

# The impact of an audience response system on evaluation results in medical lectures, a controlled study.

Mathis Wegner (✉ [Mathis.wegner@uksh.de](mailto:Mathis.wegner@uksh.de))

University Medical Center of Schleswig-Holstein

**Thorben Michaelis**

University Medical Center of Schleswig-Holstein

**Ibrahim Alkatout**

University Medical Center of Schleswig-Holstein

**Sebastian Lippross**

University Medical Center of Schleswig-Holstein

---

## Research Article

**Keywords:** Audience Response System, medical education, evaluation, lecture

**Posted Date:** May 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-500742/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Audience Response Systems (ARS) are well perceived by students and lecturers and represent a valuable instrument of quality assurance and student participation in academic teaching. Our controlled study investigated whether ARS in evaluation of medical lectures has an impact on evaluation results, enhances student learning and motivation.

**Methods:** The study was conducted at the University medical center of Schleswig-Holstein, Campus Kiel, over the course of one year. 198 students who participated voluntarily were included. In alternating lectures a paper questionnaire and ARS were offered as an evaluation tool to the audience. In both processes two items concerning teacher's behavior, three items concerning lecture's design and two items concerning lecture's content were investigated by using a seven-item questionnaire with a five-point Likert scale, two multiple-choice questions were posed at the end of each evaluation to assess knowledge transfer.

**Results:** 128 students took part using the traditional paper and pencil format, 70 students agreed to use ARS. Participants not using ARS rated the quality of teaching in both items and the structure of the lecture in one item superior to participants using ARS ( $p < 0.001$ ,  $p < 0.003$ ,  $p < 0.001$ ). The screening for participation in the additionally asked MC- questions showed a higher rate of participation in students using ARS.

**Conclusion:** The results of our study show significant differences in lectures' evaluation using ARS. Students' motivation to evaluate lectures might increase by using an audience response device.

## Introduction

The quality of teaching and thus evaluation of lectures and lecturers have increasingly drawn public's attention. Due to financial influences in education and various ranking lists of universities, quality of teaching is constantly observed in order to assess possible deficiencies.

Students' evaluation of lectures and the lecturer is used as a tool to observe teaching quality continuously with adequate effort. Based on evaluation results, salaries and financial subsidies for educational purposes are distributed [1]. Therefore, a modern challenge of medical education is to create innovative concepts to improve interaction in lectures and to represent a contemporary way of teaching.

Audience Response Systems (ARS) are resources to individually and simultaneously involve the audience in lectures. ARS allows the lecturer to interact with the audience as part of the lecture and to directly receive feedback. One of the main goals of ARS is to increase and improve interaction between the lecturer and the audience [2].

Due to the fact that most students own a smartphone (95% in the age group of 14–29 years) nowadays [3], the incorporation in lectures using browser-based ARS is easier and does not require additional

devices. There are several open-source applications that can be used for academic teaching [4].

In previous investigations the use of ARS is highly valued among students, providing a better environment for asking questions and receiving answers. It helps students to actively participate in the teaching process due to a low inhibition threshold [4–7]. The anonymity of most ARS is granted by a web-based open platform, suggesting security. As a result, ARS especially improves activation and enhances learning in students who feel uncomfortable answering questions in front of others or might rate their questions as inappropriate.

An effect of ARS use in lectures is the increase of student motivation and the generation of a learning friendly environment [8]. The improvement of assessment grades remains contradictory [4, 6, 9]. Either way students claim that using ARS enhances their understanding of the lecture. They conclude that ARS can be used as an efficient feedback tool. Lecturers can alter their teaching methods depending on direct feedback provided by the use of ARS. They receive a direct measure of their intelligibility. This makes ARS a valuable instrument of quality measurement in academic teaching [10].

Audience response technology is well perceived by students and more lecturers start to rely on new and smart ways to receive evaluations for their lectures [4–8]. Students' evaluation is established at medical schools, traditionally using a questionnaire on a handout directly after lecture. It seems to be a reliable measure when it is focused on individual courses. Furthermore, the lecturer receives a direct feedback [11, 12]. Results show that global ratings of teachers are mainly based on teaching behavior and teachers' knowledge, but there are additional effects that influence students' evaluation [13]. Expected grade, actual grade, course level, class size, course timing, student gender and course subject significantly affect student evaluation of teaching as biasing variables [14]. Investigations if ARS use has an influence on results of evaluation has shown the impact has been favorable for speakers and evaluation of lectures [15]. A main difficulty of evaluation in general is participation and the response rate. In voluntary lectures, many students leave the plenum directly after the end of the given presentation without evaluating or giving a feedback. ARS might be an interactive way to improve student's response rates [16].

While the existing literature suggests a positive perception of this relatively new technology few direct comparisons of traditional paper-pencil questionnaires and ARS have been published to our knowledge.

Our controlled study investigated whether ARS in evaluation of medical lectures has an impact on evaluation results. In addition, we investigated if ARS use in medical lectures can enhance student learning and motivation.

## **Material And Methods**

### **Study setting and design**

The study was conducted at the University medical center of Schleswig-Holstein, Campus Kiel as part of lectures of the curricular teaching program in orthopedics. In these lectures common orthopedic

conditions are presented by different lecturers through the course of two semesters. Five lectures per semester with varied content and different topics were included in our study. During the first semester of our investigation lectures' evaluation was handed out in paper and pencil (P&P; group 1). During the second semester evaluation was executed with ARS use (ARS nova; group 2). Both groups had to answer the same questionnaire to evaluate the lecture. Additionally, two multiple-choice questions concerning the lecture's content were posed to verify students' improvements and students' learning motivation. In our analysis we compared students' statements of both groups. We also screened the multiple-choice questions for participation as a separate parameter. Closing the questionnaire students had the opportunity to give a written feedback to the lecturer.

The lecturers were highly qualified experts in each topic, in group 1 the questionnaires were handed out by a medical student after each lecture. In group 2 the lecturer asked students to evaluate the lecture using an ARS. The designated URL was presented to students at the end of each lecture.

### **Sample definition and recruitment of participants**

Each semester lecture in orthopedics was part of the student curriculum. The attendance was voluntary. All of our students were invited to take part in our research study. Participation was voluntary and had no impact on students' grades. The questionnaire for evaluation was completely anonymous. Informed consent was obtained from all participants prior to the study.

### **Audience Response Technology**

Several ARS systems were evaluated by the authors and the web-based application ARSnova (TransMIT – Gesellschaft für Technologietransfer mbH, Projektbereich für mobile Anwendungen; [arsnova.eu/mobile](http://arsnova.eu/mobile)) was selected. Peer review, live feedback and live assessment, the inverted classroom are functions offered by ARSnova. This ARS system allows participants to compute responses via mobile phone or respond online using a designated URL. Students required a device to go online. No personal registration is necessary.

### **Questionnaire**

The evaluation instrument was designed to measure students' perception of lecture quality comparing ARS and paper and pencil procedure (Table 1). The dependent variables of the seven-item scale refer to teaching quality (Items one and two; The lecturer was well prepared, the lecturer seemed competent), lecture's quality (Items three to five; The amount of content was appropriate, the presentation was understandable, the lecture was structured logically.) and lecture's content (Items six and seven; The lecture had practical relevance, the lecture was entertaining. Each item was rated using a 5-point Likert scale (4 = strongly agree; 3 = agree; 2 = neutral; 1 = disagree; 0 = totally disagree) [17]. In addition, the questionnaire included two multiple-choice questions (MC-questions) to evaluate the lecture's content. Students were given the opportunity to give the lecturer a written feedback at the end of each questionnaire.

### **Sample size estimation**

The sample size of 198 students participating is comparable to other related studies [4, 5, 7]. 128 of our participants used P&P to evaluate, 70 of our students used ARS nova.

### Statistical procedure

We used descriptive statistics to assess students' evaluations (mean, SD, associated P). Each of the means of group 1 and group 2 measure were tested for statistical significance using an unpaired t-test. Significance was assumed when  $p \leq 0.05$ . The multiple-choice questions were screened for participation.

## Results

A total of 198 students through the course of two semesters were included. A mean of  $25.6 \pm 9.18$  students evaluated the lecture by using a P&P questionnaire, a mean of  $14 \pm 3.54$  students evaluated using ARS nova.

### Descriptive statistics

Means and standard deviations for each of the items (1–7) are reported in Table 1, Fig. 1.

Table 1  
Means, standard deviation and p-values for items 1–7. Comparison for significance of P&P and ARSnova.

| Table 1                               | P&P (n = 128) |      | ARSnova (n = 70) |      | p-value | significance |
|---------------------------------------|---------------|------|------------------|------|---------|--------------|
|                                       | mean          | SD   | mean (SD)        | SD   |         |              |
| The lecturer was well prepared        | 3.76          | 0.24 | 3.56             | 0.67 | 0.003   | yes          |
| The lecturer seemed competent         | 3.89          | 0.11 | 3.47             | 0.73 | < 0.001 | yes          |
| The amount of content was appropriate | 3.08          | 0.92 | 3.13             | 0.93 | 0.716   | n.s.         |
| The presentation was understandable   | 3.23          | 0.77 | 3.20             | 0.89 | 0,805   | n.s.         |
| The lecture was structured logically  | 3.35          | 0.65 | 2.81             | 1.09 | < 0.001 | yes          |
| The lecture had practical relevance   | 3.15          | 0.85 | 2.93             | 1.04 | 0.11    | n.s.         |
| The lecture was entertaining          | 2.78          | 1.22 | 2.61             | 1.18 | 0.344   | n.s.         |

The P&P group rated item 1 significantly higher than the ARS nova group (P&P: mean of 3.76 (SD 0.24), ARS nova: mean of 3.56 (SD 0.67),  $p < 0.001$ ). A comparable effect can be seen in item 2 The P&P group shows a significant higher rating compared to the ARS nova group (P&P: mean of 3.89 (SD 0.11), ARS nova: mean of 3,47 (SD 0.73),  $p < 0.001$ ). The evaluation of lecture's quality shows no difference in items

3 (P&P: mean of 3.08 (SD 0.92), ARS nova: mean of 3.13 (SD 0.93),  $p = 0.716$ ) and Item 4 (P&P: mean of 3.23 (SD 0.77), ARS nova: 3.2 (SD 0.89),  $p = 0.805$ ). Item 5 was rated significantly higher by the P&P group (P&P: mean of 3.35 (SD 0.65); ARS nova: mean of 2.81 (SD 1.09);  $p < 0.001$ ). Concerning the lecture's content, no significant differences in perception can be observed (item 6: P&P: mean of 3.15 (SD 0.85), ARS nova: mean of 2.93 (SD 1.04),  $p = 0.11$ ; item 7: P&P: mean of 2.78 (SD 1,22), ARS nova: mean of 2.61 (SD 1.18),  $p = 0.344$ ).

### Evaluation Results Regarding The Differences In Evaluation Using Ars

**Figure 1** Evaluation results regarding the differences in evaluation using ARS. In a lecture series of ten lectures 198 students of medicine evaluated the lecture and the lecturer using 7 items in a 5-point-likert-scale (from 0 „totally disagree“ to 4 „totally agree“). Shown are the means of the items, repectively SD. Statistically significant differences were described in the evaluation of Item 1 ( $p < 0.003$ ), Item 2 ( $p < 0.001$ ) and item 5 ( $p < 0.001$ ), marked with \*.

The test for participation in the additionally asked MC- questions showed that 70% of group 1 and 94% of group 2 answered the multiple-choice questions (Fig. 2).

### Screening For Participation

Students' written feedback is exemplarily shown in Table 2.

Table 2  
Exemplarily shown feedback comments of students.

|   |
|---|
| <b>"Maybe it would be a possibility to number the slides."</b>          |
| "Too much material in too short a time - could not follow in parts."    |
| "A script to follow up would be nice."                                  |
| "It was too fast to follow in writing."                                 |
| "Sometimes you need more time to think about the situation, the views." |
| "Very good presentation, very entertaining."                            |

## Discussion

The aim of the study was to question whether or not ARS has a special impact on student's perception of lecture quality. Evaluating lectures is an established instrument of quality assurance at medical schools [10]. It is an adequate method to continuously observe teaching quality almost effortless. Financial subsidies for each facility are distributed based upon students' evaluation. Therefore, lectures aim to integrate and improve students' incorporation and motivation in contemporary ways.

The traditional procedure is to hand out paper and pencil questionnaires after lectures to receive student's feedback. Different biasing variables affecting results of lecture's evaluation exist [14, 18]. Our aim was to evaluate the impact of ARS compared to the traditional way of evaluation. ARS is an innovative way to give feedback to teachers or lecturers [2, 15]. It is rated positively by students and lecturers and represents an efficient feedback tool [4–8]. Referring to our results, statistically significant differences could be demonstrated in evaluation of teaching quality and the lectures' structure. The implementation of the ARS use has an effect on students' evaluation habits. The ratings regarding the lecture's quality decreases significantly using ARS. This finding can have multiple reasons. Schmidt et. al. [4] report that students appreciate the anonymity of the ARS. Thus, a possible bias is represented in social constraints or social desirability when using paper and pencil [19]. A reason for a less favorable evaluation of lecture quality and content might be the generation of anonymity in using ARS. In 2003, Miller et. al. [15] found that lectures and speaker quality evaluated by ARS was rated more positively than when using traditional tools. These findings are in contrast to our study but might refer to the positive impact of technology tools at that time. The establishment of online evaluation portals in medical education might have changed people's perception of technology nowadays. The change of evaluation habits using ARS seems to be a relevant topic in evaluation research. Both studies, Miller et. al. [15] and our study, indicate differences in using different evaluation tools. Hence both results seem to stand in line - use of ARS can have a major impact on evaluation.

Another aim of our study was to question whether or not ARS can enhance learning and motivation. The MC-questions served to indicate the motivation to answer the additional questions. In the ARS-group 94% completed the questionnaire compared to 70% in the P&P-group (Fig. 2). These findings are consistent with Stoneking et. al. [16] who reported an increase in completion rate for lecture evaluation using ARS. P&P evaluation is traditionally executed at the end of each lecture. Raising the possibility that time limitations may have caused incomplete participation because students tend to leave punctually. In contrast, our data indicates that ARS use in evaluation may represent a solution for incomplete response rates due to temporal limitations. This is in keeping with previous study results indicating that ARS increases students' motivation by creating a learning friendly environment, increase completion rates [4, 8, 16].

The quality of answers was not included in our study. Whether an increase in responses translates to a higher quality of answers remains unknown [4, 6, 9, 15, 16].

Our study has several limitations. First of all, not all the students evaluated the lecture since it was voluntary. The individual response rate in completing the questionnaire was slightly inconsistent. Another limitation can be considered due to the fact that this study was obtained at an orthopedic lecture and may not be generalized for other medical education settings.

## Conclusion

Our data refer to differences in students' perception of lecture quality in medical education settings. The evaluation strategy using ARS differs from the traditional way using paper and pencil. Our investigation also points out that students' motivation to evaluate lectures might increase by using an audience response device. These findings stand in line with prior studies.

## Abbreviations

ARS – Audience Response System

ART - Audience Response Technology

P&P – Paper and pencil

MC – Multiple-choice

SD – Standard deviation

## Declarations

### *Ethics approval and informed consent*

The study was conducted in accordance with the WMA Declaration of Helsinki. The study protocol was approved by the local ethics committee of the medical faculty of the university of Kiel (project number D512/21). Informed consent was obtained by all participants involved in the study.

### *Consent for publication*

Not applicable.

### *Availability of data and materials*

Data and materials are available from the corresponding author on reasonable request.

### *Competing interests*

The authors declare that they have no competing interests.

### *Funding*

Not applicable.

### *Authors contributions*

TM and SL designed and conducted the study. MW did the data evaluation and wrote the manuscript. IA generated questions and supervised the process. All authors have read and approved the final

manuscript.

### *Acknowledgements*

Not applicable.

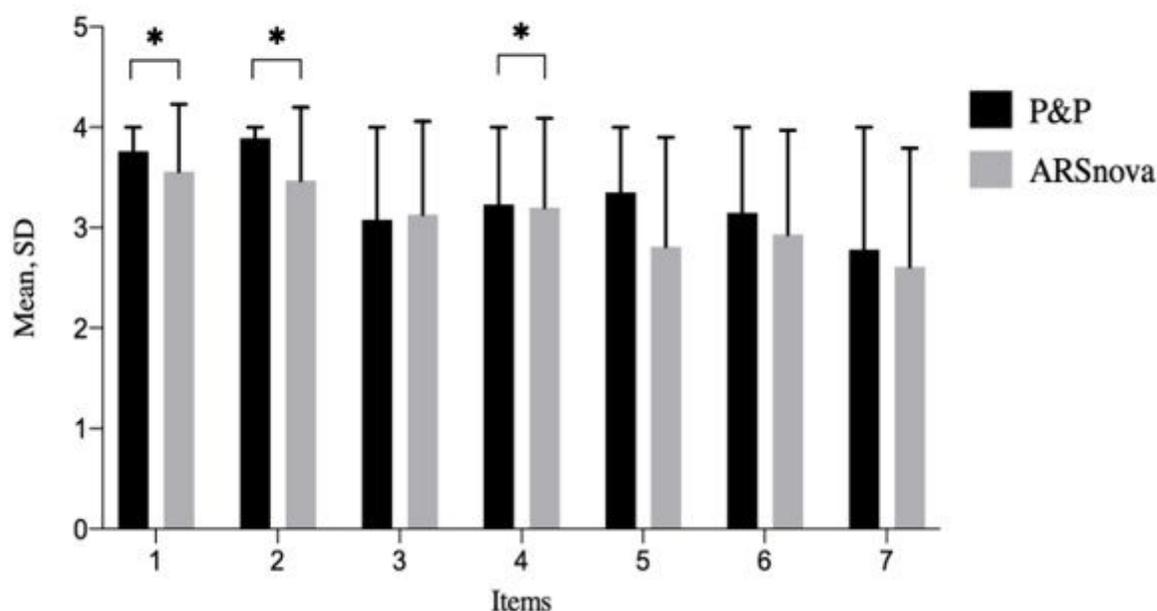
## **References**

1. Pabst R, Rothkötter HJ, Nave H, Tschernig T. Medizinstudium: Lehrevaluation in der Medizin. Dtsch Arztebl 2001; 98(12): A-747/ B-630/ C-598.
2. Kenwright K. Clickers in the classroom. TechTrends. 2009;5 3(1), 74-77.
3. <https://de.statista.com/statistik/daten/studie/459963/umfrage/anteil-der-smartphone-nutzer-in-deutschland-nach-altersgruppe/>. Accessed 26 Feb 2020.
4. Schmidt T, Gazou A, Rieß A, et al. The impact of an audience response system on a summative assessment, a controlled field study. BMC Med Educ. 2020; 20, 218.
5. Guarascio AJ, Nemecek BD, Zimmerman DE. Evaluation of students' perceptions of the Socrative application versus a traditional student response system and its impact on classroom engagement. Curr Pharm Teach Learn. 2017; 9(5): 808–812.
6. Datta R, Datta K, Venkatesh MD. Evaluation of interactive teaching for undergraduate medical students using a classroom interactive response system in India. Med J Armed Forces India. 2015; 71(3): 239–245.
7. MacGeorge EL, Homan SR, Dunning JB Jr, et al. Student evaluation of audience response technology in large lecture classes. Educ Technol Res Dev. 2008; 56:125–145.
8. Barbour ME. Electronic voting in dental materials education: the impact on students' attitudes and exam performance. J Dent Educ. 2008; 72(9):1042–1047.
9. Atlantis E, Cheema, Birinder S. Effect of audience response system technology on learning outcomes in health students and professionals, International Journal of Evidence-Based Healthcare. 2015; 13(1): 3-8.
10. Gröblinger O, Kopp M, Hoffmann B. Audience Response Systems as an instrument of quality assurance in academic teaching. In Proceedings of the 10th Annual International Technology, Education and Development Conference. 2016. Valencia, 3473–3482.
11. Diehl JM. Student Evaluations of Lectures and Seminars: Norms for two Recently Developed Questionnaires. Psychologie in Erziehung und Unterricht. 2003; 50(1), 27–42.
12. Bode SF, Straub C, Giesler M, Biller S, Forster J, Krüger M. Audience-response systems for evaluation of pediatric lectures—comparison with a classic end-of-term online-based evaluation. GMS Z Med Ausbild. 2015;32(2).
13. Blackhart GC, Peruche BM, DeWall CN, Joiner TE. Factors influencing teaching evaluations in higher education. Teaching of Psychology. 2006; 33(1): 37–39.

14. Badri MA, Abdulla M, Kamali MA, Dodeen H. Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE, [International Journal of Educational Management](#). 2006; 20(1): 43-59.
15. Miller RG, Ashar BH, Getz KJ. Evaluation of an audience response system for the continuing education of health professionals. *J Contin Educ Health Prof*. 2003; 23(2):109-15.
16. Stoneking LR, Grall KH, Min A, Dreifuss B, Spear Ellinwood KC. Role of an audience response system in didactic attendance and assessment. *J Grad Med Educ*. 2014; 6(2):335-337.
17. Likert R. A technique for the measurement of attitudes. *Arch Psychol*. 1932; 22:140–155.
18. Marsh HW. Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *J Educ Psychol*. 1984; 76(5): 707–754.
19. Ross LD, Amabile TM, Steinmetz J. [Social Roles, Social Control and Biases in Social Perception Processes](#). *Journal of Personality and Social Psychology*. 1977; 35 (7):485–494

## Figures

### Evaluation results regarding the differences in evaluation using ARS



**Figure 1**

Evaluation results regarding the differences in evaluation using ARS. In a lecture series of ten lectures 198 students of medicine evaluated the lecture and the lecturer using 7 items in a 5-point-likert-scale (from 0 „totally disagree“ to 4 „totally agree“). Shown are the means of the items, respectively SD. Statistically significant differences were described in the evaluation of Item 1 ( $p < 0.003$ ), Item 2 ( $p < 0.001$ ) and item 5 ( $p < 0.001$ ), marked with \*.

## Screening for participation

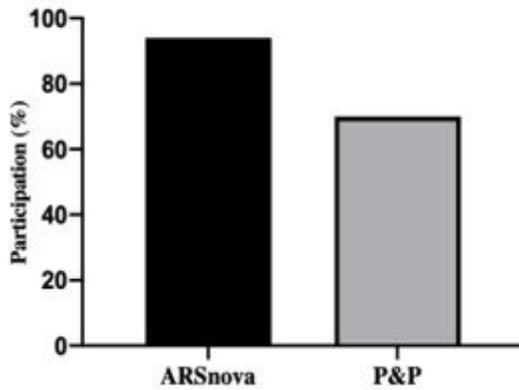


Figure 2

Results regarding participation in MC-questions. In a lecture series of ten lectures 198 students of medicine were screened for participation in two MC-questions at the end of a lecture. 70 students used ARS (ARSnova) to evaluate and answer MC-questions, 128 used a traditional handout (P&P) for answering and evaluation. With the use of ARS participation in MC-questions was higher compared to P&P (0.94 versus 0.7).