

METACOGNITIVE SCAFFOLDING TO DEVELOP CRITICAL CONSUMERS OF RESEARCH

Rachelle Kisst Hackett and Linda Bol

University of the Pacific, USA

Old Dominion University, USA

rhackett@pacific.edu

One important student learning objective (SLO) of programs in applied disciplines is developing the ability to inform one's professional practice by reading published research critically, evaluating the validity of claims by unpacking the methodology/statistical analyses employed. Continuing to do so after a course concludes depends, however, on students' perceptions of how relevant research is and self-efficacious they feel as independent, critical consumers of research. Pedagogy designed to address this SLO among graduate-level students in education is presented along with samples of student work and student reactions to such assignments. A key component involves students annotating journal articles with comments containing their evaluations of methodology (strengths / weaknesses / confusing parts) and relevance (personal / professional) in a structured system designed to induce metacognitive reflection.

INTRODUCTION

The *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report* (2016) state, "Students should become critical consumers of statistically-based results reported in popular media, recognizing whether reported results reasonably follow from the study and analysis conducted" (p.8). The authors distinguish between courses that address statistical literacy versus those focused on statistical methods (p.7). The study reported here was conducted in classes where statistical literacy goals prevail. We consider the often-quoted definition of statistical literacy used by Gal (2002), that is, "the ability to interpret, critically evaluate, and communicate about statistical information and messages" (p.1) to relate well to the terms educational researchers (and other social scientists) use when evaluating design validity. In particular, research methods course textbooks (e.g., McMillan & Schumacher, 2010) often list threats to internal, external, construct, and/or statistical conclusion validity that capture many concerns similar to those listed by Gal as "worry questions" (2002, Table 3, p.16). One important objective of programs in applied disciplines (e.g., education) is developing learners' ability to inform their professional practice by reading published research critically. The effectiveness of an intervention designed to address this objective was explored in this mixed-methods study.

INTERVENTION AND SEQUENCE

During the Fall 2017 semester, a set of four quantitative research journal articles was used to teach research methods in both a doctoral- and a master's-level class of education majors; student learning objectives partially overlap. The master's level research methods course does not require any prerequisites and literacy goals dominate, whereas both literacy and methods goals are part of the doctoral coursework. A common feature of both classes was the seven-part instructional sequence used with each article consisting of students 1) reading the research article and annotating the PDF itself with comments about the methodology and relevance, 2) transferring one relevancy and four methodological comments to the "WORD[®] template," 3) submitting their initial comments to the learning management system (LMS), 4) attending and engaging in the class discussion of the article while a small group of their peers (the article leaders) solicits input on that group's draft of the article critique, 5) modifying their methodological comments by revising and/or adding other issues, 6) resubmitting their updated list of comments to the LMS, and 7) reading feedback provided by their instructor as to the quality of their methodological comments.

On the template, students classified each research related comment as being a methodological strength, weakness, or area of uncertainty. Their relevancy comment was labeled according to whether the article has personal or professional relevance for them. After participating in the class discussion and reflecting on the most salient issues for that article, students were allowed to make modifications. Written instructions, a grading rubric, and a visual aid for

completing the template were provided. The grading rubric highlights the importance of the quality of methodology comments, use of research terminology, variety of issues noted, and salience of the set of comments. Students were encouraged to consider various types of validity and reminded that summaries of threats to each type were provided in the textbook required for the course (McMillan & Schumacher, 2010, pp.264-266). For both classes, the final, revised comments (across six articles) counted for 30% of their course grade; thus, this instructional sequence was viewed by both the instructor and students as a crucial and heavily weighted component of the course.

Each student served as a leader for just one article, working within a small group of 2-4 students. In addition to individually completing the "WORD template," the small group jointly drafted a critique of the article prior to its discussion in class. The article critique template emphasized identification of the methodological components (e.g., research question, type of design, sampling, instrumentation, data analysis) as well as a section evaluating the quality of the study. The small group's finalized article critique counted an additional 10% of their course grade. For a more detailed description of the materials and sequence, please contact the authors.

SELF-REGULATED LEARNING (SRL) AND METACOGNITIVE SCAFFOLDING

We employ self-regulation as the theoretical lens for design and data interpretation. "Self-regulated learning refers to self-generated thoughts, feelings, and actions for attaining one's learning goals" (Zimmerman & Moylan, 2009, p. 299). The ability to monitor one's learning is vital to success in all academic endeavors. Students rely on internally and externally generated feedback regarding their academic performance to make judgments about their progress towards their goals. If progress is determined to be inadequate, students must adjust their goals or performance to accomplish their tasks (Zimmerman & Moylan, 2009).

Metacognitive scaffolding can be used to prompt effective self-regulation. Scaffolding is an instructional tool designed to aid the learner by providing supportive assistance. Though there are different types of scaffolds, we focus on metacognitive or reflective scaffolds to support thinking about a task (Azevedo, Cromley, & Seibert, 2004). Guidelines, questions or other explicit reminders are scaffolds often used to encourage strategies aligned with the Forethought (e.g. task analysis, intrinsic interest/ value), Performance (e.g., task strategies, monitoring) and Self-Reflection (e.g., self-evaluation, self-satisfaction) phases of self-regulated learning (Bol & Garner, 2010).

The instructional sequence was designed to induce metacognitive processes through scaffolds such as the "WORD template" for recording initial and revised comments. The article critique prompts identification of various design components and an evaluation of the quality of the research. Student planning for this assignment was facilitated by providing clear task structure, modeling, and grading rubrics. We also attempted to spark interest by requiring students to identify content that was personally or professionally relevant. Monitoring or gauging one's mastery of the content was prompted by categorizing methodological issues as strengths, weaknesses or areas of uncertainty. Making these judgments, particularly those deemed uncertain, helped illuminate the students' own understanding of their mastery of concepts. Student reflection may have been further enhanced by asking students to recognize research design as involving difficult tradeoffs and a series of decisions that are not necessarily equally salient. Requiring students to revisit and possibly revise the comments they make on the "WORD template" after the class discussion promotes further metacognitive reflection. Reflection on one's performance was reinforced by professor feedback and instruction that emphasized the importance of self-evaluation and improvement. According to self-regulated learning, positive reflection and perceived mastery should lead to increased self-satisfaction and confidence levels.

RESEARCH QUESTIONS

1. Are the instructional techniques employed in the quantitative research methods courses helping to create educational practitioners who, upon graduation, plan to read and evaluate published research critically so as to inform their professional practice?
2. In what ways do students demonstrate their understanding of how researchers' methodological choices impact the quality of published studies?

3. What evidence suggests that the mechanisms by which the above occurs are, in part, due to scaffolds that induce metacognitive reflection and self-regulated learning?

METHOD

This study explores how instructional techniques used in teaching quantitative research methods may impact students' use of self-regulated learning strategies and their ability and intention to critically consume research once they graduate and work as professional educational practitioners. Both quantitative and qualitative data are gathered in a mixed-methods design leading to thematic analyses of student descriptions of their learning processes, descriptions of students' submitted work, and statistical analyses of self-reports by students of their perceptions, practices, and plans collected at the conclusion of the intervention. The latter survey data represent the pre-experimental post-test only design portion of the study. University approval for using human subjects was obtained.

Participants

The intervention was used with 39 students enrolled in graduate-level education courses taught by the first author. The master's-level class (n=25) consisted primarily of females (76%). Their program concentrations were mostly educational & school psychology (60%) and curriculum & instruction (20%). The doctoral-level class consisted of 14 students in educational leadership of whom 93% were female. A survey was completed by 37 students (95%) from Caucasian (~54%), African-American (~16%), Hispanic (~14%), and other (~16%) racial/ethnic backgrounds.

Data Collection Instrument and Procedures

To address the first and third research question, the Perceptions, Practices, & Plans Survey (PPPS) was developed for this project by the authors and administered to students the week after both classes had completed the 7-part instructional sequence for each of four articles. Prior to survey administration, the first author read a letter to each of her classes noting they were to respond anonymously, that it was voluntary, and one of two options available for earning extra credit in the course. It was emphasized that there were no wrong or right answers, only truthful responses. The instructor left the room while students completed the surveys and submitted them to an assistant who emailed the scanned surveys to the second author who constructed the database. The first four sections of the PPPS contained 11 closed-ended items asking about student (a) practices before the course, (b) practices during the course involving the cycle of activities, (c) perceptions about the relevance of research and their confidence in being able to read and evaluate empirical studies, (d) practices they plan to engage in after graduating that involve reading research to inform professional practice. The final section consists of four open-ended questions where students are asked to describe the approaches they used involving the cycle of activities; these prompt for self-regulation processes involving forethought (task analysis and self-motivation), performance (self-control), and self-reflection, as described above. To address the second research question, a document analysis was employed. Student comments were categorized by the types of and threats to validity they identified.

RESULTS

Research Question 1: Educational Practitioners Critically Consuming Research

Findings suggested an increase in the extent to which the doctoral and the master's level students attended to both the method and results sections and their intentions to do so in the future. Students were asked how much attention they paid to *Method* sections before taking the course and their plans for doing so after graduating. Options included noting that they (1) do not read research, (2) skip, (3) skim, (4) read, or (5) read and evaluate the section. Only one student of 37 (3%) both read and evaluated the Method sections prior to the course, whereas 16 (43%) planned to do so after graduating. Similarly, students were asked how much attention they paid to the *Results* sections (where statistical information is reported). Seven of 37 (19%) students both read and evaluated the Results sections prior to the course, whereas 19 (51%) planned to do so after graduating. Similar patterns were found when analyses were based on each course separately.

Results revealed that most students perceived the relevancy of research. Thirty-three (89%) students either somewhat or strongly agreed with the statement that, “Educational and/or psychological research is *relevant* for informing my professional practice.” When asked, “What did you find valuable about identifying the methodological strengths and weaknesses?” students spoke about thinking critically, e.g., “It made me a more critical thinker. This class changed the way I view research and studies. I would not want to major in research methods but I am glad I took this class and I will benefit from in my future learning and career.” Others commented on evaluating research quality, e.g., “It helped me have a better understanding of what makes a study valid or reliable. As a person who uses research to support my practice I think this is important.”

Research Question 2: Student Understanding of Methodology Impacting Research Quality

A brief description of three assigned articles is offered to help the reader interpret the results. Article A (Taheri, Perry, & Minnes, 2016) is a non-experimental, comparative study involving social participation and friendships of three groups of children: typically developing (TD), children with an intellectual disability (ID only), and those with autism spectrum disorder plus intellectual disability (ID+ASD). The authors of Article B (Shoulders & Krei, 2015) classify their study as causal-comparative, in which they attempt to determine the cause for differences in teacher self-efficacy by comparing subgroups based on a single factor at a time (such as level of education or years of teaching experience). The first research question stated in Article C (Scott, Hirn, & Alter, 2014) includes both a descriptive part, “To what degree are teachers engaging in instruction?” and correlational part, “What is the relation between teaching and student behaviors of engagement and disruption?” Please consult the articles (referenced below) for more contextual information, if needed.

Table 1 presents methodological comments made by students regarding internal, external, construct, and statistical conclusion validity. Their comments suggest an emerging grasp of potential strengths and weaknesses related to statistical literacy in the context of research articles.

Research Question 3: Self-Regulated Learning and Metacognitive Scaffolds

Data addressing the final research question consisted primarily of students’ written responses to open-ended questions that were aligned with the phases of SRL. See Table 2 for a sample of student comments. The most frequent themes and student quotes representative of these themes were identified within each phase. Students were asked to describe their approach to completing the task. How they planned for and began the task was meant to reflect the metacognitive strategy of task analysis in the Forethought phase. The responses to the approach adopted were primarily sequential in nature. That is, students analyzed the task by noting the steps taken. Often they began with a cursory read through the articles making notations. Some read particular sections of the article in a particular order. They then moved to comparing initial comments with text, lecture, and discussion content to inform their revisions.

To tap metacognitive strategies in the Performance phase, students were asked how they monitored and checked their progress between writing the initial and final set of comments. By far, the most common theme was the use of class discussion to help them better understand the strengths and weaknesses in studies and revise their initial comments. Others compared and contrasted their comments with content provided in the text and notes. They relied on careful note-taking during the discussions to improve their initial comments. Returning to review the rubric and instructions was yet another strategy employed by these students. Their performance frequently occurred in “real-time” as they made revisions during class discussion.

Aligned with the Self-Reflection phase, students were asked to describe how satisfied they were with the quality of their final set of comments. Students relied on formal feedback from the professor as well as explanations and clarifications during class discussions to help them evaluate their performance. One theme found in some student comments was a continued lack of confidence in their understanding of research and statistics, as might be expected midway through the course. Others noted improvement in their work over time which led to increased confidence in their mastery of the content.

Three rating scale items also assess metacognitive skills at the Forethought, Performance, and Self-Reflection phases of Zimmerman’s SRL model. A total of 23 students (62%) agreed that,

“The use of the ‘Grading Rubric for Article Comments’ helped me set goals for the quality of my work.” Twenty-six (70%) of the students either somewhat or strongly agreed with the statement that, “The use of the ‘WORD Template’ helped me to organize my thoughts about the quality of the research reported on in the article,” and 28 (76%) agreed that it, “made me reflect on the extent to which I was focusing on a variety of issues that may affect the quality of research.”

Table 1. Student comments classified by types of validity and specific threats.

Internal Validity Weakness: <i>Selection Threat</i>	The study does not address gender differences in regard to friendship. Children are generally socialized in different ways depending on their gender, which may lead to differences in quality of friendship or the behaviors of the child that we interpret. The ASD group contains a significantly larger amount of males than the other groups, which means the differences between groups could be due to gender rather than diagnosis. (<i>Article A</i>)
Internal Validity Weakness: <i>Selection Threat</i>	There were some discrepancies when the researchers looked at the level of education of their participants. The data was not analyzed in a fruitful way and instead left room for interpretation. Were the results collected because of the level of education of the participant or simply their teaching experience? Were the participants with lower educational levels newer to the practice or had they simply decided not to seek out higher degrees? (<i>Article B</i>)
External Validity Strength: <i>Selection of Subjects</i>	Researchers took extra precautions to ensure that the population they were studying was accurate. This is a strength because we can see that those who were identified as having DD (combined or single) were appropriately categorized and those who were identified as being TD were appropriately categorized. (<i>Article A</i>)
External Validity Weakness: <i>Novelty or Disruption Effect</i>	There is a problem with ecological external validity in this study. The study specifically stated, “sense of mastery could have been undermined by implementation of the Common Core and other new state standards in both locations”. The context might be affecting the results due to novelty effect by the implementation of Common Core. Therefore, the study can only be generalized to similar conditions. (<i>Article B</i>)
Construct Validity Weakness: <i>Inadequate Explication of Constructs</i>	A weakness of the study is the concerns with construct underrepresentation validity because the researchers are underreporting the amount of total teacher engagement time and teacher teaching time when the observers activated the off-task code when the teacher was not teaching a small group the target student was a part of. This is counterproductive to the first part of research question 1 which is asking “to what degree are teachers engaging in instruction?” (<i>Article C</i>)
Statistical Conclusion Validity Strength: <i>(Un) Reliability of Measures</i>	Table 1 shows the Interobserver Reliability coefficient of each behavior that has been coded. Every coded behavior has a reliability coefficient of over .8, which means that each coded behavior can be considered relatively reliable. This means that different raters, that have been extensively trained, gave consistent estimates of the same measure. It is important to have adequate reliability so that we do not underestimate the relationship of variables. (<i>Article C</i>)

DISCUSSION AND CONCLUSION

This exploratory study provides preliminary evidence to suggest that the seven-part instructional sequence can induce self-regulated learning and metacognitive reflection among graduate students enrolled in educational research methods courses. Comments suggest students are learning to consider how researchers’ methodological choices impact research quality and the validity of data-based claims. Moreover, meeting the longer-term course objective, of developing educators who plan to inform their professional practice by reading published research critically, appears promising. Space precludes an exhaustive discussion of the study, yet there are clear

limitations, most notably asking students to retrospectively and prospectively judge their behaviors. While not ideal in terms of instrument validity, asking students to report their prior behaviors was our only means available to get a sense of the intervention's impact given time, resource, and IRB approval constraints. Therefore, longitudinal designs are an obvious direction for future research. The course practice of identifying relevance when reading research coupled with focused attention on threats to validity may lead students to apply these techniques beyond the classroom and the workplace, given the broad applicability statistical literacy skills have for informing personal, as well as professional, decisions.

Table 2. Themes in student comments by phase of self-regulated learning.

Forethought: <i>Making notations in article</i>	I would read the article quickly, leaving post- its where I thought information stood out as especially interesting. After finishing the article I would return to classify the information as a strength or weakness or uncertainty.
Forethought: <i>Comparing to text and class notes</i>	Identified strengths in the articles by cross referencing course content, lecture, text, etc. Additionally, made informal inferences based on the rational of the study. Basically, apply course content to review these studies.
Performance: <i>Editing and comparing original comments</i>	When we go over the critique in class I also have my initial comments on my computer so I can edit them if necessary. Also the group's handout was useful for finalizing my comments.
Performance: <i>Taking notes for revisions</i>	I make sure that I take diligent notes during the discussion. I also make sure that each comment contains and explains targeted vocabulary. When checking my comments if they do not contain a key work, I revisit and revise.
Self-Reflection: <i>Lacking confidence</i>	I was somewhat satisfied. I still feel like I don't have a strong understanding of what it takes to properly evaluate quantitative statistics, and I am a little disappointed because of that. However I do feel as though I learned a lot.
Self-Reflection: <i>Improvement over time</i>	I am reasonably satisfied with them. They're certainly an improvement over what I started with. Once again it is my noviceness (<i>sic</i>) which still makes me tentative about some of my revision answers.

REFERENCES

- Azevedo, R., Cromley, J. G. & Seibert, D. (2004) Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29, 344–370.
- Bol, L., & Garner, J. (2011). Challenges in supporting self-regulation in distance education environments. *Journal of Computing in Higher Education*, 23, 104-123.
- GAISE College Report ASA Revision Committee, "Guidelines for Assessment and Instruction in Statistics Education College Report 2016," <http://www.amstat.org/education/gaise>.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1) 1-25.
- McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Boston: Pearson.
- Scott, T. M., Hirn, R. G., & Alter, P. J. (2014). Teacher instruction as a predictor for student engagement and disruptive behaviors. *Preventing School Failure*, 58(4), 193-200.
- Shoulders, T. L., & Krei, M. S. (2015). Rural high school teachers' self-efficacy in student engagement, instructional strategies, and classroom management. *American Secondary Education*, 44(1), 50-61.
- Taheri, A., Perry, A., & Minnes, P. (2016). Examining the social participation of children and adolescents with Intellectual Disabilities and Autism Spectrum Disorder in relation to peers. *Journal of Intellectual Disability Research*, 60(5), 435-443.
- Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299-315). New York, NY: Routledge.