The Value Of The Post-Hoc Analysis Of Responses To Questions Asked By Audience Response Technology

P Schick, S Abramson, E Gnall, S Fine, J Burke

Citation

P Schick, S Abramson, E Gnall, S Fine, J Burke. *The Value Of The Post-Hoc Analysis Of Responses To Questions Asked By Audience Response Technology*. The Internet Journal of Medical Education. 2012 Volume 2 Number 2.

Abstract

Context and background: There is a need to monitor the effectiveness of medical education. To meet this challenge, we pioneered an unappreciated application of audience response technology; the post-hoc analysis of answers to questions asked during lectures to monitor trainee knowledge and progress. . Procedure and results: Audience response technology is widely used to poll knowledge and opinions that engage audiences and promotes interactivity. Post-hoc analysis was feasible since the technology stores all answers (an underappreciated function of the technology). We used this approach in our cardiology fellowship conferences. The following information was obtained: 1) Baseline knowledge. This information helped us focus on poorly understood topics. 2) Two fellows had inadequate grades and required intervention. 3) Eleven questions were found to be ambiguous and were not graded. 4) We developed a strategy to assess the comprehension of essential core knowledge and ascertained that the fellows had a solid understanding of milestones. Impact and significance: This information was valuable for optimizing our training program. Fellow feedback: fellows strongly felt that audience response technology promoted interactivity and had educational value. In addition, grading by the technology was not intrusive, distracting or intimidating. Therefore, In comparison to stand-alone testing, monitoring knowledge by this technology was: 1) easier to implement, grade and evaluate; 2) more comprehensive and sensitive since all questions asked in the lecture series were assessed. The strategy was designed to meet the needs of our training program but could be adapted to the needs of any medical curriculum and training at all levels of medical education

INTRODUCTION – BACKGROUND

There is a need to improve the monitoring of the effectiveness of medical education programs. In order to meet this challenge, we pioneered an unappreciated application of audience response technology (ART); the post-hoc analysis of audience responses to monitor knowledge.

ART is being used in numerous medical centers and hospitals (1-6). Using this technology, medical students or trainees use remote control clickers to choose answers to questions displayed on PowerPoint slides. The group's answers are instantly displayed, and this engages participants and promotes interactivity. This approach can improve educational programs and address the limitations of the passive nature of conventional lectures. It is generally believed that comprehension and retention of new material is improved if learners are actively engaged in the learning process.

ART stores all responses to questions. Therefore, it is possible to mine the stored data and carry out post-hoc analysis to monitor participant knowledge. This is an underappreciated feature of ART. An exhaustive literature search could not find any published information about the post-hoc analysis of ART responses. In a pilot study, we used this approach in two hematology lectures. The study provided useful demographic information about differences in knowledge in medical students and residents at different levels of training. (7).

We would like to report our experience using the post-hoc analysis of audience responses in a series of lectures in our cardiology fellowship program as an example of the value of this approach. The study demonstrates that post-hoc analysis of audience answers can meet the challenge of sensitively monitoring group (aggregate) and individual knowledge and progress. We have found that this educational approach is objective, easy to implement and grade, and has been enthusiastically accepted by trainees.

The program was designed to meet the needs of our training program. However, this technology can be tailored to the needs of any medical curriculum and training programs.

Therefore, it would be of universal value at all levels of medical education.

METHODS

Audience response technology (Turning Point Technologies®, Youngstown, Ohio): Multiple choice questions were asked in Microsoft PowerPoint® slides. Trainees used remote control clickers to relay their answers to a sensor in the lecturer's laptop computer. Software in the laptop instantly processed the responses and displayed a graph of the group's answers. Anonymity was maintained during lectures since fellows did not know how their peers had answered. In addition, all answers were recorded and subjected to post-hoc analysis as described below.

Subjects: Twelve cardiology fellows participated in the study of ART during our lecture series on cardiovascular disease. A total of 34 subjects were presented in 7 blocks over a period of 8 months. Each subject was covered in two lectures, an introductory lecture and a review lecture that was given about a week later.

Application of the technology at a pivotal time point: ART was used only in the review lecture (the 2 nd lecture in the series). This time point was chosen because it was considered pivotal, since fellows still had ample time to study poorly understood topics prior to the end of the academic block. Preceptors had the option of providing additional support for topics that proved to be difficult.

Post-hoc analysis: All responses to questions were recorded by Turning Technology software®. The software enabled the transfer of raw data onto Microsoft Excel® spreadsheets for post-hoc analysis. Post-hoc analysis of the data was used to determine the following:

1) Identification of lectures in which fellows achieved outstanding and poor grades: Well understood and poorly understood subjects were identified by mean grades that were one standard deviation above or below the mean grade of all lectures.

2) The evaluation of fellow grades: We compared fellow grades to determine whether any trainee needed additional help. Remediation was implemented for fellows whose grades were one standard deviation below the mean grade achieved by all fellows.

3) Identification of unfair and poorly designed questions: The faculty evaluated individual questions that fewer than 33% of fellows answered correctly. The purpose was to determine whether any of these questions were unfair due to inappropriate distracters, improper wording or other errors. If deemed unfair, the questions were not graded.

4) Use of post-hoc analysis to determine the comprehension of essential core knowledge (milestones). The fellowship director and two other academic cardiologists involved in the lectures identified questions about essential core information. Grades in these questions were recorded to determine comprehension of milestones.

Fellow feedback about the utility and value of ART: We obtained anonymous feedback by using a Likert questionnaire after the completion of the 4 th block (about 4 months into the lecture series). The questions asked are shown in Table 1.

Study was IRB exempt when reviewed by the Lankenau Medical Center IRB (Ethical Approval).

Figure 1

Table 1: Feedback from fellows was obtained after the 4 block of lectures. There were a total of 7 blocks in the lecture series.

Table 1: Feedback (n = 12) **1 – Highest rating (Outstanding): 5 – Lowest rating**	**Mean**	**SD**
The conferences provided new information.	1.33	0.48
Audience response systems help make lectures more interactive	1.41	0.47
The lecturers adequately explained the rationale for correct and incorrect answers	1.42	0.69
The conferences were appropriate for my level of training	1.25	0.50
I was comfortable with the monitoring of my answers to questions by audience response systems	1.61	0.75

RESULTS

Figure 2

Table 2: Outstanding and Poor Grades. The mean grade in all questions was 72.8% ± 8.0%. (Mean ± SD)

The instant display of responses engaged the fellows and promoted interactivity. Lecturers seeing the real-time display of comprehension could spend more time on subjects shown to be poorly comprehended. **Table 2: Outstanding and Poor Grades**	
Outstanding grade (mean grade of all lectures plus one SD) 81% or more	**Poor grade (mean grade of all lectures minus one SD) 65% or less**
Cardiovascular physiology (84%)	Physical Exam (65%)
Revascularization (82%)	Atrial septal defect (58%)
Mitral stenosis (83%)	Ventricular septal defect (65%)
Pericarditis (81%)	Other congenital heart disorders (Lecture 1) (56%)
Tetralogy of Fallot (84%)	Other congenital heart disorders (Lecture 2) (52%)

Real time information derived from audience response technology during the lecture was available to both fellows and lecturers.

The instant display of responses engaged the fellows and promoted interactivity. Lecturers seeing the real-time display

of their comprehension could spend more time on subjects shown to be poorly comprehended..

Information derived by post-hoc analysis of fellows' answers:

1) Identification of lectures in which fellows achieved outstanding and poor grades: The average grade of audience responses in 34 lectures was 72.8% + 8.0% (mean and standard deviation). Five lectures in which fellows received outstanding grades (=> 81%) were identified (Table 2). These were lectures on cardiovascular physiology, revascularization, mitral stenosis, pericarditis and Tetralogy of Fallot. Five lectures in which fellows received poor grades (=<65%) were atrial septal defects, ventricular septal defects, physical examination, and 2 lectures on miscellaneous congenital heart defects.

2) The evaluation of fellow grades: The data revealed that two fellows had grades that were one standard deviation below the mean grade of all fellows. This provided the faculty with an objective tool for the assessment of knowledge at a pivotal time point and to implement remediation and faculty intervention.

Identification of unfair and poorly designed questions:

3) The evaluation of questions: Eleven questions were found to be ambiguous due to inappropriate distracters, confusing wording or other errors. These questions were not graded.

4) Comprehension of essential core information (milestones): Fellows achieved a grade of 83.1 ± 5.3% (mean ± SD) (8).

Fellow feedback about the utility and value of our application of ART in the cardiovascular lecture series: These data are shown in Table 1. A score of 1 was the highest rating and 5 was the lowest rating. The fellows gave outstanding feedback to queries about whether new information was presented, whether audience response systems promoted interactivity, whether lecturers provided adequate rationale for correct and incorrect answers, and whether the lectures were appropriate for their level of training. Responses to the question "I was comfortable with the monitoring of my answers to questions by audience response systems" indicated that that grading by audience response systems was not distracting or intimidating.

DISCUSSION

The application of ART to our cardiovascular lecture series engaged participants who felt that the technology improved the lectures and promoted interactivity. The instant display of fellow knowledge also helped lecturers adjust their lectures by spending more time emphasizing poorly understood topics (4).

However, the main purpose of the current study was to present the value and the strategy of the post-hoc analysis of responses to questions asked by ART. The specific information about our training program was relevant to our institution only since all training programs differ in organization and strengths and weaknesses.

We were able to ask 1,293 questions in the lecture series using ART without sacrificing time spent on discussing the topics. There were several reasons why we were able to ask a large number of questions. The length of many lectures was greater than an hour, thus providing more time for questions. There were many questions on identification of abnormalities in ECGs, echocardiograms and pathology. These subjects were covered in the introductory lectures and fellows did not need much time to answer. The lecturer could still spend time discussing slides that fellows had misinterpreted, but residents correctly answered most of the slides and hence the lecturer was able to keep a fast pace.

There are several advantages of monitoring knowledge by ART by data-mining or post-hoc analysis of ART responses instead of relying on standard standalone examinations. Post-hoc evaluation was very comprehensive. The questions covered all points and topics presented in the lectures. Therefore, we were able to detect subtle differences in knowledge. Asking as many as 1,293 questions would not be feasible by standard testing. It was easier to implement testing, to grade, and to analyze fellows' answers to questions by using audience response systems than by standard stand-alone examinations. When using ART, answers to questions were graded by a software program. Table 2 reveals that fellows thought that ART improved interactivity in lectures and that the technology was not intimidating.

Since the data about responses to questions asked in the lectures were transferred to MS Excel spreadsheets, we analyzed the data for information that we considered important to help monitor and improve our cardiovascular lecture series. However, other training programs might need additional or different types of information. Each training

institution can tailor the post-hoc analysis to their needs to optimize their educational programs.

Outstanding and poor lectures: Our criteria for identifying outstanding and poor lectures revealed that fellows had a poor understanding of congenital heart disease. There are several possibilities for the poor grades: lack of exposure to the topics, lack of interest, weak lectures, difficulty of the subject, and the fact that our hospital was not a center for these patients. Should the lack of exposure to congenital disorders be responsible, we could encourage rotations to sites that care for congenital heart disease and place more emphasis on these disorders in our training program. Should lectures be shown to be inadequate, we could encourage faculty development by improving lecture skills. Conversely, we found that fellows' grades were extremely high in subjects such as pericarditis. With this information, we could modify fellow experience and education to raise their knowledge to the level of that in pericarditis. This approach can help strengthen lectures in which fellows demonstrated poor performance.

Fellow performance: The data revealed that two fellows had grades that were one standard deviation below the mean grade of all fellows. They were given directed teaching and specific focused reading material. Their progress was monitored by oral exams. One of the fellows dramatically improved his performance. However, the other did not demonstrate improvement, and further faculty intervention was necessary. Therefore, the evaluation of individual fellows could identify and help fellows who are struggling to understand topics presented in the lectures. The intervention enabled these fellows to do well on final written in-service exams.

Identification of unfair questions: Eleven questions were not graded since post-hoc evaluation revealed that they were ambiguous or unfair. These questions were corrected prior to used in next year's curriculum

The comprehension of essential core knowledge: The assessment of the grades in questions about milestones revealed that most of the fellows had a solid comprehension of essential core knowledge and that our training program was effective (8). The assessment of grades also provides the program director the ability to identify fellows who are doing poorly at an earlier stage so that intervention is feasible.

Testing knowledge using ART at this pivotal point had several merits. The documentation of fellow knowledge at the time point provided ample time to address deficiencies in knowledge prior to the end of the academic block. Fellows became aware of their strengths and weaknesses, and this encouraged self-study. Faculty awareness of fellow performance at this pivotal point permitted faculty intervention and remediation.

The cost of an ART license, software, and clickers would range from $1,500 to $3,000. However, most institutions have purchased ART and clickers. There would be no other costs except for an administrator or secretary to help carry out the analysis that would require 10 minutes per lecture.

The implementation of ART post-hoc analysis is feasible. Many institutions are using ART in the classroom without difficulty. Most good teachers have developed skills in designed questions and using PowerPoint to give lectures. Incorporating questions into ART enabled- PowerPoint lectures is an easy task. Successful post-hoc analysis is dependent on a designing a strategy for analysis and selecting and training an individual to perform the analysis using ART software and MS Excel. One of our administrators had no difficulty in learning these tasks and could analyze a lecture within 10 minutes. Conclusions:

Post-hoc analysis of the data provides detailed information of participant knowledge, progress and the mastery of milestones. This can enable program directors to be sensitive to strengths and weaknesses of training programs, identify a below average fellow performance, and to institute measures to improve education and the training program. Faculty development can be implemented when poor lecturers are identified and encouraged to improve. There are many other options for post-hoc analysis of ART. For example, if the audience is composed of medical students, residents and fellows, post-hoc analysis can reveal demographic differences in knowledge (7). This information can be used to tailor lectures and education programs to the demographic composition of the audience. In comparison to stand-alone standard testing, monitoring knowledge by ART was: 1) easier to implement, grade and evaluate and; 2) more comprehensive and sensitive since all questions asked in the entire lecture series could be assessed.

The specific information about our lecture series is relevant only for our fellowship program since the strengths and weaknesses and structure of each training program are unique. We have presented our data derived from ART as an

example of the value of post-hoc analysis. However, the value and strategy is applicable to all educational programs and can be used at all levels of medical training and in all specialties of medicine.

A 500 word overview of the study has been published in Medical Education (Really Good Stuff) 45:1157-1158, 2011

References

1. Caldwell JE. Clickers in the large classroom: current research and best-practice tips. [Review], CBE Life Sciences Education [Electronic Resource]: 2007: 6:9-20 (Accessed 9 December 2010).
2. Collins LJ. Livening up the classroom: using audience response systems to promote active learning. Medical Reference Services Quarterly. 2007: 26: 81-88
3. Johnson JT. Creating learner-centered classrooms: use of an audience response system in pediatric dentistry education. Journal of Dental Education. 2005: 69: 378-81.
4. Kerns C. Technological tools for library user education: one library's experience. Medical Reference Services Quarterly 2007: 26:105-114
5. Streeter JL, Rybicki FJ. A novel standard-compliant audience response system for medical education. Radiographics. 2006: 26: 1243-9
6. Audience response systems in higher education: Application and cases. DA Banks, editor. Information Science Publishing, 2006
7. Schick PK, Burke MD. Post-hoc Analysis of Audience Responses-Enabled Conferences on Hematological Subjects. ISPUB Internet Journal of Hematology. 2009: 6:1-7. (Accessed 9 December 2010)
8. Schick PK, S. Abramson, JF Burke. Medical Education (Really Good Stuff). 45:1157-1158, 2011.

Author Information

Paul K. Schick, M.D.
Research Professor of Medicine, Drexel University College of Medicine

Sandra Abramson, M.D.
Thomas Jefferson Medical College

Eric Gnall, D.O.
Clinical Instructor of Medicine, Thomas Jefferson Medical College

Sandra Fine, M.B.A.
Vice-President of Medical Education, Lankenau Medical Center

James F. Burke, M.D.
Clinical Associate Professor of Medicine, Thomas Jefferson Medical College