Teaching surgical residents to evaluate scholarly articles: a constructivist approach

Tiffany Chichester, M.D., Karen Hagglund, Ph.D., Elango Edhayan, M.D., M.A.C.M.*

Department of Surgery, St. John Hospital and Medical Center, 22151 Moross Road, Suite 212, Detroit, MI 48236, USA

KEYWORDS: Educational intervention; Constructivist education; Evidence-based literature; Journal club; Surgical education

Abstract

BACKGROUND: The ability to critically appraise scholarly journals is an essential skill for surgical residents in their journey to being lifelong learners. Methods to teach evaluation of scholarly articles are scant in the educational literature.

METHODS: Residents completed a pre-test on evaluation of surgical literature. Two lectures on methodology and statistics followed. A board-certified surgeon and residents evaluated 7 articles using a scoring form. A post-test followed. Four additional sessions on evaluating surgical articles were held without the rating form. Residents completed a second post-test.

RESULTS: Residents showed improved knowledge and skills on evaluating surgical literature on the first and second post-tests when compared with the pretest (P < 0.0005 for both tests).

CONCLUSIONS: Surgical residents can be taught to evaluate literature using constructivist Educational theory. There was a significant improvement in knowledge and the skills of literature evaluation, which persisted after the educational aid was removed.

The rapid creation and easy dissemination of surgical knowledge has created a critical need for surgeons to acquire skills for understanding and incorporating change into their practices.1,2 The American Council for Graduate Medical Education (ACGME) calls for training programs to ensure that residents demonstrate “the ability to locate, appraise and assimilate evidence from scientific studies related to their patients health problems.”3 Teaching and the retention of these skills among trainees has been variable.4,5

Constructivism is an educational theory in which learner construct their own understanding and knowledge of the world primarily by experimentation and reflection.6 Learning by creating problem-solving tasks for the learners to interact with to create knowledge and understanding that is individualized by each learner.7 A “learning scaffold” is an educational tool used in constructivism to support learners when new concepts are introduced. A learning scaffold is used to decrease the cognitive load while allowing learners to develop fresh and personal insights into the problem.8,9 When the learners achieve proficiency in the task, the learning scaffold is removed. Because the knowledge is now internalized and personal to the learners, they can function without the scaffold.9

Surgical residents, being motivated adult learners, are prime candidates for learning using the constructivist epistemological method. Each resident brings previous knowledge from life and scholastic experiences to the learning environment of a residency program. These...
previous experiences are modified fresh insights, and by constant questioning of themselves they are transformed into “expert learners.” On transforming into expert learners, residents gain an armamentarium of fresh insights they use as lifelong learners.9

This study’s hypothesis states that surgical residents, being motivated adult learners, can be taught to evaluate and understand surgical literature at the level of a board-certified surgeon by using the constructivist educational theory. An evaluation form serves as the learning scaffold to aid the residents in understanding, internalizing, and retaining this skill.

**Methods**

The study was approved by the hospital institutional review board. This study was conducted at an ACGME-accredited surgical residency program. All surgical residents were enrolled. Twenty-six residents participated, including 13 junior residents (program years 1 and 2) and 13 senior residents (program years 3 to 5).

Residents were given a pretest to assess their current knowledge about the evaluation of surgical literature. The test evaluated their current knowledge of research design, basic statistics commonly used in research papers, and clinical applicability of research projects. The test was followed by 2 lectures on these 3 topics, knowledge of which is needed to evaluate surgical literature.

Over the next 4 months, residents participated in 6 workshops. Each workshop consisted of reading and evaluating an article using a rating form. Each resident independently read the article and rated the article based on criteria given in the rating form. This was done before each of the sessions. The rating form was a modified version of an evaluation form for research articles. A previous study had tested the evaluation form for intra- and inter-rater reliability.10

Session 5 had 2 articles for evaluation and rating. The remaining 5 sessions used 1 article each. The articles were primarily chosen to expose the residents to varied research designs and for their clinical relevance to surgeons.

A board-certified practicing surgeon evaluated the articles using the same rating form. The surgeon was chosen because he has acquired additional expertise in research design as part of his course work for a Masters in Academic Medicine program. It was felt that because of his additional specific training in research design, he would represent a meaningful level of competency to which residents could aspire.

The workshops consisted of interactive discussions between the faculty member and the residents. The discussions were based on the evaluation of the articles in the rating form. Concepts of study design, appropriateness of statistical testing, and clinical applicability were discussed.

The evaluation forms were collected at the end of each session. The scores on the forms were compared to assess development of congruence between the surgeon and the residents.

A month after these 6 workshops, residents took a post-test. Three additional articles were evaluated over the next 3 months. Residents and the faculty member did not use the rating form to evaluate the article and no scores were given. The faculty member led these workshops using components from the evaluation form as a guide. A second post-test was administered.

Statistical methods included paired t tests and analysis of variance (ANOVA). Paired t tests were used to evaluate improvement in test scores before and after the intervention. ANOVA was used to assess correlation between the scoring forms of the surgeon and those of the residents. The t test was used to evaluate independent groups.

**Results**

Twenty-six residents participated in this study. There were 13 junior residents (program years 1 to 2) and 13 senior residents (program years 3 to 5). All 26 residents took the initial pretest. Twenty-one residents (11 junior residents and 10 senior residents) took the first post-test.

Test scores improved between the pretest (mean ± standard deviation [SD]) (46.50 ± 14.70) and the first post-test (66.75 ± 14.07) for all residents. Paired t tests showed this difference to be statistically highly significant (P < .0005). There was no statistically significant difference in pretest scores between the junior residents (43.18 ± 14.54) and senior residents (49.55 ± 14.05). There was no statistically significant difference in the post-test scores between the junior residents (61.36 ± 13.43) and the senior residents (72.50 ± 12.08). Although the senior residents scored higher than the junior residents in both the pretest and the first post-test, there was no statistically significant difference between the 2 groups.

ANOVA was used to compare scores from the evaluation form for congruence between the residents and the attending surgeon. ANOVA showed no difference in scores between the surgeon and residents on 5 articles (articles 1, 2, 3, 4, and 7). Articles 5 and 6 were scored significantly lower than articles 1, 2, 3, 4, and 7 by residents (ANOVA P < .0005). The faculty mentor also scored articles 5 and 6 lower than articles 1, 2, 3, 4, and 7 (P < .002). When compared with the other articles used in this study, articles 5 and 6 were deemed to be inferior in research design, statistical testing, and clinical applicability by this rating method. The surgeon scored these 2 articles lower than did the residents. A difference of scores for each of these articles between the surgeon and the residents was found by the t test (P < .01).

All 26 residents took the second post-test. They showed a further improvement in the mean scores after the second post-test (75.68 ± 10.5). This improvement was highly significant (P < .0005) when compared with the initial pretest. The mean scores of the junior residents improved from
a pretest mean of 43.18 (14.54 SD) to 72.27 (10.57 SD). This difference was statistically significant at \( P < .001 \). The senior residents improved from a mean of 49.55 (14.05 SD) to a mean of 79.09 (9.7 SD). This difference was highly statistically significant at \( P < .0005 \).

There was no significant difference in the means between the first and second post-tests among the groups. There was also no statistically significant difference between the mean scores of the first and second post-tests among the group as a whole.

**Comments**

Constructivist learning theory is a process that allows learners to face a problem and develop an understanding that is individualized by this experience.\(^8,^9\) The students can then act on this knowledge to consolidate their understanding. Constructivist educational theory assumes that the learners are self-directed and motivated. Surgical residents, having declared themselves to be self-motivated adult learners by subscribing to a surgical residency, are excellent candidates for education through constructivist education theory. Surgical educators use constructivism in educating residents, usually without recognizing the pedagogy to which they are subscribing. Most surgical education comes in multiple short periods of interaction between faculty and residents. Educators mentor residents through various problem-solving tasks, which modify existing knowledge in the residents, thereby creating fresh understanding and knowledge. Educators and residents are using constructivism in their teaching and learning without overtly naming it as such.

Lev Vygotsky was a principal proponent of the constructivist learning theory.\(^1^1\) In the constructivist model, residents interact with the faculty and actively participate in the learning process by speech, memory, and action. The social and mental interactions with the faculty member are internalized by the residents in the development of problem-solving skills that ultimately are unique to each resident. Central to Vygotsky’s constructivist learning theory was the zone of proximal development (ZPD). Vygotsky defined the ZPD as “The distance between the actual developmental level of learners as defined by independent problem solving and the level of potential problem solving under adult guidance or in collaboration with more capable peers.” The role of the surgical educator is to collaborate and encourage residents to attain higher levels in the ZPD. The educator may use scaffolding techniques that include writing exercises, questioning, role modeling, and feedback, which are implemented to encourage learning and to move the residents through their personal ZPDs. When the learning becomes internalized, assistance from the faculty is no longer needed, and the residents are said to have successfully navigated their ZPDs. The scaffolds that were used to successfully navigate the ZPD may now be removed because the residents can function independently and work their way through the given task.\(^1^2\)

This study created a constructivist learning environment in which the learning task was presented without cognitive overload. A form was used to serve as a learning scaffold. This rating form was a modification of the form published by Reed et al\(^1^0\) in the *Journal of the American Medical Association* in September 2007. The rating form was chosen because of its high inter- and intrarater reliability. The rating form served as an educational scaffold to allow interns as well as senior residents to interact with the problem. The rating form was the basis for resident interaction during the workshops in which articles were discussed. Residents constructed their own knowledge on the real-world problem of evaluating scholarly articles without that knowledge being given to them.

In the appropriate environment, surgical residents can be taught to evaluate and derive conclusions from journal articles similar to the conclusions of a practicing surgeon. This task was made easier with the use of a rating form, which served as a scaffold in the development of new knowledge and skills by the residents. The form reduced the cognitive load on the residents and allowed them to construct individual meaning from the evaluation process. For many articles, the rating form eliminated the knowledge gap that exists between a resident in training and a board-certified practicing surgeon. Residents attained a level of understanding of an article similar to that of the surgeon by using the rating form. Our study showed no difference between the faculty member and the residents on rating form scores. The scores showed significant correlation when the quality of the study was good.

Only 2 initial lectures were held on this topic during the entire year. The rest of the sessions were interactive workshops in which there were 2-way interactions. Use of the rating form allowed the discussions between residents and faculty to be on an equal intellectual field. Constructivist learning depends on learners questioning and reflecting, thereby seeking their own meaning and internalizing knowledge.\(^8\) The rating form allowed such learning because resident understanding of the articles was greatly aided by the learning scaffold.

An interesting finding of this study was that correlation between the attending surgeon and the residents broke down when the quality of the study was poor. Articles 5 and 6 were deemed poor in methodology and merit by both the residents and the faculty mentor. However, residents scored the articles significantly higher than did the faculty member. This implies that residents tend to believe most published scholarly articles, whereas the attending surgeon treats poorly conceived articles with skepticism. Surgical educators should pay attention to this trend. Although residents correctly identified that the articles were scientifically inferior, the level of trepidation did not reach that of the faculty member. This could lead to surgical residents internalizing more than they should from inferior articles.
This study does have its limitations. The study used written tests as a measure of knowledge. There are many variables that can affect performance on a written test. The study tried to overcome this problem by not announcing the test. Residents did not know the study design and were not aware they were participating in a study. They were not aware of when they were going to be given a test or if they were going to be tested at all. We believe that this study replicates true baseline knowledge and gain in knowledge.

Another major constraint of this study is that there was only one faculty member who served as the reference to whom residents were compared. The faculty member was chosen primarily because of his advanced degree in education in addition to his board certification. The faculty member had completed Masters-level courses in study design as part of his Masters in Academic Medicine program. It was felt that his ability to evaluate articles would be greater than that of an average board-certified surgeon. The study was designed to see if residents can attain a level of competency similar to a surgeon with advanced training in research design. A more robust study design would have been to have all faculty members participate. The study could have been improved by having specialists in each field discuss each article.

This study has shown that the residents retained knowledge and skills. It does not prove that they have applied their skills in direct patient care. Such a longitudinal study will be difficult to perform logistically, and outcomes will be difficult to measure because they are multifactorial. However, like most aspects of a training program, residents learned and retained an essential skill.

In conclusion, our study shows that residents can be taught to evaluate surgical articles on the level of a practicing surgeon. This knowledge is retained by repetitive practice. This intervention not only satisfies an ACGME program requirement, it also has prepared this group of residents to critically appraise scholarly journal articles. A more robust study design would have been to have all faculty members participate. The study could have been improved by having specialists in each field discuss each article.

Discussion

Dr Norman Estes (Peoria, IL). I read this article with the belief that the best way to learn to evaluate a scholarly article is to do scholarly research and to have publication experience. I still have that belief. However, this article has great importance for those of us leading the challenge for surgical education. Do you believe the rating form that served as the learning scaffold provided a means of interpreting the scholarly article or did it just serve as an aid for scoring well on the post-test? Did previous research or publication experience benefit the learner on the pretest? Would the value of the interactive discussion have been enhanced if additional attending physicians had participated and given their varied input, as occurs in many journal clubs. Lastly, was a final assessment made as to whether the residents felt that they had developed increased confidence to assess the scholarly article after the program?

Dr Chichester. The intervention did help the residents learn statistics and ways of research design, but it was done in the form of an evaluation form, so it was a way to interpret the articles, their merit, and to see if they were actually of high quality, and the residents were able to discern that. As to the previous experience of our residents, we think that the residents come in with their own previous knowledge and experience, and publications and research help that, but our residents were pretty much on the same playing field because no one had an advanced degree in research design or statistics, and no one had advanced course work. We chose our preceptor because of his advanced course work, his Masters in academic medicine, and how he had course work in research design and statistics.

Dr Scott M. Wilhelm (Cleveland, OH). Is journal club at your institution a mandatory thing for all residents to attend? If not mandatory, the people who come are people

References

who are already interested in reading. Do you think that puts any bias into your study or not?

**Dr Chichester.** This study was actually time set aside from journal club, so it was mandatory for the residents to come, as well as the attending physician who served as precept. This was special time set aside in addition to journal club, so we had great participation from our residents.

**Dr Vijay Mittal** (Southfield, MI). Did the article belong to the specialty of your attending physician? Because you need to have variability. A vascular surgeon attending physician, a general surgeon, or a trauma surgeon may not be the same thing.

**Dr Chichester.** We did have the one attending physician who we chose because of his qualifications, but he was a general surgeon, so he was able to, you know, give us background on all aspects of the surgical literature that we assessed. Some of our articles were trauma, some were general surgery, and some were vascular, so we had the competency of a general surgeon who would be able to offer his perspective.