Pediatric Radiology (2018) 48:1406–1409
https://doi.org/10.1007/s00247-018-4193-x

MINISYMPOSIUM: EDUCATION IN RADIOLOGY

Designing and testing an educational innovation

Janet R. Reid [1] · Ryan S. Baker [2]

Received: 18 April 2018 / Revised: 24 May 2018 / Accepted: 19 June 2018 / Published online: 15 August 2018

Abstract
Technology provides an opportunity to develop innovations to improve the way we teach. Developing educational innovations must be approached in a methodical and thorough manner to identify the educational gap and validate the innovation's success in filling the gap. A successful educational innovation requires a stepwise approach that includes needs assessment, design of intervention, testing and analysis, evaluation of intervention, and determination of learning retention.

Keywords Education · Design · Evaluation · Innovation · Needs assessment · Radiology · Testing

Introduction

A wise professor once said that when you come up with a great idea, it is likely that at least five other people have independently come up with that "unique" idea — success is in the execution. One could add that success is also in the validation. Physician researchers conduct scientifically rigorous research to advance knowledge and skills and provide relevant, current and high-quality clinical care to patients. Clinician educators should use that same level of rigor to validate the methods they use to teach their trainees and attempt to link these to outcomes, as has been done in several clinical specialties [1–4]. Successful research in medical education must identify stakeholders, gaps in knowledge and skills, and novel or existing interventions that are appropriate to address the gap. So often one develops a novel technology without first determining the need and appropriateness, in effect building the hammer without first identifying the loose nail. As with clinical research, it is time for medical education research to move beyond cohort qualitative assessments toward a sound scientific process that involves developing a hypothesis, identifying separate control and experimental groups, applying the independent variable (the "intervention") and using rigorous statistical principles and tools for analysis [4]. Unfortunately, limited grant funding has led to less rigorous practice and stalls the cultivation of established medical education researchers who in turn could mentor the next generation [5]. In addition, there are challenges in studying the moving and developing population in medical education and in finding similar cohorts in similar settings to serve as "quasi-experimental" controls, but this should not discourage the medical education researcher from optimizing the research design [4]. In this review we suggest a five-step approach, common in other areas of education, toward designing and assessing a medical education research project: (1) needs assessment, (2) design of intervention, (3) testing and analysis, (4) evaluation of intervention and (5) determination of learning retention.

Needs assessment

No matter how nifty a new idea, it will fail unless a proper needs assessment is undertaken. The purpose of such an exercise is to flesh out the gaps in knowledge and skills and to identify and critique the established methods to teach these. With a strong grasp of what is currently taught and how it is taught, together with an assessment of how efficacious current approaches are, one is in an excellent position to develop focused goals and objectives for new educational approaches. These can then be rolled into an appropriate and effective intervention.

✉ Janet R. Reid
reidj@email.chop.edu

[1] Department of Radiology, Children's Hospital of Philadelphia, Perelman School of Medicine, University of Pennsylvania, 3401 Civic Center Blvd., Philadelphia, PA 19104, USA

[2] Learning Analytics Laboratory, Graduate School of Education, University of Pennsylvania, Philadelphia, PA, USA

First, one must identify key and peripheral stakeholders. For example, for an education project in radiology, the stakeholders are the trainees, and the peripheral stakeholders might include all the clinician educators and the program director. Next, one develops a list of items that should be learned and assessed. This list could be formulated through casual conversation with colleagues and students; through formal discussion with the education committee; or based on performance data on past in-service examinations, observed skills and knowledge in the workplace, complaints and concerns filed with the program director, a survey of the relevant literature, and topical issues raised through groups setting the examinations such as the American College of Radiology (ACR) in-service and American Board of Radiology (ABR) certifying examinations.

Needs assessments can be conducted in several ways, each with inherent strengths and weaknesses. These include questionnaires (electronic or paper-based), interviews, focus groups, observation and tests. Each method must be planned carefully with the method for data analysis in mind. There is extensive published work on how to craft questionnaires, going back decades, but several key principles in creating a questionnaire/survey and tests are: (1) include questions that are free from bias and one-and-a-half-handedness (e.g., "Do you believe that slide identification is important, or do you not care about effective cancer detection?); (2) avoid negatively worded questions (e.g., "Which one of the following is NOT true?"); (3) assign labels for all values on a number scale and not just the lowest and highest; (4) clearly word the questions so there is no ambiguity or undue complexity; and (5) avoid emotional or extreme language or labels such as "absolute" or "never" [6]. Survey response rates, especially for those sent electronically, are inherently low but can be increased by including a personalized cover note indicating the importance of their participation in your research, and why the research itself is important. In addition, you will improve your response rate and data retrieval by keeping the survey short and including the expected time for completion in the note. As with all research, an assurance of anonymity always improves the response rate.

Interviews produce a greater response rate only if you can get the interview! The problem with interviews is that they imply lack of anonymity, and so this must be addressed at the time of the invitation. Focus groups are an extended method to interview several people in one sitting, and as the name suggests, they must be focused on a narrow area of the research in order to be fully effective. The optimal size of a focus group is 5–8 people.

Design of intervention

A full discussion of design of the learning tool or intervention is beyond the scope of this article. However this very important step is predicated on a close examination of the results of the needs assessment. In many cases failure to properly evaluate these needs reduces the educational value or appeal of the educational tool or intervention. The design must address the content (pertaining to the knowledge and skills to be learned) and the mode of delivery of the intervention, the latter adhering closely to accepted principles of learning [7, 8]. Although both are important, failure to align with the needs assessment in either aspect of the design can lead to a poor result when the intervention is evaluated.

In modern education research, increasing evidence shows that a design effort is more likely to be effective if it is conducted by an inter-disciplinary team involving both education designers and members of the stakeholder community, for example, medical instructors who will become the first users of the resulting curriculum — a practice referred to as co-design [9].

Testing and analysis

The initial development and deployment of the educational intervention is only the beginning. Common education research and development practice calls for multiple cycles of iterative refinement and redeployment, involving repeated assessment used to determine how well the intervention is working and improve it [10]. At the conclusion of that process, an educational intervention must undergo stringent quantitative and qualitative assessment in order to be validated as a better solution to address the teaching gap. The basic principle is to apply an intervention (the independent variable) to one group and measure the response (the dependent variable) and compare it to a control group to determine significant change. Because of so many confounding variables in any educational intervention, and the inability to truly randomize exposure and timing of exposure to the intervention, it is more common for education researchers to adopt quasi-experimental methods than fully randomized approaches [4, 11]. Some key goals in setting up an educational experiment that is considered valid by the broader education research community include: (1) having a control group with characteristics similar to the experimental group, either through random assignment, random group assignment (typically referred to as a quasi-experiment) or matched-pair random assignment; and (2) having a pretest and posttest for both groups. Where this is not possible, one should consider the one-group time series where there is only one group that undergoes both the experimental and control conditions at different times (typically for different content), with a pretest and a posttest for each condition. In this case, each subject serves as her own control. Where one can expect no participants to have advance knowledge of the material, or true random assignment has occurred, it is acceptable to forego the pretest and only conduct a posttest on the two groups because the difference in this case can be attributed to the intervention.

Next, one must design the tests to best capture the learning that the intervention is hoped to produce. For this step in the study, consider the anticipated effects of the intervention on the student's acquisition of skills, knowledge, expertise and mastery, and design tasks or items that would address the essence of the hypothesis. The test items or tasks must be appropriate and relevant to the hypothesis. Additional questions could also be included in a separate section to address such things as attitudes, confidence and motivation. Knowledge and skills tests can take a variety of formats and can be written or practical/performance-based.

Some attention must be given to the timing of the tests. There should be minimal time between the pretest and the start of the experiment for both the control and intervention groups to minimize effects from confounding variables that appear during students' ongoing learning. A posttest given immediately can indicate immediate recall bias that might not reflect true retention. The test given too late is affected by confounding variables from interim learning, and perhaps the effects of studying for this later test, the latter of which is arguably not a reflection of the intervention's effect on learning at all [11]. Many researchers therefore recommend giving both immediate posttests and later retention tests [8]. Statistical analysis then typically measures the difference in scores between the control and experimental groups. The statistical method chosen should take into account the relationship between the groups, whether the variables are parametric or non-parametric, whether they are continuous or categorical, the size of the samples and the number of groups. Variables are also analyzed and described using descriptive statistics (percentages and ranges, frequencies, correlations and relationships, reliability, etc.; Table 1 [12]).

Evaluation

Studies that report on evaluations of educational interventions often include assessments of the appeal, value and potential reach of the education tool (the evaluation) in addition to the effect of the intervention on learning. One way to evaluate these aspects of learning is to provide a short questionnaire on ease of use, speed, reliability, access, likelihood that user would use it again, and potential for widespread use beyond the given institution. These responses serve as feedback to the design team for potential future revisions of the learning tool (Table 2 gives an example used in one study [13]). Ultimately, many programs of education design research also investigate implementation fidelity as an intervention scale [14] and investigate how effectively an intervention scales toward understanding how to more effectively spread future educational innovations.

Learning retention

While the goal of novel educational tools and interventions is to enhance the acquisition of knowledge and skills, it is equally important that these enhancements "stick" or have some longer-lasting impact [8]. Using established principles developed by learning scientists, the educator attempts to create a tool that is contextual, relevant, easily accessible, engaging, self-directed, and includes social interaction [7]. An intervention that is engaging and interactive and that allows immediate application to the learner's work environment enhances immediate learning [7], which will be evident on the pretest-to-posttest analysis. In most education experiments, a repeat of the posttest at 6 months accurately gauges long-term retention of knowledge and skills imparted by the intervention.

Table 1 Most commonly utilized tests for statistical analysis in education research [12]

Indications	Nonparametric	Parametric
Qualitative description	Frequencies, percentages, modes	Frequencies, mode-mean-median, standard deviation
Correlation	Spearman rho	Pearson
Best fit	Chi square	
Difference — two unrelated groups	Mann–Whitney	Student's *t*-test
Difference — two related groups	Wilcoxon	Paired *t*-test
Difference — 3+ unrelated groups	Kruskal–Wallis	Analysis of variance (ANOVA; 1- or 2-way)
Difference — 3+ related groups	Friedman	ANOVA (repeated measures)
Impact of independent on dependent variables	Ordinal or logistic regression	Linear regression

Table 2 Evaluation questionnaire for an educational intervention for pediatric ultrasound includes sections about appeal, value and reach (used with permission from [13]

1. Appeal: Rate the following statements on a scale of 1 (disagree/low) to 5 (fully agree/high)
   - _ I found the tutorial too long
   - _ I found the tutorial too short to cover the necessary information
   - _ The tutorial lasted a perfect length of time to cover necessary information
   - _ The level of difficulty was appropriate for a pediatric radiology fellow
   - _ The level of difficulty was appropriate for a radiology resident
   - _ The tutorial could be used to teach sonographers
   - _ The tutorial captured all features of the examination necessary to be able to perform the scan
   - _ The tutorial addressed all of my questions about the examination technique
   - _ The video quality for the technique was good
   - _ The video feed quality for the US images was good
   - _ The audio quality was good
   - _ The use of graphics and text is necessary to make the learning experience complete
   - _ The addition of test questions is necessary to make the learning experience complete
   - _ I would review this module for future reference
   - _ Overall grade for appeal (acceptability as a learning tool) of the tutorial method
2. Learning effectiveness: Rate the following statements on a scale of 1 (disagree/low) to 5 (fully agree/high)
   - _ I could competently perform this examination without assistance before the tutorial
   - _ I could competently perform this following completion of the tutorial
   - _ The tutorial improved my knowledge of indications for the study
   - _ The tutorial improved my knowledge of patient preparation
   - _ The tutorial improved my knowledge of patient draping and positioning
   - _ The tutorial improved my knowledge of probe selection
   - _ The tutorial improved my knowledge of image optimization
   - _ The tutorial improved my knowledge of scan planes
   - _ The tutorial gave helpful recommendations to aid in difficult scans
   - _ Overall grade for impact on learning for this tutorial
3. Reach: Rate the following statements on a scale of 1 (disagree/low) to 5 (fully agree/high)
   - _ The tutorial contains information that is only useful for our trainees
   - _ The tutorial could be helpful for trainees outside our institution in USA programs
   - _ The tutorial could be useful in international training programs
   - _ The tutorial could be useful for emergency department physicians for point-of-care diagnosis
   - _ The tutorial is only useful on one type of US machine and therefore cannot be used in all programs
   - _ I would use this only during academic time
   - _ I would use this while studying at home
   - _ I would use this at point of care at the workstation
   - _ I would review this tutorial while exercising or listen while driving
   - _ Overall grade for reach (audience size) for this tutorial

US ultrasound

Conclusion

The dynamic and multifactorial nature of learning makes it a challenging subject for quantitative research, but nonetheless a considerable amount of learning research in medicine has not fully validated the interventions created, nor followed thorough processes in designing interventions. It is important to do so to ensure that the interventions that are deployed are of maximal benefit to learners and, eventually, their patients.

Compliance with ethical standards

Conflicts of interest None

References

1. Dauphinee WD, Wood-Dauphinee S (2004) The need for evidence in medical education: the development of best evidence medical education as an opportunity to inform, guide, and sustain medical education research. Acad Med 79:925–930
2. Prystowsky JB, Bordage G, Feinglass JM (2002) Patient outcomes for segmental colon resection according to surgeon's training, certification, and experience. Surgery 132:663–672
3. Tamblyn R, Abrahamowicz M, Brailovsky C et al (1998) Association between licensing examination scores and resource use and quality of care in primary care practice. JAMA 280:989–996
4. Shea JA, Arnold L, Mann KV (2004) A RIME perspective on the quality and relevance of current and future medical education research. Acad Med 79:931–938
5. Collins J (2006) Medical education research: challenges and opportunities 1. Radiology 240:639–647
6. Cohen L, Manion L, Morrison K (2011) Chapter 13: surveys, longitudinal, cross-sectional and trend studies. In: Cohen L, Manion L, Morrison K (eds) Research methods in education, 7th edn. Routledge, New York, pp 256–289
7. Collins J (2004) Education techniques for lifelong learning: principles of adult learning. Radiographics 24:1483–1489
8. Koedinger KR, Corbett AT, Perfetti C (2012) The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. Cogn Sci 36:757–798
9. Roschelle J, Penuel WR (2006) Co-design of innovations with teachers: definition and dynamics. In: Proceedings of the 7th International Conference on Learning Sciences, pp 606–612
10. Koedinger KR, Sueker EL (1996) PAT goes to college: evaluating a cognitive tutor for developmental mathematics. In: Proceedings of the 1996 International Conference on Learning Sciences, pp 180–187
11. Cohen L, Manion L, Morrison K (2011) Chapter 16: experiments, quasi-experiments, single-case research and internet-based experiments. In: Cohen L, Manion L, Morrison K (eds) Research methods in education, 7th edn. Routledge, New York, pp 312–334
12. Cohen L, Manion L, Morrison K (2011) Chapter 38: choosing a statistical test. In: Cohen L, Manion L, Morrison K (eds) Research methods in education, 7th edn. Routledge, New York, pp 697–704
13. Back SJ, Darge K, Bedoya MA et al (2016) Ultrasound tutorials in under 10 minutes: experience and results. AJR Am J Roentgenol 207:653–660
14. Feng M, Roschelle R, Murphy R et al (2014) Using analytics for improving implementation fidelity in a large scale efficacy trial. In: Proceedings of the International Society of the Learning Sciences, pp 527–534