(36)$ = 3.56, p = .001.

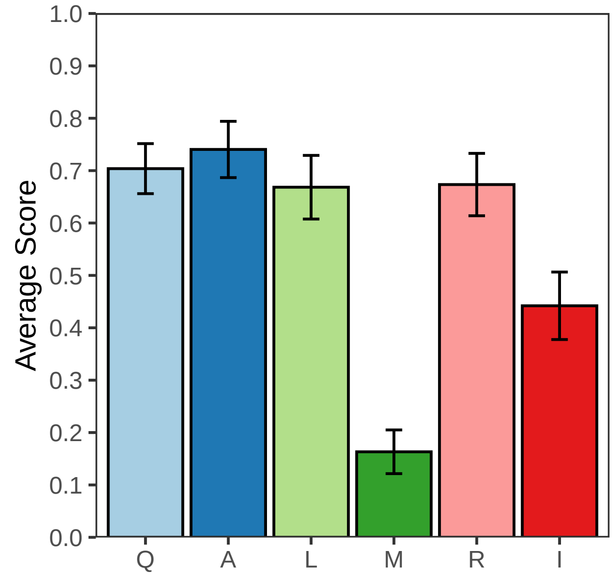

Figure 4. Results from Study 2 showing QALMRI item scores averaged across QALMRI sessions 2 through 5.

College (N = 85). Participants completed QALMRIs as part of their required course assignments.

Materials

Five empirical articles were chosen to align with the course content. In this case, articles were chosen based primarily on their experimental design and increased in difficulty according experimental design (e.g., single factor with two groups, repeated measures designs, two by two factorial designs, etc.). See Appendix A for a complete list of the chosen articles.

Procedure

The procedure is largely the same as Study 1. In addition to the in-class QALMRI summaries, the College of Staten Island students were also required to complete traditional written summaries of the same articles as part of their course assignments.

Results and Discussion

QALMRI performance

To evaluate QALMRI performance we compared scores on the QALMRIs prior to any instruction (pre-intervention) to the subsequent four QALMRIs. Participants who failed to complete at least 3 out of the 5 QALMRIs were excluded, which excluded 10 participants. We used a linear mixed effects model with QALMRI session as a fixed effect and participant and article as random effects. We collapsed over class, as we were not interested in differences between classes, however, we included article as a random effect to take into account the variability in difficulty between articles. The resulting analyses showed that participants scored significantly higher on the QALMRI after receiving the instruction intervention on each of the following sessions (see Figure 1): Session 2, $t(9.39)$ = 3.99 , p = 0.003; Session 3, $t(9.39)$ = 4.92 , p < .001; Session 4, $t(9.78)$ = 2.57 , p = 0.029; and Session 5, $t(9.21)$ = 3.34 , p = 0.008.

QALMRI summaries as compared to traditional written summaries

The College of Staten Island participants also completed traditional written summaries of the same articles. We compared performance between the average written QALMRI summary scores and found a significant positive correlation (see Figure 3), r = .51, 95% CI [.23, .71], $t(36)$ = 3.56, p = .001. This result is perhaps unsurprising in that students who do well in the written summaries also do well on the traditional written summaries. However, it validates the QALMRI method as not only a instructional tool, but as an assessment tool, in place of more traditional summary assignments.

QALMRI performance by item

Finally, we interested in how we could improve and revise the QALMRI method. Therefore, we analyzed performance across QALMRI items to determine where students were having difficulty. To that end, we analyzed QALMRI scores from sessions 2 through 5 using a repeated measures ANOVA with QALMRI item (Q,A,L,M,R, and I) as the repeated measures factor and found a significant difference across items, $F(5, 365) = 90.87$, $MSE = 0.04$, $p < .001$, $\hat{\eta}^2_p = .555$, 90% CI [0.5, 0.59].

General Discussion

In two studies we evaluated student performance using the QALMRI method to summarize empirical articles. In general, we found that students, both in a first-year introductory course and a third-year experimental methods course, learned very quickly how to use the QALMRI after only a single instructional session. This suggests that the QALMRI method is very simple tool that students easily adopt and demonstrates its usefulness across varying educational levels. We also found a strong correspondence between QALMRI performance and performance on a more traditional written summary assignment. This result suggests that QALMRI assignment could subsitute for more traditional forms of scientific literacy assessments.

We also note however, that students did not improve to the same degree across all the QALMRI items. In study 2, we found that students performed much worse, on average, on the methods section, and to a lesser degree, the inferences section. The scores the methods section of the QALMRI particularly stand out: students scored very low across all the QALMRI sessions. In some ways perhaps this is also unsurprising. The methods section contains the most factual information and missing any of the required information could result in a score of 0. So this may be an artifact of how we assessed the QALMRIs. However, it may also speak to the difficulty students have identifying what aspects of the methods section are important. Qualitatively, student responses varied a great. For instance, some students focused heavily on minor details (e.g., number of trials, detailed experimental procedures) and missed important aspects (e.g., identifying the independent variables), while other students presented only vague responses. The inferences section was also challenging for students. Again, qualitatively we found that students had difficulty separating their own conclusions and inferences from those of the authors. This might be the result of the vaguely worder prompt which does not specifically indicate which it is referring to ("What inferences about the hypotheses and questions can be made based on the results?"). A final limitation we found with using the QALMRI method, is its inability to deal with multi-experiment studies. The sections simply do not have enough space and are not organized in a way that allows students to summarize each experiment.

Given these limitations and student performance on the QALMRI, we have revised the QALMRI (see Table 2). First, we have added more specific questions about the methods that still generalize across studies, but provide more guidance for students. Second, we have re-worded the inferences section to clarify that that the students should identify the conclusions the authors are drawing from their results. Finally, we have created an alternative multi-experiment QALMRI which simply bookends the original QALMRI with the broader questions the entire article is asking and the inferences the author's draw from the entirety of the article. Instructors can include as many experiment-specific QALMRIs as needed.

References

Anisfeld, M. (1987). A Course to Develop Competence in Critical Reading of Empirical Research in Psychology. *Teaching of Psychology*, *14*(4), 224–227. https://doi.org/10.1207/s15328023top1404_8

Bachiochi, P., Everton, W., Evans, M., Fugere, M., Escoto, C., Letterman, M., & Leszczynski, J. (2011). Using Empirical Article Analysis to Assess Research Methods Courses. *Teaching of Psychology*, *38*(1), 5–9. https://doi.org/10.1177/0098628310387787

Baram-Tsabari, A., & Yarden, A. (2005). Text genre as a factor in the formation of scientific literacy. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *42*(4), 403–428.

Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist*, *49*(2), 104–122. https://doi.org/10.1080/00461520.2014.916217

Brosowsky, N. P., & Parshina, O. (2017). Using the QALMRI Method to Scaffold Reading of Primary Sources . In A. Schwartz, C. Shane-Simpson, P. J. Brooks, & R. Obeid (Eds.), *How We Teach Now: The GSTA Guide to Student-Centered Teaching* (pp. 312–328). Society for the Teaching of Psychology.

Carkenord, D. M. (1994). Motivating Students to Read Journal Articles. *Teaching of Psychology*, *21*(3), 162–164. https://doi.org/10.1177/009862839402100309

Chamberlain, K., & Burrough, S. (1985). Techniques for Teaching Critical Reading. *Teaching of Psychology*, *12*(4), 213.

Christopher, A. N., & Walter, M. I. (2006). An Assignment to Help Students Learn to Navigate Primary Sources of Information. *Teaching of Psychology*, *33*(1), 37–63. https://doi.org/10.1207/s15328023top3301_9

Coil, D., Wenderoth, M. P., Cunningham, M., & Dirks, C. (2010). Teaching the Process of Science: Faculty Perceptions and an Effective Methodology. *CBE—Life Sciences Education*, *9*(4), 524–535. https://doi.org/10.1187/cbe.10-01-0005

Dasgupta, A. P., Anderson, T. R., & Pelaez, N. (2014). Development and Validation of a Rubric for Diagnosing Students' Experimental Design Knowledge and Difficulties. *CBE—Life Sciences Education*, *13*(2),

Table 2: An outline of the revised QALMRI method (2019)

Questions	What was the broad question being asked by this research project? What was the specific question being asked by this research project?
Alternatives	What was the author's hypothesis? What were the alternative hypotheses?
Logic	What was the logic of the hypothesis? i.e., if the hypothesis was true, what should we expect to happen?
Methods	Briefly describe the study design What were the independent and dependent variables? Briefly describe the study procedure in everyday terms
Results	What were the important results?
Inferences	What did the authors conclude from their study and how do those conclusions relate to the questions and hypotheses?

Adapted from Kosslyn and Rosenberg (2003); See also Brosowsky and Parshina (2017) for a more detailed description and additional instructional materials

265–284. https://doi.org/10.1187/cbe.13-09-0192

Day, J. D. (1986). Teaching Summarization Skills: Influences of Student Ability Level and Strategy Difficulty. *Cognition and Instruction*, *3*(3), 193–210. https://doi.org/10.1207/s1532690xci0303_3

Easterday, M. W., Aleven, V., & Scheines, R. (2007). Tis better to construct than to receive? The effects of diagram tools on causal reasoning. *Frontiers in Artificial Intelligence and Applications*, *158*, 93.

Gillen, C. M. (2006). Criticism and Interpretation: Teaching the Persuasive Aspects of Research Articles. *CBE—Life Sciences Education*, *5*(1), 34–38. https://doi.org/10.1187/cbe.05-08-0101

Halonen, J. S., Buskist, W., Dunn, D. S., Freeman, J., Hill, G. W., Enns, C., & others. (2013). APA Guidelines for the Undergraduate Psychology Major (Version 2.0). *Washington, DC: APA*.

Hmelo-Silver, C. E. (2006). Design principles for scaffolding technology-based inquiry. *Collaborative Learning, Reasoning, and Technology*, 147–170.

Hmelo-Silver, C. E., Duncan, D., Ravit Golan, & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and. *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Karcher, S. J. (2001). Student reviews of scientific literature: Opportunities to improve students' scientific literacy and writing skills. In S. J. Karcher (Ed.), *Tested studies for laboratory teaching* (Vol. 22, pp. 484–487).

Kardash, C. M., & Noel, L. K. (2000). How Organizational Signals, Need for Cognition, and Verbal Ability Affect Text Recall and Recognition. *Contemporary Educational Psychology*, *25*(3), 317–331. https://doi.org/10.1006/ceps.1999.1011

Kershaw, T. C., Lippman, J. P., & Fugate, J. M. B. (2018). Practice makes proficient: Teaching undergraduate students to understand published research. *Instructional Science*, *46*(6), 921–946. https://doi.org/10.1007/s11251-018-9456-2

Kosslyn, S. M., & Rosenberg, R. S. (2003). *Fundamentals of Psychology: The Brain, the Person, the World*. Allyn; Bacon.

Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press.

Levine, E. (2001). Reading your way to scientific literacy. *Journal of College Science Teaching; Washington*, *31*(2), 122–125.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and Learning Argumentative Reading and Writing: A Review of Research. *Reading Research Quarterly*, *46*(3), 273–304. https://doi.org/10.1598/RRQ.46.3.4

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education*, *76*(1), 59–

92.

Oldenburg, C. M. (2005). Use of Primary Source Readings in Psychology Courses at Liberal Arts Colleges. *Teaching of Psychology*, *32*(1), 25–29. https://doi.org/10.1207/s15328023top3201_6

Pennington, H. (1992). Excerpts from Journal Articles as Teaching Devices. *Teaching of Psychology*, *19*(3), 175–177. https://doi.org/10.1207/s15328023top1903_15

Price, D. W. W. (1990). A Model for Reading and Writing about Primary Sources: The Case of Introductory Psychology. *Teaching of Psychology*, *17*(1), 48–53. https://doi.org/10.1207/s15328023top1701_12

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., … Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, *13*(3), 337–386.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, *13*(3), 273–304.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* Oxford university press.

Sanchez, R. P., Lorch, E. P., & Lorch, R. F. (2001). Effects of Headings on Text Processing Strategies. *Contemporary Educational Psychology*, *26*(3), 418–428. https://doi.org/10.1006/ceps.2000.1056

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102.

Sego, S. A., & Stuart, A. E. (2016). Learning to Read Empirical Articles in General Psychology. *Teaching of Psychology*, *43*(1), 38–42. https://doi.org/10.1177/0098628315620875

Suppe, F. (1998). The Structure of a Scientific Paper. *Philosophy of Science*, *65*(3), 381–405. https://doi.org/10.1086/392651

Suter, W. N., & Frank, P. (1986). Using Scholarly Journals in Undergraduate Experimental Methodology Courses. *Teaching of Psychology*, *13*(4), 219.

Taylor, K. K. (1983). Can College Students Summarize? *Journal of Reading*, *26*(6), 524–528.

Van Lacum, E. B., Ossevoort, M. A., & Goedhart, M. J. (2014). A Teaching Strategy with a Focus on Argumentation to Improve Undergraduate Students' Ability to Read Research Articles. *CBE—Life Sciences Education*, *13*(2), 253–264. https://doi.org/10.1187/cbe.13-06-0110

Van Lacum, E., Ossevoort, M., Buikema, H., & Goedhart, M. (2012). First Experiences with Reading Primary Literature by Undergraduate Life Science Students. *International Journal of Science Education*, *34*(12), 1795–1821. https://doi.org/10.1080/09500693.2011.582654

Varnhagen, C. K., & Digdon, N. (2002). Helping students read reports of empirical research. *Teaching of Psychology*, *29*(2), 160–165.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34–41.

Ware, M. E., Badura, A. S., & Davis, S. F. (2002). Using student scholarship to develop student research and writing skills. *Teaching of Psychology*.

Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving*. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yarden, A. (2009). Reading scientific texts: Adapting primary literature for promoting scientific literacy. *Research in Science Education*, *39*(3), 307–311.

Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, *8*(1), 26.

Table 3: A multi-study QALMRI

Questions	What was the broad question being asked by this research project?
Study 1:	
Questions	What was the specific question being asked in this study?
Alternatives	What was the author's hypothesis? What were the alternative hypotheses?
Logic	What was the logic of the hypothesis? i.e., if the hypothesis was true, what should we expect to happen?
Methods	Briefly describe the study design What were the independent and dependent variables? Briefly describe the study procedure in everyday terms
Results	What were the important results?
Inferences	What did the authors conclude from their study and how do those conclusions relate to the questions and hypotheses?
Study 2:	
Questions	What was the specific question being asked in this study?
Alternatives	What was the author's hypothesis? What were the alternative hypotheses?
Logic	What was the logic of the hypothesis? i.e., if the hypothesis was true, what should we expect to happen?
Methods	Briefly describe the study design What were the independent and dependent variables? Briefly describe the study procedure in everyday terms
Results	What were the important results?
Inferences	What did the authors conclude from their study and how do those conclusions relate to the questions and hypotheses?
Inferences	What did the authors conclude from the results of all the studies and how do those conclusions they relate to the broad question?