

CONFERENCE PANEL DISCUSSION

Overcoming Barriers to Addressing Education Problems With Research Design: A Panel Discussion

Lalena M. Yarris, MD, MCR, Larry D. Gruppen, PhD, Stanley J. Hamstra, PhD, K. Anders Ericsson, PhD, and David A. Cook, MD, MHPE

Abstract

A plenary panel session at the 2012 *Academic Emergency Medicine* consensus conference "Education Research in Emergency Medicine: Opportunities, Challenges, and Strategies for Success" discussed barriers educators face in imagining, designing, and implementing studies to address educational challenges. This proceedings article presents a general approach to getting started in education research. Four examples of studies from the medical education literature that illustrate a distinct way to approach specific research questions are discussed. The study designs used are applicable to a variety of education research problems in emergency medicine (EM). Potential applications of studies are discussed, as well as effects and lessons learned.

ACADEMIC EMERGENCY MEDICINE 2012; 19:1344-1349 © 2012 by the Society for Academic Emergency Medicine

Most instructional methods in medical education are based on tradition and philosophy rather than the results of well-designed studies. Despite the plethora of educational problems that emergency medicine (EM) educators face, critical appraisals of education research studies in the EM literature in the past 3 years have identified few original contributions that reach publication.¹⁻³ Common barriers medical education researchers encounter include lack of training opportunities, protected time, or funding for

research; competing administrative and leadership roles; small numbers of learners; and difficulty defining relevant and measurable outcomes.⁴ Among published studies, deficiencies in reporting quality are common.⁵⁻¹⁰

The purpose of this article is to provide readers with strategies to overcome the barriers they face in the education research process, from idea to publication. We begin by summarizing a three-step approach to education research design and presenting practical tips for getting started. We then summarize four original research studies in health professions education to illustrate how educational problems can be addressed with research studies.

GETTING STARTED IN EDUCATION RESEARCH

Three Steps in Planning a Scholarly Project

Beckman and Cook¹¹ have presented a three-step approach to developing scholarly projects in medical education:

1. Define the Study Question. The most important part of any study is the research question (or study goal or hypothesis). Even the most rigorous study will fail to have an effect on the field if it does not answer a question that is both important (i.e., the answer would influence practice or future research) and novel (i.e., the answer is unknown).

The best way to establish a question's importance is to employ a strong conceptual framework, which is a theory, approach, or model for how things work. This allows educators at future times and in other settings to adopt and build on a study's results.¹² Conceptual

From the Department of Emergency Medicine, Oregon Health & Science University (LMY), Portland, OR; the Department of Medical Education, University of Michigan Medical School (LDG), Ann Arbor, MI; the Academy for Innovation in Medical Education, University of Ottawa Skills and Simulation Centre (SJH), Ottawa, Ontario, Canada; the Department of Psychology, Florida State University (KAE), Tallahassee, FL; and the Division of General Internal Medicine, Office of Education Research, Mayo Clinic College of Medicine (DAC), Rochester, MN.

Received June 27, 2012; accepted June 28, 2012.

This paper reports on a panel discussion of the 2012 *Academic Emergency Medicine* consensus conference, "Education Research in Emergency Medicine: Opportunities, Challenges, and Strategies for Success," May 9, 2012, Chicago, IL.

The authors have no relevant financial information or potential conflicts of interest to disclose.

Supervising Editor: Terry Kowalenko, MD.

Address for correspondence and reprints: Lalena M. Yarris, MD, MCR; e-mail: yarrisl@ohsu.edu.

frameworks also assist educators in defining and selecting the study variables and in predicting and interpreting the results.¹³ A question's importance also derives from its timeliness or effect on current practice, but in the absence of a conceptual framework such answers will usually have limited influence over time. Cook and others^{14,15} have argued that the most important questions—those with the greatest potential for long-term effect—are those that *clarify* how things work, for whom, and in what circumstances.

The best way to establish a question's novelty is to rigorously review the literature. Even if no studies are found for the question being considered, the researcher should diligently seek studies that address similar questions or that may offer suggestions for key elements of the framework, training intervention, or assessment. Such relevant work may often be found in other fields such as surgery, nursing, or non-health professions literature. The literature review should culminate in a focused problem statement—a clear summary of what remains unknown. The research question (or objective) then follows naturally from the problem statement: "The evidence thus far shows _____. What remains unknown is _____. Thus, we sought to _____."

Educators can identify important questions in their daily teaching activities each time they ask, "I wonder if we should do it [this way] or [that way]?" If a literature search for evidence and theory-based principles does not answer the question, this may be the start of a worthwhile project.

2. Identify Study Designs and Methods. Three guiding principles for research study design are:

First, identify the general class of study design most relevant to the research question. This is essential, because the standards for excellence and potential pitfalls vary depending on the design. Studies of educational interventions (e.g., a new course or training approach) will require an experimental study design, studies that evaluate educational assessments (e.g., a new testing approach or tool) will invoke a validity study design, studies seeking to understand current attitudes or learning needs may employ a cross-sectional (e.g., survey) design, and studies attempting to summarize published literature will use a systematic or nonsystematic literature review.^{16–20}

Second, realize that all studies have flaws. However, some flaws matter more than others. In addressing study flaws it is helpful to focus on reproducibility and validity threats, i.e., threats to the validity of study interpretations. Ask, "If another researcher at another institution were to conduct a study using different methods but addressing the same question, would he or she be likely to arrive at the same answer?" Answering this question will focus primarily on threats to study validity—issues such as preexisting differences among participants, selection bias, instrumentation problems, and implementation bias.¹⁶ The most important threats, and the specific ways in which they can manifest, vary for different research designs.

Third, be realistic. Carefully consider factors such as the number of potential participants, the logistics of implementing an intervention, and the feasibility of assessing outcomes. The standards of Glassick et al.²¹

for scholarship are useful in judging the rigor of your scholarly project, however small. Estimate the needed sample size using a credible approximation of the anticipated effect.²² It is also important that they ensure that institutional review board approval is obtained, as learners are a vulnerable study population.²³

3. Select Outcomes. Nearly all research studies require some type of outcome. To avoid prematurely settling on an inferior outcome measure, we suggest first identifying the broad class of outcomes (e.g. "knowledge" or "skill" [in a controlled setting] or "behaviors" [at the bedside]), then considering different measures available for a given outcome, and then selecting a specific instrument.²⁴ Ten tips for success in education research have been published previously and are summarized in Table 1.²⁵

EXAMPLES OF EDUCATION RESEARCH STUDIES

Example 1: Study of Fidelity in Simulation-based Education

Our first example is a randomized trial comparing different approaches (different fidelity) for simulation-based education.²⁶ The primary investigator of this study began with a general interest in developing a simulator to train urologists and answering two questions: how can simulation be effective, and what are the essential key steps for training a particular procedure? Guided by a literature search, he defined a specific research question: is a bench model superior to didactic training for learning a technical skill and is there a difference in training effect between low-fidelity and high-fidelity models? The conceptual framework for this study drew upon the literature regarding transfer of learning, which addresses the factors that influence how learned knowledge or skills from one setting can be effectively applied to another. Instead of comparing training on a simulator to training in a clinical setting, the investigators chose to compare one simulator to another. The research hypothesis was twofold: 1) that both hands-on training models (low- and high-fidelity) would lead to greater gains than the didactic training and 2) there would be no difference in effect between high-fidelity training and low-fidelity training.

The primary investigator identified a procedure that he wanted to teach (midureter kidney stone extraction) and defined the key steps that needed to be performed to complete this procedure. These steps were crucial to developing both a procedure checklist and a low-fidelity model that the authors hypothesized might be as effective as currently available, expensive, high-fidelity models. Forty fourth-year medical students on their surgical rotation were recruited to participate in a randomized, prospective three-arm study. This study had a pre-post design: all subjects were tested before and after the intervention on a high-fidelity simulator. The three interventional arms were a control arm (didactic presentation only), a low-fidelity training arm, and a high-fidelity training arm.

The results supported that hands-on training was superior to didactic instruction. Students trained on both the high- and low-fidelity models did significantly better than the group that received didactic instruction

Table 1
Tips for Getting Started in Education Research

Tip	Brief Notes
Get some training	Training can occur locally (local workshops), at national meetings (in particular at education-oriented conferences), and through formal degrees (e.g., masters-level training).
Find a mentor	The importance of a mentor cannot be overemphasized. If no single person at your institution possesses the needed skills or sufficient time, consider working with multiple mentors or looking outside your institution.
Ask important questions	See main text.
Start small and grow	Powerhouse investigators did not start off that way; they started with small projects and grew into their current position. A poster at a national meeting may be the first step in a series of progressively insightful studies.
Aim high	It is okay to start small, but don't settle for mediocrity. Whatever you do, do it well. The standards of Glassick et al. ²¹ for scholarship are useful in judging the rigor of your scholarly project, however small.
There's no such thing as a perfect study	If you wait for the perfect study (or perfect opportunity), you'll be waiting a very long time. If the question is important and novel, any answer (even if incomplete) will be better than none.
Where will you find the time?	Be realistic in budgeting time, money, and other resources.
Remember ethical issues	Learners are a vulnerable population. ²³ Many education journals now require institutional review board approval for all studies involving human subjects.
Network	Build relationships locally and nationally through involvement in committees and professional organizations. Approach new contacts with a specific purpose (e.g., a question).
This is hard work	Education research is fun and rewarding, but it is not easy.

only on all outcomes (global rating scale, checklist score, pass rating, and time to task completion.) However, there were no significant differences between the low- and high-fidelity groups with respect to the same measures. These results suggest that as long as a low-fidelity model is designed to incorporate essential features required in an actual procedure, a realistic visual appearance may not be necessary.²⁶

Some lessons can be taken away from this example.

1. In general, generating project ideas is the easy part. The majority of the time invested in the project will be spent refining the specific research questions and hypotheses, evaluating their relation to the literature, and navigating logistical issues and project management issues as they arise.
2. Do not underestimate the potential downstream effect of even small-scale studies. This example answered a relevant research question, added new knowledge to the existing literature, and established a foundation for this young researcher to develop a niche in education research, which led to several subsequent studies.²⁷⁻³⁰
3. It is helpful to start with general intuition or "gut instinct." This will help ensure that the researcher has a passion for the research question, which is necessary because the path toward scholarship will require a significant amount of work. However, it is important to follow that up with discussions with an expert in the field and a thorough literature review to make sure one's gut is not leading one astray.
4. Be flexible and thoughtful in developing the specific research question. The final question might be quite different than the initial draft.

Example 2: Study of an Innovative Approach to Interprofessional Communication

Our second example is an evaluation of an innovative approach to improving the team communication skills

of fourth-year medical students.³¹ The investigator team set out to develop a novel approach to measure the communication and decision-making involved in telephone consultations with nurses about patient care problems and an accompanying measurement instrument.

A multidisciplinary team of physicians, nurses, and education experts developed a set of standardized surgical patient scenarios that might prompt a nurse to page a physician (mock pages). They also developed a structured checklist of desirable and undesirable questions or actions that the responding physician might ask or take. Study participants included 79 fourth-year medical students participating in surgery "boot camp" programs at five institutions (trained group) and 10 new surgery interns at two institutions (untrained group). A trained nurse periodically paged participants and presented them with hypothetical surgical scenarios and then assessed their performance using the checklist.

Inter-rater reliability and internal consistency were acceptable to high (alpha from 0.65 to 0.92). The results demonstrated clear differences in the level of performance of the trained versus the untrained participants, with large effect sizes. This provides evidence for the validity of skill judgments made on data from this assessment tool.³¹

Although designed as an assessment, the mock page technique has also proven to be a valuable training resource and has been well received by students and faculty from each of the participating institutions. The mock page curriculum has subsequently been adopted by other departments and institutions, and there is current discussion about including it in the "National Boot Camp Curriculum" for surgery clerkships.

As an example of how an educational problem can be addressed through research, this study offers several lessons:

1. Prior to developing the mock pages cases for this study, the team conducted a careful review of the

existing literature to avoid duplicating efforts and to take full advantage of findings from other researchers.

2. The investigator team worked diligently to provide a consistent intervention among the participating medical schools.
3. The multi-institutional nature of the research was another strength and highlights the importance of communication and attention to the well-being of the collaborative research team.
4. The results of the study are meaningful because they reflect clinically and educationally relevant outcomes (communication and decision-making) as well as both statistically and practically relevant effects from the intervention. Therefore, this project had both beneficial research outcomes and practical implications for educational practice.

Example 3: Study of Improving Radiology Interpretation Using Deliberate Practice

Our third example is a cross-sectional study that assessed the individual learning curves of trainees as they engaged in deliberate practice in interpreting ankle radiographs.³² Research in a wide range of domains, such as chess, sports, and music, has suggested that experiential learning is not optimal for developing expertise.³³ The most potent predictor of performance improvement in these domains is the amount of accumulated engagement in deliberate practice, where: 1) individuals engage in tasks with the explicit goal of improving a particular aspect of performance without a teacher; 2) tasks offer accurate and immediate feedback; 3) individuals can engage in the same or similar tasks and can thus gradually improve by repetition, successive refinement, and problem-solving; and 4) individuals can engage in the training when they are ready and rested, until concentration is reduced.³⁴

The authors of this study wanted to build on prior assessments of the effectiveness of deliberate practice for improving performance in medical tasks by specifically demonstrating how learning curves can describe proficiency improvements.³⁵ The authors aimed to demonstrate that learning curves could be used by medical educators to define at which point practice is most efficient and how much practice is required to achieve a defined level of mastery for a specific task.

Researchers created a case bank of 234 pediatric ankle radiographs with official case reports. Each case presented a clinical summary of patient symptoms, followed by a three-view ankle radiograph series. Thirty-two pediatric trainees rotating through a pediatric ED were recruited to participate in the study, along with six pediatric EM attending physicians to serve as a reference standard. Subjects were prompted to classify the case as normal or abnormal, and identify the location of the abnormality if applicable. Subjects were then provided with immediate feedback in the form of the official radiology report. Cases were scored and longitudinal learning curves were generated based on calculated test characteristics (such as accuracy, sensitivity, and specificity). While individuals displayed dramatic differences in learning curves, group-level summary curves for each of the test characteristics were similar.

An overall pattern demonstrated a phase of irregular performance until 20 cases and then a maximal increase in performance over time from 21 to 50 cases. An inflection point where learning slowed was noted at 50 cases, but learning did not stop even after all cases had been completed.

This example illustrates how deliberate practice can be studied in an EM setting and also how learning curves can describe performance improvements in learners over time. This may allow medical educators to understand both their learners' rate of learning and the effectiveness and efficiency of a specific learning intervention for either a group or an individual. These results may have applicability to other types of medical skills, such as exam skills (cardiac auscultation, pulmonary auscultation), other imaging interpretation, visual diagnosis, or perceptual skills training.

Example 4: Study of a New Tool for Simulation-based Assessment

Central venous catheterization (CVC) is a procedure commonly performed by practitioners in many medical specialties. Several studies have evaluated how to train health professionals in CVC using simulation.³⁶ However, assessments of CVC proficiency have infrequently been described. The purpose of this study was to evaluate an instrument designed to determine if residents have mastered the key steps of inserting an internal jugular central line.³⁷ The study team began by carefully developing an instrument based on institutional practice guidelines and by having the instrument reviewed by several internal focus groups. They videotaped the participants of a training workshop as they performed a CVC. These videotapes were then independently reviewed by two reviewers. However, the team encountered difficulty interpreting the results of data analyses. At this point, a new investigator was added to the team—a researcher with experience planning, conducting, and publishing validity studies. This new addition to the team suggested a new conceptual framework for thinking about validity studies. Over the course of several meetings, the study team refined the scope and aims of the study, defined the target audience, selected a target journal, and revised the analysis plan. The project was slightly delayed when one video reviewer was delayed in completing his review assignment, but ultimately the study was published in a well-respected journal.³⁷

Key lessons from this experience include:

1. Plan the study well in advance. In this case, a mid-stream reformulation was possible without sacrificing the rigor of the final product. However, not all studies are so fortunate; failure to plan carefully will often allow a flaw that might otherwise have been easily avoided with a minor adjustment in methods. Even in this case, advance planning might have permitted collection of additional useful information during the earliest stages of the study.
2. Pay close attention to team composition. If the team lacks necessary expertise, it helps to get an expert involved early. Ensure that all team members are committed to the project, that roles are clearly defined, and that each member has sufficient

resources (including time) to complete required tasks. Adjustments in the roles and expectations of individual members may sometimes be necessary midway through a project.

3. Conceptual frameworks apply not only to study interventions, but also to study designs. In this case, the use of a modern framework for thinking about the validity of assessment scores added substantially to the rigor and clarity of the final report.
4. The old framework for validity studies of face validity, content validity, criterion validity, and construct validity has been replaced by a new model. In the new model, validity is viewed as a hypothesis that can be tested through the collection and synthesis of validity evidence.³⁸ Evidence can derive from sources of content, internal structure, relations with other variables, response process, and consequences. This evidence is then synthesized to construct a rational validity argument.¹⁷

SUMMARY

Early career education researchers can overcome barriers to conducting quality studies by defining relevant research questions that stoke their passion, applying a structured approach to the research process, and selecting appropriate designs and methods. A well-designed study that addresses a timely educational problem can create a ripple effect by raising new questions and inspiring follow-up studies that will affect educational practice, contribute to the collective knowledge, and promote researcher career advancement. By engaging in sound education research practices and challenging our colleagues and collaborators to set and achieve high standards for study quality, we can transform the state of medical education to meet the ever-changing needs of educators, learners, and ultimately patients.

References

1. Farrell SE, Coates WC, Kuhn GJ, Fisher J, Shayne P, Lin M. Highlights in emergency medicine medical education research: 2008. *Acad Emerg Med.* 2009; 16:1318–24.
2. Kuhn GJ, Shayne P, Coates WC, Fisher J, Lin M, Maggio LA, Farrell SE. Critical appraisal of emergency medicine educational research: the best publications of 2009. *Acad Emerg Med.* 2010; 17(Suppl 2):S16–25.
3. Shayne P, Coates WC, Farrell SE, et al. Critical appraisal of emergency medicine educational research: the best publications of 2010. *Acad Emerg Med.* 2011; 18:1081–9.
4. Dyrbye LN, Thomas MR, Papp KK, During SF. Clinician educators' experiences with institutional review boards: results of a national survey. *Acad Med.* 2008; 83:590–5.
5. Wolf FM. Methodological quality, evidence, and research in medical education (RIME). *Acad Med.* 2004; 79(Suppl 10):68–9.
6. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. *Med Educ.* 2007; 41: 737–45.
7. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. *JAMA.* 2008; 300:1181–96.
8. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA.* 2007; 298:1002–9.
9. Baernstein A, Liss HK, Carney PA, Elmore JG. Trends in study methods used in undergraduate medical education research, 1969–2007. *JAMA.* 2007; 298:1038–45.
10. Cook DA, Levinson AJ, Garside S. Method and reporting quality in health professions education research: a systematic review. *Med Educ.* 2011; 45:227–38.
11. Beckman TJ, Cook DA. Developing scholarly projects in education: a primer for medical teachers. *Med Teach.* 2007; 29:210–8.
12. Bordage G. Conceptual frameworks to illuminate and magnify. *Med Educ.* 2009; 43:312–9.
13. Giacomini M. Theory-based medicine and the role of evidence: why the emperor needs new clothes, again. *Perspect Biol Med.* 2009; 52:234–51.
14. Cook DA, Bordage G, Schmidt HG. Description, justification, and clarification: a framework for classifying the purposes of research in medical education. *Med Educ.* 2008; 42:128–33.
15. Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: what are they and what can they contribute? *Med Educ.* 2012; 46:89–96.
16. Cook DA, Beckman TJ. Reflections on experimental research in medical education. *Adv Health Sci Educ Theory Pract.* 2010; 15:455–64.
17. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med.* 2006; 119:166e7–16.
18. Dillman D. *Mail and Internet Surveys: The Tailored Design Method*, 2nd ed. New York, NY: JohnWiley & Sons, Inc., 2000.
19. Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach. *Med Educ.* 2012; 46:943–52.
20. Eva KW. On the limits of systematicity. *Med Educ.* 2008; 42:852–3.
21. Glassick CE, Huber MT, Maeroff G. *Scholarship Assessed: Evaluation of the Professoriate*. San Francisco, CA: Jossey-Bass, 1997.
22. Cook DA. If you teach them, they will learn: why medical education needs comparative effectiveness research. *Adv Health Sci Educ Theory Pract.* 2012; 17:305–10.
23. Roberts LW, Geppert C, Connor R, Nguyen K, Warner TD. An invitation for medical educators to focus on ethical and policy issues in research and scholarly practice. *Acad Med.* 2001; 76:876–85.
24. Cook DA. Twelve tips for evaluating educational programs. *Med Teach.* 2010; 32:296–301.
25. Cook DA. Getting started in medical education scholarship. *Keio J Med.* 2010; 59:96–103.
26. Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The effect of bench model fidelity on endourologic

- skills: a randomized controlled study. *J Urol.* 2002; 167:1243–7.
27. Grober ED, Hamstra SH, Wanzel KR, et al. Laboratory based training in urological microsurgery with bench model simulators: a randomized controlled trial evaluating the durability of technical skill. *J Urol.* 2004; 172:378–81.
 28. Matsumoto ED, Pace KT, Honey R, John D. Virtual reality ureteroscopy simulator as a valid tool for assessing endourological skills. *Int J Urol.* 2006; 13:896–901.
 29. Samuelson ML, Cadeddu JA, Matsumoto ED. Laparoscopic decision making: impact of preoperative reading and laparoscopic experience. *J Urol.* 2006; 176:1553–7.
 30. Sabbagh RA, Chatterjee SB, Chawla AB, Hoogenes JC, Kapoor AB, Matsumoto ED. Transfer of laparoscopic radical prostatectomy skills from bench model to animal model: a prospective, single-blind, randomized, controlled study. *J Urol.* 2012; 187: 1861–6.
 31. Hauge LS, Frischknecht AC, Boehler M, et al. Simulating management of patient care: how prepared are your interns to take call? Presentation: Surgical Education Week (ASE/APDS). Boston, MA, March, 2011.
 32. Pusic M, Pecaric M, Boutis K. How much practice is enough? Using learning curves to assess the deliberate practice of radiograph interpretation. Using learning curves to assess the deliberate practice of radiograph interpretation. *Acad Med.* 2011; 86:731–6.
 33. Ericsson KA. The influence of experience and deliberate practice on the development of superior expert performance. In: Ericsson KA, Charness N, Feltovich P, Hoffman RR (eds). *Cambridge Handbook of Expertise and Expert Performance.* Cambridge, UK: Cambridge University Press, 2006, pp 685–706.
 34. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med.* 2008; 15:988–94.
 35. Thurstone LL. The learning curve equation. *Psychological Monog.* 1919; 26:1–51.
 36. Ma IW, Brindle ME, Ronksley PE, Lorenzetti DL, Sauve RS, Ghali WA. Use of simulation-based education to improve outcomes of central venous catheterization: a systematic review and meta-analysis. *Acad Med.* 2011; 86:1137–47.
 37. Dong Y, Suri HS, Cook DA, et al. Simulation-based objective assessment discerns clinical proficiency in central line placement: a construct validation. *Chest.* 2010; 137:1050–6.
 38. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ.* 2003; 37:830–7.

Call for Papers

The Evidence-based Diagnostics section is seeking submissions. These manuscripts will evaluate a single emergency medicine-relevant diagnosis using a systematic review and meta-analysis to summarize high quality clinical research focusing on history, physical exam, readily available lab tests, and common imaging strategies. Evidence quality will be graded using the Quality Assessment Tool for Diagnostic Accuracy Studies. The highest quality evidence will then be summarized to report point-estimates or ranges for pre-test probability, diagnostic accuracy including interval likelihood ratios, and test-treatment thresholds for definitive tests. Authors are encouraged to contact the section editor, Christopher Carpenter, MD (carpenterc@wusm.wustl.edu) with specific questions for this series.