

# Teaching in Geriatrics: The Potential of a Structured Written Feedback for the Improvement of Lectures

Theresa Pohlmann<sup>1,2</sup>, Volker Paulmann<sup>3</sup>, Sandra Steffens<sup>2</sup>, Klaus Hager<sup>4</sup>

<sup>1</sup>Klinikum Dritter Orden, München, Germany

<sup>2</sup>Dean's Office-Evaluation Office, Hannover Medical School, Hannover, Germany

<sup>2</sup>Dean's Office-Curriculum Development, Hannover Medical School, Hannover, Germany

<sup>4</sup>Hannover Medical School (MHH), Institute for General Medicine and Palliative Medicine, Hannover, Germany

## Abstract

**Objective:** Lectures are worldwide still a widespread concept of knowledge transfer. The module "Medicine of Ageing and of People of Age" (geriatrics) at the Hannover Medical School uses lectures as one means of knowledge transfer.

**Materials and Methods:** This study aimed to analyze whether a criteria-based written feedback for the lecturers can improve their teaching. In a prospective longitudinal design 17 lectures are rated by a trained student reviewer in two consecutive trimesters according to a questionnaire covering 22 items. The students' perceptions are evaluated using a standardized query with five additional questions.

**Results:** The overall rating of the lectures (1= not apparent; 5= excellent) improves from 3.8 (T0) to 4.4 points in the second evaluation (T1) (+0.59 points,  $p < 0.001$ ). Ratings in all three main categories (content/structure, presentation, visualization) increase significantly in the second series of lectures. A significant amelioration can be seen in six of the 22 items, especially in "content/structure". The perceptions of the students show a trend for a better rating, too.

**Conclusion:** Lecturers can benefit from an additional feedback to their lectures. The review should follow a standardized procedure and should be communicated transparently. Therefore, an individual criteria-based review by a trained student reviewer is a viable solution.

**Keywords:** Geriatrics, teaching, evaluation, university, lectures

## Introduction

Lectures are the basis of knowledge transfer and should be evaluated according to content and structural criteria. In order to continuously improve the quality of teaching, measurement and evaluation of lectures is crucial (1). The quality of good teaching is based on different criteria examined in the literature. However, there is no unanimous definition of "good teaching", but rather many different points of view, e.g., student satisfaction, outcome of teaching, or the qualification of the teachers. There is also a strong heterogeneity among evaluation questionnaires. This study focuses on an individualized, criteria-based written feedback from a trained student reviewer. Each lecture is evaluated separately with respect to content, organization, and quality.

The module "Medicine of Ageing and of People of Age" (geriatrics) at the Hannover Medical School is taught in the fourth of six years of undergraduate medical education and is divided into a theoretical and a practical part. The optional 20 lectures with 45 minutes each (= one teaching hour) take place within one week. Practical aspects are covered in 10 mandatory teaching units of 90 minutes each, which also include patient contact in the hospital. With a total teaching time of 20 hours, the module is above the national average of 8.3 hours (2). Because of their large proportion and voluntary nature, it is especially important to make the lectures attractive for the students. When it comes to quality assessment, students' evaluations are widely recognized as a feedback tool. However, it is sometimes difficult for the module organizer to decide whether a lecturer

**Address for Correspondence:** Klaus Hager, Hannover Medical School (MHH), Institute for General Medicine and Palliative Medicine, Hannover, Germany

**Phone:** +491776627561 **E-mail:** Hager.Klaus@mh-hannover.de **ORCID:** orcid.org/0000-0003-4497-3878

**Received:** 26.11.2021 **Accepted:** 14.02.2022

**Cite this article as:** Pohlmann T, Paulmann V, Steffens S, Hager K. Teaching in Geriatrics: The Potential of a Structured Written Feedback for the Improvement of Lectures. Eur J Geriatr Gerontol 2022;4(3):123-128

©Copyright 2022 by the Academic Geriatrics Society / European Journal of Geriatrics and Gerontology published by Galenos Publishing House.



is teaching successfully since, the central university evaluation forms usually cannot provide feedback for every single lecture. Instead, as a compromise, an overall assessment is recorded that often combines different forms of instruction (seminars, bedside teaching, lectures) as well as different lecturers.

For high-quality lecturing, some features are important. Copeland et al. (1) validated some predictors for successful learning such as clear and organized lectures, a case-based format, encouraging to engage the audience's attention, identifying important points or presenting relevant material with readable slides. According to the Kirkpatrick model, all levels (reaction, learning, behavioral change, organizational performance) should be implemented when delivering feedback to the instructors (3). The study aims to analyze if a criteria-based written feedback for the lecturers can improve the lectures in terms of content, organization and quality. In addition, the study also considered whether the consequence of this feedback was reflected by the general students' evaluations.

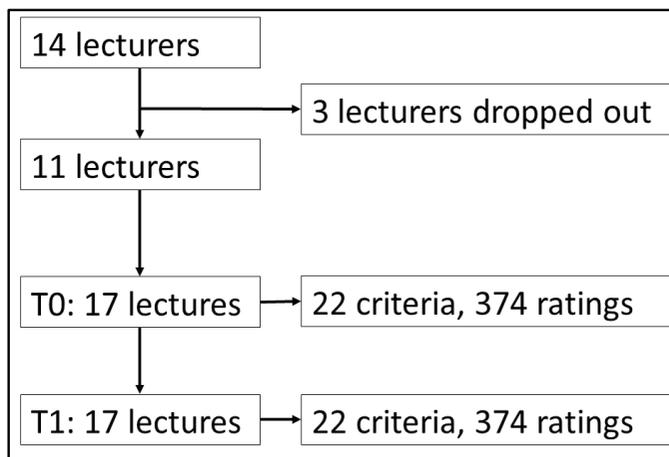
## Materials and Methods

**Study Design:** This study is a prospective longitudinal analysis. A total of 14 lecturers are involved in the lectures of the Geriatrics module (October 2017 to March 2018). The lecturers were recruited from different departments of the medical school and among geriatricians from a nearby geriatric hospital. These include the departments of general medicine, cardiology, nephrology, trauma surgery, neurology, history/ethics/philosophy, forensic medicine, clinical pharmacology and psychiatry. The lecturers had no special training before teaching the geriatrics module and there was a wide range of teaching experiences and didactical training. The lecturers were informed in advance, both verbally and in writing, about how the study would be conducted. Of the 14 lecturers, 13 agreed to participate in the study. Subsequently one lecturer withdrew from the study and one lecturer could not be included in the study due to a missing comparison lecture in the second lecture week. As a result, a total sample size of  $n=11$  lecturers (three female and eight male) participated and gave their written consent for the evaluation, thus the willingness to participate was 86%. During the entire module, one lecturer held five lectures, two lecturers held two lectures each, and the remaining lecturers held one lecture (Figure 1).

The study design was reviewed and approved by the Ethics Committee of the Hannover Medical School (no. 3634-2017).

### Data Collection

Using a five-point Likert scale (1= not apparent; 5= excellent), the lectures are rated according to a questionnaire that consisted of 22 items in the categories "content/structure", "presentation", and "visualization" (Table 1). The questionnaire



**Figure 1.** Scheme of the study (T0: first lecture period; T1: second lecture period)

was developed by Ruessler et al. (4) and based on criteria for effective teaching identified in the literature. As a validated assessment instrument, it was put forth by Newman et al. (5,6). The questionnaire has already been used successfully to evaluate lectures on emergency medicine and surgery (4,6,7). The geriatric lectures at MHH were evaluated over two consecutive trimesters (fall and winter trimester). In total, 17 lectures were evaluated two times employing 22 criteria ( $n=748$  ratings).

The evaluation is carried out by a trained female student from the fifth year who had already completed the module. During a training session prior to the evaluation cycle, a five-member expert (experienced teachers, a MHH alumnus, a trained social scientist of the central evaluation unit) group evaluated a video-taped prototype lecture as an example. The results were presented and discussed in the group, explicitly pointing out possible observation and evaluation errors, such as the halo effect, the primacy effect and the error of central tendency (4,5,7).

Based on the first evaluation in the fall trimester, individual written feedback was emailed to each lecturer for each lecture given. The feedback contains a general summary of strengths and suggestions for improvement, including free comments as well as "closed" items. Furthermore, a comparative rating of the individual aspects compared to the other lecturers is included (Figure 2).

In addition, all students who attended the geriatrics module (T0= 96 students, T1= 76 students) were informed about the study and were invited to participate in the central standard, end-of-trimester student evaluation (Table 2). In the first trimester  $n=75$  students participated and  $n=60$  students in the second trimester (T0 = 78%; T1 = 79%). Among other things, this includes an overall evaluation of the module (scale: 0 points = deficient <> 15 points = very good). In addition to a standardized query, five additional questions were asked that specifically address the teaching objectives, lecture structure, the sequence of the lectures, relevance to routine medical

**Table 1. Individual item analysis and test values for the Wilcoxon test (n=17)**

Aspect	T0 Mean/median	T1 Mean/median	p	Z
<b>Content/structure</b>				
Clear learning objectives	1.41/1	3.82/5	<b>0.001</b>	-3.443
Transparent sequencing	1.76/1	4/5	<b>0.001</b>	-3.336
Clear organization	3.24/3	4.12/4	<b>0.003</b>	-2.950
Connections to prior knowledge	3.94/4	4.35/5	0.112	-1.588
Memorable visualization	4.18/4	4.59/5	0.083	-1.732
Clear instructions	4.29/5	4.47/5	0.454	-0.749
Active inclusion	4.35/5	4.47/5	0.739	-0.333
Appropriate amount of data	4.65/5	4.82/5	0.257	-1.134
Content summaries	2.82/3	4.18/5	<b>0.002</b>	-3.108
Adequate time management	3.35/3	4.24/5	0.060	-1.879
<b>Presentation</b>				
Speaking rate	4.41/5	4.53/5	0.480	-0.707
Volume/pronunciation	4.76/5	4.82/5	0.564	-0.577
Enthusiasm for the topic	4.65/5	4.59/5	0.564	-0.577
Respect for listeners	3.71/4	4/4	0.132	-1.508
Inviting questions from listeners	4/4	4.76/5	<b>0.046</b>	-1.997
Discussion moderation	4.08/4	4.77/5	0.053	-1.933
Language of the slides	4.24/5	4.53/5	0.131	-1.512
<b>Visualization</b>				
Adequate design	3.88/4	4.18/4	0.132	-1.508
Graphics/diagrams/images	3.76/4	3.82/4	0.903	-0.122
Amount of text on slides	3.76/4	4.06/4	0.166	-1.387
Congruence between image and language	4.18/4	4.65/5	0.097	-1.660
Adequate number of slides	4.18/4	4.94/5	<b>0.006</b>	-2.739
Scale 1= not apparent, 5= excellent, bold= significant				

practice, and the students' prior knowledge (scale: 1=agree completely <> 6=disagree completely) (Table 3).

**Statistics**

Statistical analysis is performed using Microsoft Excel 2018®, version 6.13.1, and SPSS (version 25). A paired-sample t-test is used for the rating differences in the overall evaluation before and after feedback (T0 and T1), a p-value of p<0.05 indicates statistical significance. The data for the evaluated items do not show a normal distribution in most cases, which is why the Wilcoxon test for dependent samples is carried out.

The student evaluations, including the additional five questions, are analyzed using the t-test for independent samples after verifying the pre-requisites for this.

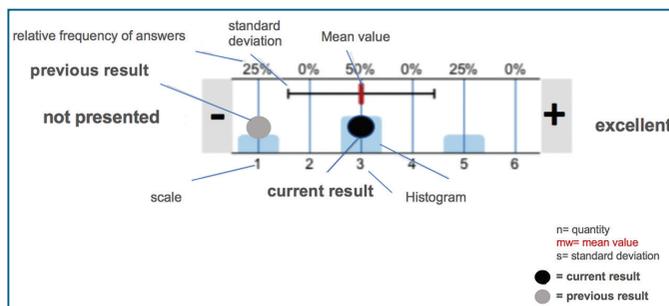


Figure 2. Scheme of visual feedback

**Table 2. Some results of the general student evaluation (T0: n=75. T1: n=60)**

Question	T0	T1	p	Standard error of the difference
Course content (1= very good <> 6= failure)	1.84	1.62	0.052	0.112
Instructors (standardized academic grade 1- 6)	1.56	1.48	0.469	0.106
Overall evaluation (0 pts.= failure <> 15 pts.= very good)	12.8	13.18	0.198	0.298

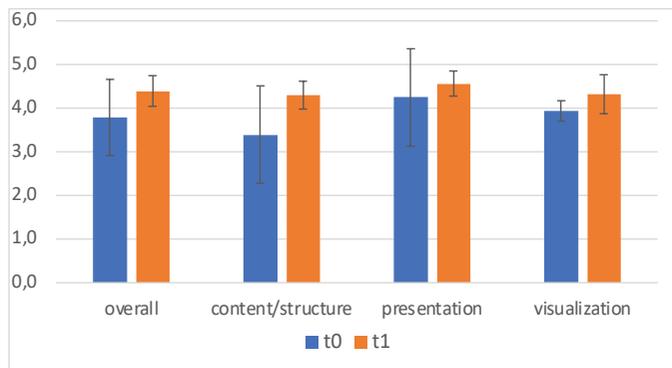
**Table 3. Results of the additional student survey (T0: n=75. T1: n=60)**

Question	T0	T1	p	Standard error
The lecturer has clearly recognizable teaching objectives.	1.757	1.638	0.363	0.131
The content taught in the lecture was well organized.	1.861	1.746	0.332	0.119
The narrative thread of the course was clearly visible.	1.903	1.554	<b>0.006</b>	0.125
The relevance of the topics important for future medical practice became clear.	1.608	1.357	<b>0.022</b>	0.113
The lecture built upon prior knowledge.	1.676	1.5	0.137	0.117

Scale: 1= agree completely 6= disagree completely, bold= significant

**Results**

Reviewing the lectures, a mean rating of 3.8 out of 5 points for all items is calculated at the first evaluation (T0) in the fall trimester and a mean of 4.4 points at the second evaluation (T1) in the winter trimester (±0.59 points, p<0.001) (Figure 3). All three main categories (content/structure, presentation, visualization) are rated significantly better in the second series of lectures. A significant improvement can be seen in six of the 22 items (Table 1). The most significant improvement for a single lecture is more than one point, the largest increase can be



**Figure 3.** Comparison of the overall review of the lectures and the three major categories at the first (T0) and the second (T1) evaluation (five-point Likert scale; 1=not apparent; 5=excellent)

seen in the category "content/structure".

Similar to the significantly improved results in the reviewer evaluations, there is also a trend of an improvement in the students' general evaluation of the module (Table 2). At first, the geriatrics module is rated with 12.8 out of a possible 15 points. After the intervention this already solid result improves to 13.2 ( $\pm 0.4$ ) points (n.s.). A positive trend can be seen between the trimesters in the "instructor ratings" starting at 1.56 and moving to 1.48 (n.s.) and as well regarding "course content" moving from 1.84 to 1.62 ( $p=0.052$ ) (Table 2).

With regard to the additional items that cover the learning outcome and overall satisfaction, the students responded with significantly higher ratings to the question about "being able to recognize the narrative thread (sequencing) running through the lecture series" ( $p<0.001$ ) and to the question about the "relevance of the topics covered to future medical practice being clear" ( $p=0.022$ ) (Table 3).

## Discussion

Lectures as a means of teaching: Despite the criticism of this format at German medical schools, knowledge is still imparted in over 90% of the time through lectures (8). In the module geriatrics at the Hannover Medical School lectures account for two thirds of the curriculum. This reflects a general tendency in geriatrics—as well as in other small subjects—with their limited teaching resources. Only few different formats for teaching geriatrics in undergraduate medical education are described in the literature (9–11). Most of these studies focus on innovative teaching formats and not on improving the standard lectures themselves. Also, many of these evaluations are based only on student feedback, which gives an overall rating of the module but does usually not rate the single lectures held by individual instructors.

Previous studies have shown that student feedback from the lecture hall does not always appropriately rate the quality of the

course content or the materials used (12). Student feedback on instructors can be influenced by other factors that are beyond or only partially within the control of the instructors, for instance the influence of prior knowledge and interest, gender or expectations regarding test scores (13). In contrast to the student feedback, the evaluation by independent reviewers is not influenced by these factors similar to peer reviewing. This has clear advantages compared to a student feedback, as shown by the study of Sterz et al. (7). Furthermore, training of the reviewer prior to the evaluation can minimize the risk of bias (7).

The individual feedback in our study was accepted by the lecturers, because it may have been easier to accept a feedback from a trained student who already passed the module than from the university or a colleague.

In addition, individual feedback on specific lectures is more valuable than a summative feedback on the entire module (14).

A criteria-based feedback is one of the best methods for generating differentiated feedback. It has been shown that a personal written feedback improves the extent and quality of the feedback, especially when it is structured with specific criteria (15).

What did the feedback change? The study shows the biggest improvements in the sub-section "content/structure". This could be due to the fact, that this area offered the most potential for improvement and that the related didactical suggestions could be implemented by the instructors with relative ease. Another explanation could be that lecture content or organization is easier to improve than other aspects since lecturers can reorganize lecture content or structure without changing deeply rooted personal traits or routinely adapted skills.

In contrast, "speaking rate" and "speaking volume" each received the same ratings at both measurement points. Ruessler et al. (4) saw similar results and pointed out that it is very difficult to change individual characteristics based on a single instance of written feedback.

In addition, there was also a significant improvement in "inviting questions from listeners". In contrast, the category "active inclusion" of students has remained mostly similarly assessed in both periods in our study. Knight and Wood (16) show similar results, although pure interactive classroom activities also have disadvantages. The significant change in "inviting questions" in this study may indicate that instructors were already placing more emphasis on interacting with students because of the feedback given, even though this occurred in the context of traditional lectures. Furthermore, the improvement in the section "content summaries" shows that the feedback encouraged the lecturers to summarize the key facts at the end, which was also directly acknowledged by the students as their

rating in the category "narrative thread of the course" increased significantly in the second lecture week.

Benefit for the lecturers? Breaking down written feedback into identifying strengths and making suggestions for improvement is useful for promoting intrinsic motivation for faculty, as it is a direct recognition of individual performance.

Our survey also found that the lecturers, despite the increased amount of work and the feeling of being observed, viewed the feedback favorably and found added value in it. Moreover, all of the surveyed lecturers were prepared to revise their lectures making it possible to use the feedback as a source of concrete improvements. Reviewing one's own lecture using a criteria-based method and benchmarking it with the other lecturers (Figure 1) May have facilitated acceptance.

The perception of the students: The improved overall rating of the module by the students in the central evaluation suggests that they also saw an improvement not only in the quality of teaching, but also in the quality of their personal learning success. Therefore, the improvement due to the structured feedback was not only noticeable in the evaluation of the instructors, but also in the evaluation of the students.

### Study Limitations

Due to the limited number of lectures this study does not make use of a control group, which received no written feedback or an alternative format for feedback. Furthermore, there could be a potential ceiling effect in some categories, because good results have already been achieved in T0. Despite the good prior results, it is still possible to show that written feedback triggered significant improvements in some categories. Another limitation is the conduction of the study with only one reviewer who may have been biased despite the training. Using multiple reviewers could have allowed for a greater reliability. In addition to all of this, a great willingness of the lecturers to participate is necessary for such a study. In total, participation of the lecturers in our study is as high as 86%. Another limiting factor was that learning improvement in students regarding the content taught was not directly tested using an objective competency assessment, but rather by compiling the students' subjective perceptions.

### Conclusion

This study shows that a significant improvement in teaching is possible by means of an individualized, criteria-based written feedback for each lecture by an independent as well as trained student reviewer and that students acknowledge the resulting improvements positively.

### Acknowledgements

Special thanks to Klaas Brandt, Constantin Büttner and Kristina

Schaubert for their extensive support. The authors also wish to thank the students and instructors for their willingness to participate in the study.

### Ethics

**Ethics Committee Approval:** The study design was reviewed and approved by the Ethics Committee of the Hannover Medical School (no: 3634-2017).

**Informed Consent:** Informed consent was obtained.

**Peer-review:** Internally and externally peer-reviewed.

### Authorship Contributions

Concept: T.P., S.S., Design: T.P., S.S., Data Collection or Processing: T.P., K.H., Analysis or Interpretation: T.P., V.P., S.S., K.H., Literature Search: T.P., V.P., S.S., K.H., Writing: T.P., V.P., S.S., K.H.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study received no financial support.

### References

1. Copeland HL, Longworth DL, Hewson MG, Stoller JK. Successful lecturing: a prospective study to validate attributes of the effective medical lecture. *J Gen Intern Med* 2000;15:366-371.
2. Kolb G. Unterricht Q7 (Medizin des Alterns und des alten Menschen) an 36 deutschen medizinischen Fakultäten. In: Kolb G, Leischker, AH, ed. *Medizin des alternden Menschen – Lehrbuch zum Gegenstandskatalog der neuen ÄAppO*. Wissenschaftliche Verlagsgesellschaft; 2009.
3. Kirkpatrick D. Evaluating training programs: the four levels. vol 46. *Evaluation in Education and Human Services*. Berrett-Koehler; 1994.
4. Ruesseler M, Kalozoumi-Paizi F, Schill A, Knobe M, Byhahn C, Müller MP, Marzi I, Walcher F. Impact of peer feedback on the performance of lecturers in emergency medicine: a prospective observational study. *Scand J Trauma Resusc Emerg Med* 2014;22:71.
5. Newman LR, Lown BA, Jones RN, Johansson A, Schwartzstein RM. Developing a peer assessment of lecturing instrument: lessons learned. *Acad Med* 2009;84:1104-1110.
6. Newman LR, Brodsky DD, Roberts DH, Pelletier SR, Johansson A, Vollmer CM Jr, Atkins KM, Schwartzstein RM. Developing expert-derived rating standards for the peer assessment of lectures. *Acad Med*. 2012;87:356-363.
7. Sterz J, Hofer SH, Bender B, Janko M, Adili F, Ruesseler M. The effect of written standardized feedback on the structure and quality of surgical lectures: A prospective cohort study. *BMC Med Educ* 2016;16:292.
8. Singler K, Stuck AE, Masud T, Goeldlin A, Roller RE. [Catalogue of learning goals for pregraduate education in geriatric medicine. A recommendation of the German Geriatric Society (DGG), the German Society of Gerontology and Geriatrics (DGGG), the Austrian Society of Geriatrics and Gerontology (OGGG) and the Swiss Society of Geriatric Medicine (SFGG) on the basis of recommendations of the European Union of Medical Specialists Geriatric Medicine Section (UEMS-GMS) 2013]. *Z Gerontol Geriatr* 2014;47:570-576.
9. Granero Lucchetti AL, Ezequiel ODS, Oliveira IN, Moreira-Almeida A, Lucchetti G. Using traditional or flipped classrooms to teach "Geriatrics and Gerontology"? Investigating the impact of active learning on medical students' competences. *Med Teach* 2018;40:1248-1256.
10. Sauer M, Gornig M, Voigt K, Schubel J, Bergmann A. [Interprofessional

- multistation practical training in the auditorium : Implementation of the national competence-based catalog of learning objectives for undergraduate medical education in the cross-sectorial area medicine of aging and the aged at the TU Dresden]. *Z Gerontol Geriatr* 2018;51:903-911.
11. Eckardt R, Nieczaj R, Steinhagen-Thiessen E, Arnold T. [Cross-sectional field Q7"medicine of aging and the elderly" at the Charité - Universitätsmedizin Berlin: Curriculum and evaluation by students]. *Z Gerontol Geriatr* 2013;46:548-555.
  12. Irby DM. Peer review of teaching in medicine. *J Med Educ* 1983;58:457-461.
  13. Hatfield CL, Coyle EA. Factors that influence student completion of course and faculty evaluations. *Am J Pharm Educ* 2013;77:27.
  14. Knol MH, Dolan CV, Mellenbergh GJ, van der Maas HL. Measuring the Quality of University Lectures: Development and Validation of the Instructional Skills Questionnaire (ISQ). *PLoS One* 2016;11:e0149163.
  15. Newton PM, Wallace MJ, McKimm J. Improved quality and quantity of written feedback is associated with a structured feedback proforma. *J Educ Eval Health Prof* 2012;9:10.
  16. Knight JK, Wood WB. Teaching more by lecturing less. *Cell Biol Educ* 2005;4:298-310.