INSTRUCTIONAL DESIGN AND ASSESSMENT

Instructional Scaffolding to Improve Students’ Skills in Evaluating Clinical Literature

Stefani Dawn, MA, MS, Karen D. Dominguez, PharmD, William G. Troutman, PharmD, Rucha Bond, PharmD, and Catherine Cone, PharmD

University of New Mexico

Submitted November 22, 2010; accepted February 8, 2011; published May 10, 2011.

Objective. To implement and assess the effectiveness of an activity to teach pharmacy students to critically evaluate clinical literature using instructional scaffolding and a Clinical Trial Evaluation Rubric.

Design. The literature evaluation activity centered on a single clinical research article and involved individual, small group, and large group instruction, with carefully structured, evidence-based scaffolds and support materials centered around 3 educational themes: (1) the reader’s awareness of text organization, (2) contextual/background information and vocabulary, and (3) questioning, prompting, and self-monitoring (metacognition).

Assessment. Students initially read the article, scored it using the rubric, and wrote an evaluation. Students then worked individually using a worksheet to identify and define 4 to 5 vocabulary/concept knowledge gaps. They then worked in small groups and as a class to further improve their skills. Finally, they assessed the same article using the rubric and writing a second evaluation. Students’ rubric scores for the article decreased significantly from a mean pre-activity score of 76.7% to a post-activity score of 61.7%, indicating that their skills in identifying weaknesses in the article’s study design had improved.

Conclusion. Use of instructional scaffolding in the form of vocabulary supports and the Clinical Trial Evaluation Rubric improved students’ ability to critically evaluate a clinical study compared to lecture-based coursework alone.

Keywords: evidence-based medicine, literature evaluation rubric, literature evaluation, scaffolds in instruction, instructional model, clinical trial evaluation

INTRODUCTION

Critical analysis of the medical and scientific literature is one element of evidence-based medical practice and is necessary for today’s pharmacists to help physicians and patients use medications safely and effectively. The Accreditation Council for Pharmacy Education recognizes this need for evidence-based pharmaceutical care and includes requirements for training in evidence-based medicine in Standards 2007.

Pharmacy and medical education publications have presented various approaches for training students to read and critically evaluate the medical literature. The most common instructional support tools used to teach students to evaluate medical literature are standalone questions and simple checklists. Although standalone questions and checklists provide some structure for students when learning to evaluate medical literature, there are weaknesses to this approach. Students must possess either extensive background knowledge or depend on faculty members or experts to assist them in answering the questions. Also, standalone questions and checklists may not teach students how to judge the quality of a study or its associated publication. Having faculty members/experts present to support student learning is important and arguably they are irreplaceable; however, use of faculty expertise should be strategic and combined with explicit and targeted instructional support to build solid, transferable skills in students, particularly in complex learning situations.

Reading and evaluating medical literature is a complex process because of the layered knowledge required. Not only is domain-specific knowledge about the content in the article needed (eg, the pharmacokinetics of a specific drug or the anatomy and physiology of a body system), the reader also must have general knowledge of the scientific process and specific knowledge about the particular type of research (eg, clinical trial, cohort, case-control study). When learning complex, domain-specific science skills and knowledge, students need instructional guidance that promotes an integrated conceptual framework.
Targeted instructional techniques such as scaffolding can be used to help students develop conceptual frameworks. Scaffolding is “the systematic sequencing of prompted content, materials, tasks, and teacher and peer support to optimize” independent learning,\textsuperscript{19} and can include instructional elements such as guided questioning; comparing ideas; identifying connections and distinguishing characteristics between concepts; and identifying valid relationships. When used in such complex, knowledge-based learning situations, scaffolding is more effective than open-ended inquiry-based science instruction and more effective than traditional lecture-based instruction.\textsuperscript{14,17-20}

The purpose of this research was to develop and implement an instructional model using targeted scaffolding and an associated activity to teach students to critically evaluate clinical literature. The hypothesis was that the use of targeted instructional scaffolding using highly interactive, student-oriented, and cognition-based approaches would improve student abilities to evaluate clinical literature more critically than traditional lecture-based instruction.

**DESIGN**

The goal for this research was to develop and test a model and an associated activity to teach students how to read and evaluate clinical literature. This included testing the hypothesis that the use of targeted instructional scaffolding using highly interactive, student-oriented, and cognition-based approaches would improve student ability to critically evaluate clinical literature compared to more traditional lecture-based didactic instruction. This hypothesis was tested by implementing the clinical literature evaluation activity in the Pharmaceutical Care Laboratory near the end of the semester after students had completed the majority (>80%) of material and assignments for the Pharmacy Informatics and Research course, including writing a major literature review paper. The research was approved by the University of New Mexico Human Research Review Committee prior to study implementation.

The University of New Mexico College of Pharmacy is located in Albuquerque and has a 4-year doctor of pharmacy (PharmD) program with a diverse student population (47% white, 32% Hispanic, 12% Asian, 6% Native American/Alaskan native, and 3% black). This research was conducted in fall 2007 with second-year pharmacy (P2) students (class of 2009). The clinical literature evaluation activity was implemented in the second-year Pharmaceutical Care Laboratory because it provided large blocks of time for implementation. Also, the course is taught concurrently with the lecture-based Pharmacy Informatics and Research course, which the authors used as a comparator course.

Activity implementation and faculty observation occurred in 2 different physical settings: a fixed-seat lecture hall, where all 84 of the P2 students attended the Pharmaceutical Care Laboratory lecture simultaneously; and a computer laboratory that could accommodate 20 to 23 students per 3-hour laboratory session. The majority of the activity and observations occurred in the smaller laboratory setting.

Students were informed of the study 2 weeks prior to the activity and were asked to complete consent forms. Students could participate if they allowed their work on the blinded activity to be evaluated by the study investigators and/or agreed to serve as a case study student. In order to become part of the case study, students had to consent to have their work undergo product analysis.

For the scaffolding activity, students read and evaluated a single clinical research article on the effects of red yeast rice on lowering cholesterol. The activity knowledge content consisted of 2 elements: the layered-learning around how to evaluate a clinical study (eg, clinical study design, common strengths and weaknesses in clinical studies, and applying the knowledge to the study they read), and the content knowledge regarding the effects of red yeast rice on cholesterol. The topic of cholesterol and red yeast rice was selected to expand students’ knowledge in self-care therapeutics and support concurrent coursework; however, the primary intention of the activity was to develop students’ knowledge and skills in critically evaluating a clinical study, which included developing an awareness of their own knowledge and skill limitations.

Although the activity and model were highly structured (a common characteristic of instructional scaffolding), they were heavily learner-centered. The instructor primarily served as a facilitator and provided expertise as needed if information was missing from the activity materials. The model and activity used individual, small group, and large group instruction, with carefully structured, evidence-based scaffolds and support materials centered around 3 educational themes important to support science reading comprehension and analysis: (1) the reader’s awareness of text organization,\textsuperscript{13,21-25} (2) contextual/background information and vocabulary,\textsuperscript{13,22,24} and (3) questioning, prompting, and self-monitoring (metacognition).

**Text Organization.** Students were oriented to the text organization of clinical research articles (introduction, methods, results, discussion, conclusion) through use of the Clinical Trial Evaluation Rubric developed by the study authors and based upon the CONSORT Statement.\textsuperscript{23} Additionally, the rubric provided extensive details to familiarize students with the type of information that should be included in each section of the study.\textsuperscript{23,24}

**Contextual Background Information and Vocabulary.** Scientific and medical literature typically has advanced vocabulary, often without definitions, which can hinder reading comprehension,\textsuperscript{13,25} especially if a student’s
vocabulary is limited.\textsuperscript{26,27} The instructional model and activity had extensive vocabulary-related scaffolds in the form of individual, small group (3-4 students), and large group (20-23 students) activities related to the Vocabulary and Concepts Worksheets and the rubric.

Prompting and Self-Monitoring (Metacognition). The vocabulary scaffolds were intended not only to expand students’ vocabulary but also to identify their knowledge gaps (metacognition). Instructional scaffolds that develop metacognition improve science learning and science reading comprehension.\textsuperscript{28,29}

Student and instructor materials used in the activity are listed in Table 1. Each step of the activity and its associated educational outcomes is described:

1. Students independently read and evaluated a clinical research article using the Clinical Trial Evaluation Rubric. Students wrote an initial summary of their evaluation of the article justifying the initial scores they gave.
2. Students individually revisited the article and identified and defined 4 to 5 vocabulary/concept knowledge gaps and placed them in a blank Vocabulary and Concepts Worksheet, which contained guiding questions.
3. Students worked in groups of 3 where they helped each other further define and clarify the self-identified terms and concepts.
4. Instructors led a large-group discussion (20 to 23 students per laboratory session) centered on vocabulary and concepts to address any gaps in student definitions. The instructors asked each of the small groups to share 1 or 2 vocabulary words or concepts with the larger group, along with the definition or information they had obtained. Students from the larger group were invited to add information or ask clarification questions. The 2 instructors who implemented the activity used the same Article Vocabulary and Concepts Worksheet key to ensure that important concepts and vocabulary were covered and similar information was given to the classes to reduce variability. The revised version of this step is described in the Evaluation and Assessment section. (The blank and completed Article Vocabulary and Concepts Worksheet are available from the author.)
5. Following the large-group discussion, students individually reevaluated the article using the rubric. Students then wrote a second summary explaining any changes in the rubric scores they assigned to the clinical article.

The refined instructional model based upon this research is presented in Figure 1.

### EVALUATION AND ASSESSMENT

Of the 84 second-year students, 51 (61\% of the class) consented to participate in the blinded phase of the study; 30 also consented to be included in the unblinded case study phase for which 6 students were selected. Case study students were selected to ensure 1 group within each PCL instructor’s laboratory (R.B. and C.C.) Students who chose not to participate in the study were still required to complete the activity because it counted as part of the course grade; however, participation in the study did not affect students’ grades.

Each student had been assigned a number to blind the investigators who conducted the work product analysis and still allow for future matching of the survey and multiple work products. Data sources included completed pre- and post-activity Clinical Trial Evaluation Rubric
scores/assessments, pre- and post-activity written summaries, in-class observations, and case studies with interviews.

Students read and scored the clinical article using the Clinical Trial Evaluation Rubric before and after the activity. The pre-activity article rubric score and evaluation summary were used as comparative measures to characterize the students’ knowledge base following their completion of the majority of the Pharmacy Informatics and Research course (the comparator course). The post-activity rubric score and summary content indicated changes in students’ critical evaluation of the clinical literature. Average pre- and post-activity rubric scores were compared using the Wilcoxon matched-pairs signed-rank test with a Bonferroni correction for multiple analyses. Statistical analyses were performed using SPSS.

Written summaries from case-study students were analyzed for changes before and after the activity with respect to depth of analysis by counting the number of specific reasons for a given score using the rubric and analyzing the quality of the analysis provided by the student (eg, were the reasons “superficial” and/or did they simply quoted information from the article, or did they demonstrate extended critical thinking and an application of understanding the features of a stronger or weaker article?). Class observations were made by the lead author (S.D.) during activity implementation in 1 laboratory for each of the 2 instructors. No direct interactions occurred between the faculty observer and the students or instructors during the observations. The purpose of the observations was to note general activity flow and areas of strength or difficulty for students, to observe case-study student interactions during the activity (which also were videotaped), and to observe differences in activity implementation between the Pharmaceutical Care Laboratory instructors. The information from the observations was used to make changes to the activity and instructional model and to identify potential study weaknesses as a result of differences in activity implementation.

Interviews with the case-study students were video recorded and occurred within 1 week of the activity to maximize activity experience recall. Five of the 6 case study students completed the interview following the activity. The case study interview questions are in Table 2.

In-class observations revealed some weaknesses in the initial activity design. Modifications included expanding the “big-picture” introduction to the activity and distributing the completed Vocabulary Sheet to all students.

Case-study student interviews and comments shared by non-case-study students with the instructors indicated...
that the activity was valuable, despite students’ general dislike of reading clinical literature. On a scale of 1 to 10, with 1 indicating “dump the activity” and 10 indicating that it was a “fantastic” activity, the case study students gave the activity a mean score of 7. The least popular element of the activity, which reduced its score, was the extensive writing (the pre- and post-activity written summaries). Students felt that writing at the beginning and end was arduous and did not really fit as a laboratory activity. The Pharmaceutical Care Laboratory faculty members also raised concern about the extensive writing as it was time-consuming to grade. In response to the concerns expressed by students and the faculty members who implemented the activity, the revised version of the activity eliminated the first writing summary. In subsequent implementations of the activity, the end-of-activity written summary was further reduced to coherently written bullet points where students highlighted the new information/awareness about the article that had led them to modify their score.

The students identified some improvements needed to the rubric including weighting certain elements of the evaluation and adding even more detailed descriptions. As a result of this feedback, the rubric was expanded to include additional details, particularly in areas where, based upon observations, students appeared to need additional support. (The revised rubric can be obtained online at http://hsc.unm.edu/pharmacy/DawnetalCTERubric.html.) Changes in student performance were assessed quantitatively and qualitatively by using pre- and post-activity Clinical Trial Evaluation Rubric scores that students assigned to the article and the differences between students’ pre- and post-activity written article evaluation summaries.

The article’s overall rubric evaluation score given by the students decreased significantly \((p < 0.005)\) from a pre-activity score of 76.7% to a post-activity score of 61.7%. The decrease in the rubric score indicated that students identified additional weaknesses in the article’s study design (a lower article score indicated a weaker study/publication), reporting of study data, and the authors’ interpretation of the data. Figure 2 shows article analysis scores given by students for specific elements in the major article sections (Introduction, Study Design, Statistics, and Results/Conclusions). Starred items in Figure 2 show significant differences (95% CI, \(p < 0.003\)) between pre- and post-activity scores for each of the article sections.

Detailed analysis of the pre- and post-activity summaries written by the case-study students \((n = 6)\) showed that the students added a mean of 4 additional discussion points per major article section in their post-activity evaluation of the article. Students also were much more specific in providing evidence and details to support their statements or arguments. For example, 1 case-study student went from basic, noncritical descriptions of the study,
which were often provided directly in the article, such as “The study was a double-blind, placebo controlled, randomized 12-week trial at a University research center” to more astute observations such as, “the authors of this paper are co-chairs for the [company that makes the] proprietary drug, therefore bias or potential conflict of interest [may be] present and could impact the study.” In the post-activity analysis of the article, this student identified 2 additional weaknesses of the study design, as well as 2 strengths, none of which were mentioned in her pre-activity analysis.

One of the most experienced case study students also modified her post-activity article evaluation summary and stated in her post-activity interview that she was “more critical of the Discussion section, I feel less satisfied that the results were conclusive based upon (1) the reduced power, (2) non-matched objective patient section, and (3) potential bias.”

To determine whether changes in article scoring reflected improved understanding of how to evaluate a research study, faculty experts independently evaluated the red yeast rice clinical article using the same rubric that the students used and then the student scores were compared with the consensus faculty expert score. The article score given by the faculty experts was 62.0%, compared to the students’ pre-activity score of 76.7% and post-activity score of 61.7% (Table 3). Student post-activity article scores were similar to faculty expert scores (< 1 point difference) for 12 of the 15 components.

**DISCUSSION**

The goal of this research was to collect data to test and refine an instructional model and activity to teach students to evaluate clinical literature. The hypothesis associated with this research was that an instructional approach using targeted scaffolding would improve student abilities to critically evaluate clinical literature compared to more traditional lecture-based didactic instruction. We found that using scaffolding in the form of the Clinical Trial Evaluation Rubric, the Vocabulary and Concepts Worksheets, and individual and small- and large-group instruction appeared to improve students’ ability to critically evaluate a clinical study better than lecture-based coursework alone. This conclusion was reached because students’ had already completed 80% of the Pharmacy Informatics and Research course (which consisted of lecture, examinations, and a major group literature evaluation project) before they began the Pharmaceutical Care Laboratory activity. Students demonstrated growth in identifying strengths and weaknesses in clinical literature, as evidenced by the changes in the students’ pre- and post-activity rubric scores and the additional explanations and details that they provided in the post-activity written summaries.

This study has several limitations including differences in the styles of instruction used by the 2 instructors, potential introduction of bias through phrasing of follow-up questions during case study interviews, and the potential that students’ article evaluation rubric scores improved simply as a result of a second readingreview of the study. The activity has been revised based upon observations and feedback from case study students and faculty members. Modifications include reducing the writing to a single end-of-activity bulleted summary in which students explain their reasoning for the Clinical Trial Evaluation Rubric score assigned to the article, increasing the details within the rubric, and distributing the completed Vocabulary and Concepts Worksheet directly to the students.

The question of whether changes in the article scoring reflected an improved or more accurate understanding among students of how to evaluate the quality of a research study remains and may be addressed further by exploring whether this instructional model for scaffolding enables students to independently differentiate a high-quality study from a weaker study. The authors have developed and implemented a second activity, based upon the model presented here, for which students evaluate 2 clinical studies on the same drug. Findings will be presented in a future publication.

It would be useful to determine whether students are able to apply the evaluation skills and knowledge gap self-assessment techniques independently over time and in different elements of evidence-based medicine. Anecdotal observations of the same students in the year following this study indicated that although the students were better able to evaluate the quality of studies (compared to students in previous years), their knowledge and understanding of how to apply primary literature in clinical settings was

---

**Table 3. Expert Consensus CTE Rubric Score Compared to Students’ Pre- and Post-Activity CTE Rubric Scores (n=51)**

<table>
<thead>
<tr>
<th></th>
<th>Introduction</th>
<th>Study Design</th>
<th>Statistics</th>
<th>Results/Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Expert</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Student Pre-Score</td>
<td>2.5</td>
<td>2.2</td>
<td>3.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Student Post-Score</td>
<td>2.3</td>
<td>0.7</td>
<td>3.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

* CTE Rubric score difference of greater than 1 point.
still weak. This suggests that additional explicit instruction is needed to assist students in developing other/stronger evidence-based medicine skills, such as assembling a body of literature and accurately applying data in clinical situations.

**SUMMARY**

Using scaffolding in the form of the Clinical Trial Evaluation Rubric, the Vocabulary and Concepts Worksheets, and individual and small- and large-group instruction appeared to improve students’ ability to more critically evaluate a clinical study compared to the didactic coursework alone. The article’s overall rubric evaluation score given by the students decreased significantly ($p < 0.005$) from a pre-activity score of 76.7% to a post-activity score of 61.7%, similar to the article score assigned by faculty experts (62.0%). The decrease in the students’ post-activity rubric score indicated that they identified additional weaknesses in the article’s study design (a lower article score indicated a weaker study/publication), reporting of study data, and the authors’ interpretation of the data. Detailed analysis of summaries written by the case-study students (n = 6) showed that the students added a mean of 4 additional discussion points per major article section in their post-activity article evaluation.

**ACKNOWLEDGMENTS**

Grant funding from the New Mexico Center for Teaching Excellence supported this research. The authors extend special thanks to Melanie Dodd, PharmD, who served as an expert reviewer to score the red yeast rice clinical article and Ludmila Bakhireva, MD, PhD, who provided constructive input into the Clinical Trial Evaluation Rubric.

**REFERENCES**

25. McNamara DS, Kintsch E, Songer NB, Kintsch W. Are good texts always better? interactions of text coherence, background knowledge,

