

# Three years of experience with a workshop for medical specialty examiners in South Baden: A project report and initial evaluation results

## Abstract

**Introduction:** An oral exam (30-60 minutes) is administered at the end of every post-graduate medical specialty program or is required to attain additional specialized qualifications. In both undergraduate and post-graduate medical education oral exams are not considered to be very objective or reliable. To improve the quality of exams in medical specialties, the Regional Medical Association for South Baden (Bezirksärztekammer Südbaden) decided in 2013 to offer a training program for head examiners and others responsible for administering exams in medical specialties.

**Project Description:** Following a survey of examiners and examinees conducted from January through June, 2013, on the difficulty level of examination questions, satisfaction with the test, and the need for training in administering exams, the first workshop of its kind was designed. Since 2013, six workshops with a total of 93 participants have been held and evaluated.

**Results:** The evaluations (response rate: 86%) showed a high level of acceptance for the concept behind the training. A large number of participants felt the need to define minimum standards for exams, to standardize the required level of difficulty and the assessment criteria in each subject, and to give examiners the appropriate tools needed to improve the validity and reliability of the exams.

**Conclusion:** Offering a training program for those responsible for administering medical specialty exams appears to be both meaningful and necessary in order to meet the existing need for increased validity and reliability. In light of the initial experiences with this workshop and the differing percentages of failed exam attempts nationwide, the implementation of examiner training is to be recommended in other regions in Germany. In other European countries examiners conducting medical specialty exams undergo appropriate training before administering their first exam.

**Keywords:** Medical specialty exam, oral assessments, examiner workshop, post-graduate training

Irmgard  
Streitlein-Böhme<sup>1</sup>  
Wilhelm Niebling<sup>2</sup>  
Götz Fabry<sup>3</sup>  
Klaus Böhme<sup>1</sup>

1 Universität Freiburg,  
Lehrbereich  
Allgemeinmedizin, Freiburg,  
Germany

2 Universität Freiburg, Freiburg,  
Germany

3 Universität Freiburg, Abt. für  
Med. Psychologie, Freiburg,  
Germany

## Introduction

In Germany oral assessments have a long tradition not only in undergraduate medical education, but also in post-graduate medical specialization. However, the validity and reliability of oral exams have been viewed as overall low, primarily because they are very susceptible to grading errors and biases. This has been demonstrated by numerous studies [1], [2]. In particular, the tendency described by Wakeford et al. to test only for factual knowledge, not to give separate scores for partial performance and the “minor errors” belong to the most important examples of grading biases and errors that can occur in the context of a medical specialty exam. Wakeford et al. see a solution in the selection and training of examiners. In the

training sessions the examiners should measure themselves against the exam criteria [2].

Jefferies et al. also consider it important to explore ways to increase the quality of practical exams, above all, to implement a stronger focus on competency [3].

To achieve a higher degree of objectivity, reliability and thus validity in an oral exam, it is necessary to structure and standardize both the test situation and the test questions [4], as previously called for in 1999 in the recommendations regarding structured oral assessments in the medical specialty exams. These stress the necessity of ensuring identical test conditions for examinees in order to achieve standardization of the exam. It is therefore necessary to define and weight the assessment objectives and content in advance for each particular medical spe-

cialty exam. It would be most logical if all of the examiners agreed upon exam content and identified which grading criteria apply to candidates who fail the exam. An improvement in structuring will be evident mainly in a written record of the test questions formulated in advance along with the corresponding expectations regarding the correct answers and a scoring rubric. The exam questions should be case-based and contain aspects relevant to practice (e.g. x-rays, ECGs, sonographic images, etc.).

Currently in Germany there are only few examples of how to improve the medical specialty exam. For instance, in Thuringia a change was made to the exams in 2014 in response to demands made by examiners: topics and sets of questions were defined to better standardize the exam, and exam content was modernized, for example, through digital images in radiology or anchoring practical components in the exam. As a result, the 30-minute timeframe for the exam was extended to 45-90 minutes per examinee [5].

The question arises as to which improvements in assessment quality are possible within the existing situation. Among different European countries (Denmark, Germany, The Netherlands, Poland, and Great Britain) some mandatory training courses for examiners have been established. This was ascertained as part of a pilot survey of an expert panel consisting of examiners from five major EU countries who were responsible for administering the medical specialty exam in general practice. Only in Poland and Germany was this kind of training not offered. In addition, the survey revealed that all of the countries except for Germany carry out longitudinal evaluation of performance over the entire course of post-graduate medical training [6].

That training sessions can completely change grading behavior is shown by an analysis of scoring on the oral/practical M2 state examinations at the end of undergraduate medical education. Schickler et al. identified differences in the scores of examinees who were evaluated by examiners with or without prior training. The highest grade ("very good") was given much less often by examination committees with members who had attended a workshop addressing the issue of grading [7].

### Initial situation in South Baden

In April 2012 as part of an extraordinary meeting of the Regional Medical Association for South Baden addressing the current state of post-graduate clinical and practical training, the decision was made to define (minimum) standards for medical specialty exams and to initiate dialogue among the examiners [8]. As a consequence, in 2013 the Regional Medical Association offered a workshop for the first time aimed at examiners responsible for administering medical specialty exams. The most urgent goal of the workshop was to effect qualitative improvements in the content and method of the 30- to 60-minute specialty exam administered in South Baden.

The first step involved evaluating the medical specialty exams. From January to June, 2013, the head examiners,

examiners and examinees were surveyed immediately after the 102 medical specialty exams administered during this period. The questionnaire with 13 items, plus open-ended questions, was primarily focused on the level of difficulty of the exam questions, the level of satisfaction with the assessment, and the need for extra training on the topic of assessments.

The qualitative analysis of the initial evaluations after the medical specialty exams showed that only 27.3% of the head examiners and 23.4% of the examiners saw a need to train examiners to administer the medical specialty exams [8]. The open-ended responses, however, indicated numerous suggestions for content-based and structural improvements to the previous medical specialty exams. For example, 20.8% of examiners and 36.4% of the head examiners desired a standardization of the exam questions for each subject area and/or determination of grading criteria prior to the exam. A total of 7.8% of examiners felt 30 minutes for a medical specialty exam were too short. The main criticisms were mainly the varying levels of difficulty of the exam questions, the absence of coordination among examiners regarding exam content, and the lack of required subject-specific content and grading criteria.

With only a few exceptions the examinees demonstrated a high level of acceptance for the exams that had taken place. It must be pointed out that the annual failure rate for the medical specialty exams in South Baden is <3% [8].

In view of the high level of satisfaction among the examiners and head examiners with the mode of testing at the time, the question arose whether this group would be accepting of a training measure offered by the Regional Medical Association precisely for examiners. What could the content of such a training program encompass and which improvements could be achieved as a result?

## Project Description

Based on the decision of the representatives and the executive board of the Regional Medical Association for South Baden, supported by international examples [2], [3], [4] and a wide variety of suggestions for improvement made by examiners in the initial survey, a training program comprised of six units (4.5 hours) was developed at the beginning of 2013 and heavily influenced by the workshop offered at the time by the medical schools in Baden-Württemberg for examiners conducting the oral/practical section of the M2 exam (now the M3 exam) [9]. Adaptation for the medical specialty exams is shown in table 1, and the most important improvements described above were implemented in the medical specialty exams:

Prior to the workshop, each participant received example materials and was asked to prepare two case vignettes with exam-relevant patient cases from practice, including the grading criteria for correct answers, and to bring other exam materials such as x-ray images, ECG leads,

**Table 1: Organization of the workshops in terms of content and teaching**

Content	Teaching	Duration
Familiarization with the formal testing procedure (primarily for new examiners),  Significance and implementation of structured oral assessments	Interactive presentation	1 hour, 15 min.
Weighting the specialized medical exam content and standardizing performance expectations,  Coordination of topics within the medical specialty groups	Discussion in small groups according to medical specialty	45 min.
Creation of structured exam questions in the form of case vignettes,  Definition of grading criteria	Individual work (preparatory task),  Workshop: peer review in small groups according to medical specialty	1 hour, 30 min.
Familiarization with and reflection on the exam situation in the roles of examiner and examinee.	Role play: example exam situations are presented, then reflected upon and discussed primarily with a focus on the role of the examiner	1 hour

and lung function results to the workshop. The case vignettes were first reviewed internally as part of the workshop to then be used later as exam questions during the subsequent exam simulation.

The workshop has taken place six times since 2013 with a total of 93 participants coming from a wide range of medical disciplines and has been evaluated each time at the end. During this time period it has been slightly modified and adjusted for the particular number of attendees and composition of medical disciplines.

A participant evaluation was administered at the end of each workshop. The evaluation survey had 12 items on the relevancy and usefulness of the workshop's content regarding the medical specialty exam and level of satisfaction with the workshop. It was possible to include open-ended responses for the following questions:

- Do you have any suggestions for improving the forms and/or materials?
- For which aspects would you like more information and/or support?
- What was the most important thing you learned from this workshop?

Analysis of the open-ended responses included sorting and categorizing as the first step. Then, four different raters independently assigned the responses to the pre-defined categories.

## Results

Of the 449 examiners registered with the Regional Medical Association for South Baden only around 250 are active, and only ten of the 33 head examiners. Of these 260 active examiners and head examiners, 93 (36%) have taken part in the training program. Of these, 42 participants (52.5%) indicated they had more than five years of experience as an examiner, ten participants (12.5%) indicated three to five years, and 12 (15%)

between one and two years. Eleven participants (13.7%) had not yet administered an exam, and five (6.3%) gave no indication of experience.

The specific medical specialties of the examiners who participated in the six training programs are presented in table 2. Mostly, the examiners are specialized in internal medicine, surgery, general practice, gynecology, and emergency medicine.

The evaluations administered after each training session (response rate: 86.0%) showed a high level of acceptance for the training strategy, primarily the opportunity to actively participate during the workshop and the positive atmosphere among medical colleagues. In addition to the structured presentation of information and the theoretical principles underlying the need for a structured oral exam, the participants found the generation of case vignettes, the internal review of the exam questions by medical colleagues and the simulated exam especially useful (see figure 1).

The most meaningful result to arise from analysis of the open-ended responses was seen in answers to the question about the most important insight from the workshop. Participants primarily found proposals to standardize and formally structure exams to be meaningful. Likewise, exchange and coordination between examiners for a specific specialty was of great importance. A verifiable and comparable structuring of the medical specialty exams, standardization of grading, prior coordination among examiners, and the creation of an internal catalogue of case vignettes for each specialty were explicitly mentioned. In addition, the examiners and head examiners had a clearer understanding of the legal and formal aspects of the medical specialty exam. Many experienced examiners felt affirmation regarding how they had previously handled exams but were still able to take away additional suggestions from the workshops for future exams.

Several open-ended responses on the most important insights from the workshops are listed in figure 2 as ex-

**Table 2: Distribution of subjects in the evaluation surveys analyzed (N=80)**

Specialty	Number of surveys	Additional medical qualifications	Number of surveys
Internal Medicine (Cardiology, Angiology, Phlebology, Oncology)	N = 12	Emergency Medicine	N = 8
Surgery	N = 11	Acupuncture	N = 4
General Practice	N = 9	Physical Therapy	N = 2
Gynecology	N = 7	Addiction Medicine	N = 1
Anesthesiology	N = 5		
Neurology	N = 5		
Psychiatry	N = 4		
Oral and Maxillofacial Surgery	N = 3		
Otorhinolaryngology	N = 3		
Dermatology	N = 1		
Ophthalmology	N = 1		
Psychosomatics	N = 1		
No information given	N = 3		

**Workshop for Medical Specialty Examiners**

Number of evaluations: N = 80 (6 workshops, 2013-2016) Response rate: 86.0%

**Rating**

- 1 = very good  
2 = good  
3 = satisfactory  
4 = sufficient  
5 = deficient  
6 = fail

**Please evaluate the following aspects regarding usefulness for administering the medical specialty exam**

Formal exam procedure	1,72
Theoretical principles behind the exam	1,76
Creation of case vignettes + grading criteria	1,71
Peer review of exam questions	1,71
Creation of an exam blueprint	2,06
Exam simulation	1,76

**Please evaluate the workshop**

How did you perceive the general atmosphere?	1,19
How would you rate the organizational aspects?	1,58
How would you rate the variation of the teaching methods?	1,91
How would you rate the learning materials?	2,01
How would you rate the opportunity to actively participate?	1,34
How would you rate the trainers regarding their ability to impart information?	1,36

**Figure 1: Evaluation of the examiner workshops, 2013-2016 (N=80)**

What was the most important thing you learned from this workshop?	
Comments from the most common categories	
Standardization of exams	“Standardization makes sense.” “Effort should be made to standardize the grading of the exams even more.” “Defining required criteria makes scoring easier.”
Structuring of exams	“Creating a blueprint is very useful for structuring and scoring.” “Tools for structuring are extremely helpful for holding uncomplicated medical dialogue.” “I will structure my exams better and more efficiently.”
Coordination/exchange among colleagues	“Coordination with the co-examiners prior to the exam regarding exam questions” “Sharing experiences with other examiners” “Consensus among examiners is necessary.”

**Figure 2: Selected responses to questions about the most important personal observations from the workshop.**

amples from the three most frequently named topics (standardization, structuring and coordination, and professional exchange with other examiners for the same specialty):

## Discussion

The evaluations at the end of the examiner workshop showed that the examiners see a need to develop shared

minimum standards and a uniform level of requirements for each medical specialty and medical specialty exam and, above all, to give examiners new tools that enable improvements in the objectivity, reliability, and thus validity of the oral assessments [1], [2], [4]. The need for good structuring and standardization appeared particularly important to the examiners in terms of the oral exam. This can primarily be seen in the subjective assessments of the usefulness of the case vignettes, their peer review and the many comments on the most important observations made during the workshops (see figure 2).

Öchsner et. al come to a similar result in a study they conducted in Ulm, Germany, on the effects and sustainability of a training workshop aimed at the M2 exam [10]. In that study, 86% of the workshop attendees indicated that they strived to apply the concept of the “structured oral examination.”

The simulated exam as a practical application of the workshop content, and thus a central element of the medical specialty examiner workshop, was viewed by the examiners as extremely useful for the administration of the medical specialty exam.

A second evaluation of the medical specialty exams with a survey of the examiners, head examiners, and examinees is planned with a focus on answering the following questions:

- Were changes initiated in how the examiner conducted exams as a result of the workshop, and if so, what changes?

Possible criteria for this would include the creation of a shared, subject-specific blueprint and definition of the grading criteria for passing and failing scores.

- Are there differences between examiners who have undergone training and those who have not when administering medical specialty exams?

Possible criteria for this would be whether or not written case vignettes including previously defined grading criteria had been created, if these were used during the exam, and if the formal criteria for a good case vignette had been fulfilled.

The <3% failure rate on the medical specialty exam prior to implementation of the examiner workshop has shown no substantial change in the past three years, so that it must be investigated if this criteria can be drawn upon as a quality marker for a changed testing culture.

If the published failure rates for exams leading to additional qualifications in medical disciplines and specialties given by the different medical associations across Germany are considered, then it is possible to see that these rates vary significantly. Not all of the state medical associations provide these numbers annually or in a transparent manner, and the published failure rates range widely, for instance, from 3.5% in Hamburg [11] to 6.4% in Westfalen-Lippe [12] for all types of specialty exams and from 2.5% in Hamburg [11] to 7.2% in Westfalen-Lippe [12] for exams leading only to an additional occupational designation in a medical specialty.

## Analysis of strengths and weaknesses

The strengths of this project lie in the fact that the training program for medical specialty examiners is the first of its kind in Germany, and thus acts as a role model. For example, the representatives of the Regional Medical Association for South Baden have not only contemplated qualitative improvements to the medical specialty exams earlier than other medical associations, but have also decided to offer and implement appropriate training for examiners and head examiners. In addition, a successful workshop design was developed, one that has been very positively received by the participants.

One of the challenges faced by this project was that the workshop participants came from a wide variety of medical specialties so that it was not always possible to bring at least four examiners from a single specialty (see table 2) together for professional exchange on assessment objectives or exam content or to carry out a simulated exam as part of the workshop. During the first training sessions, this was only a small problem – it had been possible to invite a sufficient number of representatives from three to four medical specialties. In recent training sessions, however, it has become more frequent that only a few representatives have participated in the workshop. In these cases it was necessary for the trainers to closely mentor the representatives during the review phase. During the simulation, these representatives are only able to serve in an observational capacity.

One further limitation is found in that up until now no outcome could be assessed in respect to the training. This is being considered in follow-up investigations.

## Conclusions

Training programs for examiners administering medical specialty exams and other assessments leading to additional post-licensure medical qualifications appear to be both valuable and necessary despite, or perhaps precisely because of, the low failure rates. Since failure rates for these exams vary widely across Germany, aiming for examiner qualification would be worth considering, not just in South Baden, but also in other regions, as this is already standard in other European countries. For improvement of quality across the board, it seems important to develop and implement a uniform concept for all of Germany.

Regardless, the establishment of other assessment procedures for post-licensure medical specialization should also be considered in Germany. Relying only on knowledge-based tests is not sufficient to assess and document the competencies required of medical specialists. It is much more critical to assess how medical expertise is applied in the practical setting [4]. Examples from other European countries point to a better measurement of competency by using longitudinal formats [6]. For this reason, it would make sense to consider a new assessment format for the medical specialty exams in Germany

and implementing a combination of multiple longitudinal assessments as well as a final exam. Similar considerations have been voiced in the position paper published in 2013 by the GMA committee on post-graduate education [13]. In addition to this, a 90-minute assessment at the end of the specialist training, such is already done in Thuringia [7], could be considered.

These new assessment formats could replace the current 30-60 minute-long oral examination in front of two specialist physicians and a head examiner. However, any changes to the assessment format will be accompanied by an increase in costs, as is expressly mentioned by Flum et. al in 2015 [6].

## Note

Dr. med. Irmgard Streitlein-Böhme, Dr. med. Götz Fabry and Dr. med. Klaus Böhme received financial compensation from the Regional Medical Association for South Baden to develop and conduct the examiner workshops.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Van der Vleuten CP. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract.* 1996;1(1):41-47. DOI: 10.1007/BF00596229
2. Wakeford R, Southgate L, Wass V. Improving oral examinations: selecting, training and monitoring examiners for the MRCGP. *BMJ.* 1995;311(7010):931-935. DOI: 10.1136/bmj.311.7010.931
3. Jefferies A, Simmons B, Ng E, Skidmore M. Assessment of multiple physician competencies in postgraduate Training: utility of the structured oral examination. *Adv Health Sci Educ Theory Pract.* 2011;16(5):569-577. DOI: 10.1007/s10459-011-9275-6
4. Bloch R, Hofer D, Krebs R, Schläppi P, Weis S, Westkämper R. *Kompetent prüfen. Handbuch zur Planung, Durchführung und Auswertung von Facharztprüfungen.* Bern, Wien: Universität Bern und Med. Universität Wien; 1999.
5. Landesärztekammer Thüringen. *Tätigkeitsbericht der Landesärztekammer 2014.* Jena: Landesärztekammer Thüringen; 2014. S.23-31. Zugänglich unter/available from: <http://www.laek-thueringen.de>
6. Flum E, Maagard R, Godycki-Cwirko M, Scarborough N, Scherpbier N, Ledig T, Roos M, Steinhäuser J. Facharztprüfung Allgemeinmedizin – was können wir von den europäischen Nachbarn lernen? *GMS Z Med Ausbild.* 2015;32(2):Doc21. DOI: 10.3205/zma000963
7. Schickler A, Brüstle P, Biller S. Mündlich-praktischer Teil des Zweiten Abschnitts der Ärztlichen Prüfung in Freiburg 2012 - Analyse der Notenvergabe zur Überprüfung von Qualitätssicherungsmaßnahmen. *GMS Z Med Ausbild.* 2015;32(4):Doc39. DOI: 10.3205/zma000981
8. Streitlein-Böhme I, Giesler M, Niebling W, Böhme K, Fabry. *Facharztprüferworkshop in Südbaden. Strukturierte Prüfungen nach der Facharztweiterbildung.* Ärztebl Baden-Württemberg. 2014;69(4 Suppl):142.
9. Kompetenznetz "Lehre in der Medizin" in Baden-Württemberg. *Handbuch zur Vorbereitung der Prüfer auf M3.* Freiburg: Kompetenznetz "Lehre in der Medizin" Baden-Württemberg; 2014.
10. Öchsner W, Geiler S, Huber-Lang M. Effekte und Nachhaltigkeit von Trainingsworkshops für den mündlich-praktischen Teil des M2-Examens. *GMS Z Med Ausbild.* 2013;30(3):Doc36. DOI: 10.3205/zma000879
11. Ärztekammer Hamburg. *Tätigkeitsbericht.* Hamburg: Ärztekammer Hamburg; 2014. S.21-29. Zugänglich unter/available from: <http://www.aerztekammer-hamburg.org>
12. Ärztekammer Westfalen-Lippe. *Bericht des Vorstandes der Ärztekammer Westfalen-Lippe.* Münster: Ärztekammer Westfalen-Lippe; 2015. S.48. Zugänglich unter/available from: <http://www.aekwl.de>
13. David DM, Euteneier A, Fischer MR, Hahn EG, Johannink J, Kulike K, Lauch R, Lindhorst E, Noll-Hussong M, Pinilla S, Weih M, Wenneke V. *Die Zukunft der ärztlichen Weiterbildung in Deutschland – Positionspapier des Ausschusses Weiterbildung der Gesellschaft für Medizinische Ausbildung (GMA).* *GMS Z Med Ausbild.* 2013;30(2):Doc26. DOI: 10.3205/zma000869

### Corresponding author:

Dr. med. Irmgard Streitlein-Böhme  
Universität Freiburg, Lehrbereich Allgemeinmedizin,  
Elsässerstr. 2 m, D-79110 Freiburg, Germany, Phone:  
+49 (0)761/270-72461, Fax: +49 (0)761/270-72480  
[irmgard.streitlein-boehme@uniklinik-freiburg.de](mailto:irmgard.streitlein-boehme@uniklinik-freiburg.de)

### Please cite as

Streitlein-Böhme I, Niebling W, Fabry G, Böhme K. Three years of experience with a workshop for medical specialty examiners in South Baden: A project report and initial evaluation results. *GMS J Med Educ.* 2017;34(5):Doc55. DOI: 10.3205/zma001132, URN: <urn:nbn:de:0183-zma0011323>

### This article is freely available from

<http://www.egms.de/en/journals/zma/2017-34/zma001132.shtml>

Received: 2016-11-15

Revised: 2017-06-19

Accepted: 2017-08-07

Published: 2017-11-15

### Copyright

©2017 Streitlein-Böhme et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

# 3 Jahre FacharztprüferInnen-Workshop in Südbaden – ein Projektbericht und erste Evaluationsergebnisse

## Zusammenfassung

**Einleitung:** Eine 30- bis 60-minütige mündliche Prüfung steht am Ende jeder Facharztweiterbildung bzw. ist für den Erwerb von Gebiets-, Schwerpunkts- und Zusatzbezeichnungen erforderlich. In der medizinischen Aus- und Weiterbildung gelten mündliche Prüfungen als wenig objektiv und reliabel. Um eine Verbesserung der Qualität von Facharztprüfungen zu erreichen, hat sich die Bezirksärztekammer Südbaden 2013 entschieden, eine Qualifizierungsmaßnahme für FacharztprüferInnen und Prüfungsvorsitzende anzubieten.

**Projektbeschreibung:** Nach einer von Januar bis Juni 2013 durchgeführten Befragung der PrüferInnen, Prüfungsvorsitzenden und Prüflinge u.a. zum Niveau der Prüfungsfragen, zur Zufriedenheit mit der abgelaufenen Prüfung sowie zum prüfungsmethodischen Fortbildungsbedarf, erfolgte die Konzeption eines bundesweit ersten Workshop-Angebotes. Seit 2013 wurden 6 Workshops mit insgesamt 93 TeilnehmerInnen durchgeführt und evaluiert.

**Ergebnisse:** Die Evaluationen (Rücklauf: 86%) zeigten eine hohe Akzeptanz des Schulungskonzeptes. Eine große Anzahl von TeilnehmerInnen sahen die Notwendigkeit, Minimalstandards für die Prüfungen festzulegen, das Anforderungsniveau und Bewertungskriterien in den jeweiligen Fächern zu vereinheitlichen und den FacharztprüferInnen entsprechende Instrumente an die Hand zu geben, um Validität und Reliabilität der Prüfungen zu verbessern.

**Schlussfolgerung:** Das Angebot eines Schulungsprogramms für FacharztprüferInnen und Prüfungsvorsitzende erscheint sinnvoll und notwendig, um einem bestehenden Bedarf an mehr Validität und Reliabilität zu entsprechen. Nach den ersten Erfahrungen mit diesem Workshop-Angebot und vor dem Hintergrund der bundesweit sehr unterschiedlichen Nichtbestehensquoten wäre die Durchführung von Qualifizierungsmaßnahmen von FacharztprüferInnen auch in anderen Ärztekammerbezirken zu empfehlen. In verschiedenen europäischen Ländern durchlaufen FacharztprüferInnen bereits vor ihrem ersten Einsatz entsprechende Programme.

**Schlüsselwörter:** Facharztprüfung, mündliche Prüfungen, FacharztprüferInnen-Workshop, Weiterbildung

## Einleitung

Mündliche Prüfungen haben in Deutschland sowohl in der Ausbildung des Studiengangs Humanmedizin als auch in der Facharztweiterbildung eine lange Tradition. Ihre Validität und Reliabilität gilt insgesamt als gering, vor allem weil sie sehr anfällig für Beurteilungsfehler und Beurteilungstendenzen sind. Das ist durch eine Vielzahl von Studien belegt [1], [2]. Insbesondere die von Wakeford et al. beschriebene Tendenz, reines Faktenwissen zu prüfen sowie keine Einzelnoten für Teilleistungen zu geben und der sogenannte „Milde-Fehler“ gehören zu den wichtigsten Beispielen für Beurteilungstendenzen und -

fehler, die im Rahmen einer Facharztprüfung auftreten können. Einen Ausweg sehen Wakeford et al. in der Auswahl und Schulung der PrüferInnen. In den Schulungen sollten sich die PrüferInnen selbst den Anforderungen der Prüfung unterziehen [2].

Auch Jefferies et al. halten es für wichtig, sich Gedanken über eine Qualitätssteigerung von praktischen Prüfungen zu machen und vor allem eine stärkere Kompetenzorientierung zu implementieren [3].

Um eine hohe Objektivität, Reliabilität und damit auch Validität in einer mündlichen Prüfung zu erreichen, bedarf es einer „Strukturierung und Standardisierung sowohl der Prüfungssituation als auch der Prüfungsaufgaben“ [4], wie bereits 1999 in den Empfehlungen zu strukturierten mündlichen Prüfungen in Facharztprüfungen gefor-

Irmgard  
Streitlein-Böhme<sup>1</sup>  
Wilhelm Niebling<sup>2</sup>  
Götz Fabry<sup>3</sup>  
Klaus Böhme<sup>1</sup>

1 Universität Freiburg,  
Lehrbereich  
Allgemeinmedizin, Freiburg,  
Deutschland

2 Universität Freiburg, Freiburg,  
Deutschland

3 Universität Freiburg, Abt. für  
Med. Psychologie, Freiburg,  
Deutschland

dert. Hierin wird dargelegt, dass identische Bedingungen für die Prüflinge gewährleistet sein müssen, um eine Standardisierung in der Prüfung zu erreichen. Es ist daher notwendig, im Vorfeld der Prüfung die Prüfungsziele und Prüfungsinhalte für die jeweilige Facharztprüfung festzulegen und zu gewichten. Am sinnvollsten wäre es, wenn alle beteiligten PrüferInnen die fachspezifischen Prüfungsinhalte gemeinsam abstimmen und Kriterien vereinbaren, die für ein Nichtbestehen der Facharztprüfung gelten sollen. Eine Verbesserung der Strukturierung wird vor allem erreicht durch eine vorherige, schriftlich ausformulierte Festlegung der Prüfungsfragen inklusive des dazugehörigen Erwartungshorizontes für die richtigen Antworten sowie des Bewertungsschemas. Die Prüfungsfragen sollen fallorientiert sein und praxisrelevante Anteile (z.B. Röntgenbilder, EKG-Aufzeichnungen, Sonographiebilder etc.) aufweisen.

Zurzeit gibt es in Deutschland nur einige wenige Beispiele zur Optimierung der Facharztprüfung. So wurde in Thüringen beispielsweise aufgrund einer „Forderung der Prüfer“ in 2014 eine Änderung der Facharztprüfungen vorgenommen. Zum einen wurden Themen und Fragenkomplexe festgelegt, um die Standardisierung der Prüfung zu verbessern, zum anderen wurden die Prüfungsinhalte z.B. durch die digitale Bildbetrachtung in der Radiologie oder die Verankerung praktischer Prüfungsteile modernisiert. Dadurch ergab sich allerdings eine Verlängerung der Prüfungszeit von 30 auf 45 bis 90 Minuten pro Prüfling [5].

Es stellt sich auf jeden Fall die Frage, welche Verbesserungen der Prüfungsqualität im Rahmen des Bestehenden möglich sind. In verschiedenen europäischen Ländern (Dänemark, Deutschland, Niederlande, Polen und Großbritannien) sind beispielsweise obligatorische Qualifizierungskurse für FacharztprüferInnen etabliert. Dies wurde unter anderem im Rahmen einer Pilotbefragung einer Expertenrunde von allgemeinmedizinischen FacharztprüferInnen aus fünf großen EU-Ländern erhoben. Nur in Polen und Deutschland gibt es diese Kursangebote nicht. Außerdem zeigte sich in der durchgeführten Befragung, dass vier der fünf untersuchten Länder außer Deutschland „eine longitudinale Überprüfung der Leistungen über die gesamte Dauer der Weiterbildung durchführen“ [6]. Dass Schulungen grundsätzlich das Beurteilungsverhalten verändern können, zeigte eine Analyse zur Notenvergabe bei den mündlich-praktischen M2-Staatsexamensprüfungen am Ende des humanmedizinischen Studiengangs. Schickler et al. haben Unterschiede in der Benotung von Prüflingen, die von PrüferInnen mit bzw. ohne vorherige Schulungsmaßnahme geprüft wurden, gesehen. In den Kommissionen mit Mitgliedern, die einen Prüfer-Workshop absolviert hatten, wurde seltener die Note „sehr gut“ vergeben [7].

## Ausgangssituation in Südbaden

Im April 2012 wurde im Rahmen einer außerordentlichen Vertreterversammlung der Bezirksärztekammer Südbaden mit dem Thema „Weiterbildung in Klinik und Praxis –

Wunsch und Wirklichkeit“ beschlossen, (Mindest)Standards für Facharztprüfungen zu definieren und zu diesem Zweck einen fachlichen Austausch der FacharztprüferInnen zu initiieren [8]. In der Folge hat die Ärztekammer Südbaden 2013 erstmalig für FacharztprüferInnen und Prüfungsvorsitzende einen Workshop angeboten. Vordringliches Ziel des Workshops war es, eine Verbesserung der inhaltlichen und methodischen Qualität der minimal 30-minütigen, maximal 60-minütigen Facharztprüfungen in Südbaden zu erreichen.

In einem ersten Schritt erfolgte eine Evaluierung der Facharztprüfungen. Von Januar bis Juni 2013 wurden sowohl die jeweiligen Prüfungsvorsitzenden und PrüferInnen sowie alle Prüflinge direkt im Anschluss an die in diesem Zeitraum durchgeführten 102 Facharztprüfungen befragt. Der 13 Items sowie offene Fragen umfassende Evaluationsbogen war vor allem auf das Niveau der Prüfungsfragen, die Zufriedenheit mit der abgelaufenen Prüfung sowie den Fortbildungsbedarf zum Thema Prüfungen ausgerichtet.

In der quantitativen Analyse der Eingangsevaluationen nach den durchgeführten Facharztprüfungen zeigte sich, dass nur 27,3% der Prüfungsvorsitzenden sowie 23,4% der PrüferInnen einen Schulungsbedarf für FacharztprüferInnen sahen [8]. Die Freitextkommentare wiesen im Gegensatz dazu zahlreiche Anregungen für inhaltliche und strukturelle Verbesserungen der bisherigen Facharztprüfungen aus. So wünschten sich 20,8% der PrüferInnen und 36,4% der Prüfungsvorsitzenden eine Standardisierung der Prüfungsfragen für das jeweilige Fach bzw. eine vorherige Festlegung von Vorgaben für die Bewertung. 7,8% der PrüferInnen empfanden eine Prüfungszeit von 30 Minuten für eine Facharztprüfung als zu kurz. Die Hauptkritikpunkte waren vor allem das unterschiedliche Niveau der Prüfungsfragen, die fehlende Abstimmung unter den PrüferInnen in Bezug auf die Prüfungsinhalte sowie fehlende fachspezifische Inhaltsvorgaben und Bewertungskriterien.

Die Prüflinge selbst zeigten bis auf wenige Ausnahmen eine hohe Akzeptanz der abgelaufenen Prüfung. Hier muss angeführt werden, dass die jährlichen Durchfallquoten in den Facharztprüfungen in Südbaden bei <3% liegen [8].

In Anbetracht der hohen Zufriedenheit der PrüferInnen und Prüfungsvorsitzenden mit den bisherigen Prüfungsmodalitäten stellte sich die Frage, ob es bei diesem Personenkreis eine Akzeptanz für eine von der Ärztekammer intendierte Prüferschulung geben würde. Wie könnten die Inhalte einer solchen Schulung aussehen und welche Verbesserungen ließen sich dadurch erzielen?

## Projektbeschreibung

Aufgrund des Beschlusses der Vertreterversammlung und des Vorstandes der Bezirksärztekammer Südbaden, inhaltlich unterstützt durch die internationalen Vorbilder [2], [3], [4] sowie die diversen Verbesserungsvorschläge und Anregungen von Seiten der PrüferInnen und Prüfungs-

Tabelle 1: Inhaltlicher und didaktischer Aufbau des Workshops

Inhalt	Didaktische Umsetzung	Dauer
Kennenlernen des formalen Prüfungsablaufes (vor allem für neue PrüferInnen) Bedeutung und Umsetzung strukturierter mündlicher Prüfungen	Interaktive Präsentation	1 h 15
Gewichtung der fachspezifischen Prüfungsinhalte und Standardisierung des Leistungsniveaus Thematische Absprache der jeweiligen FacharztprüferInnen-Gruppen und	Diskussion in fachspezifischen Kleingruppen	0 h 45
Erstellung strukturierter Prüfungsfragen in Form von Fallvignetten Definition von Bewertungskriterien	Einzelarbeit (Vorbereitungsaufgabe) Im Workshop: kollegiales Review in fachspezifischen Kleingruppen	1 h 30
Kennenlernen bzw. Reflexion der Prüfungssituation - sowohl in der Rolle als Prüfer als auch in der Rolle als Prüfling	Rollenspiel: exemplarische Prüfungssituationen werden dargestellt, anschließend wird das Geschehen reflektiert und diskutiert, v.a. mit dem Fokus auf der Rolle des Prüfenden	1 h 00

vorsitzenden in der Eingangsbefragung, wurde Anfang 2013 ein sechs Unterrichtseinheiten (4,5 Stunden) umfassendes Schulungsprogramm entwickelt, das sich stark am damaligen Workshop-Angebot der Medizinischen Fakultäten in Baden-Württemberg für PrüferInnen der mündlich-praktischen M2-Prüfung (jetzt M3-Prüfung) orientierte [9]. Die Adaptation an die Facharztprüfungen war folgendermaßen aufgebaut (siehe Tabelle 1) und enthielt die wichtigsten in der Einleitung beschriebenen Aspekte zur Verbesserung der Facharztprüfungen:

Alle TeilnehmerInnen wurden vor der Veranstaltung gebeten, entsprechend einer zugesandten Vorlage zwei Fallvignetten mit prüfungsrelevanten Patientenfällen aus der Praxis einschließlich des Erwartungshorizontes korrekter Antworten zu erarbeiten und zusammen mit weiteren Prüfungsmaterialien wie z.B. Röntgenbilder, EKG-Ableitungen und Lungenfunktionen am Veranstaltungstag mitzubringen. Die Fallvignetten wurden zunächst im Rahmen des Workshops fachintern einem Review unterzogen, um später als Prüfungsfragen für die anschließende Prüfungssimulation Verwendung zu finden.

Der Workshop wurde seit 2013 sechsmal mit insgesamt 93 TeilnehmerInnen aus den unterschiedlichsten Fachrichtungen durchgeführt und im Anschluss daran evaluiert. In diesem Zeitraum wurde er mehrmals leicht modifiziert und an die jeweilige TeilnehmerInnenzahl und Facharztgruppenzusammensetzung angepasst.

Am Ende jedes PrüferInnen-Workshops wurde eine TeilnehmerInnen-Evaluation durchgeführt. Der Evaluationsbogen umfasste 12 Items zur Nützlichkeit der einzelnen Workshop-Inhalte für die Facharztprüfung sowie zur Zufriedenheit mit dem Workshop-Angebot und es gab die Möglichkeit, Freitextkommentare zu folgenden Fragen abzugeben:

- Haben Sie Verbesserungsvorschläge für Formulare/Materialien?
- Zu welchen Aspekten wünschen Sie weitere Informationen/Unterstützung?
- Was ist für Sie die wichtigste Erkenntnis aus dieser Veranstaltung?

Die Auswertung der Freitextkommentare erfolgte im ersten Schritt durch Sortieren und Kategorisieren der freien Antworten. Danach wurden die Freitextkommentare von 4 verschiedenen BewerterInnen unabhängig voneinander den vorgegebenen Kategorien zugeordnet.

## Ergebnisse

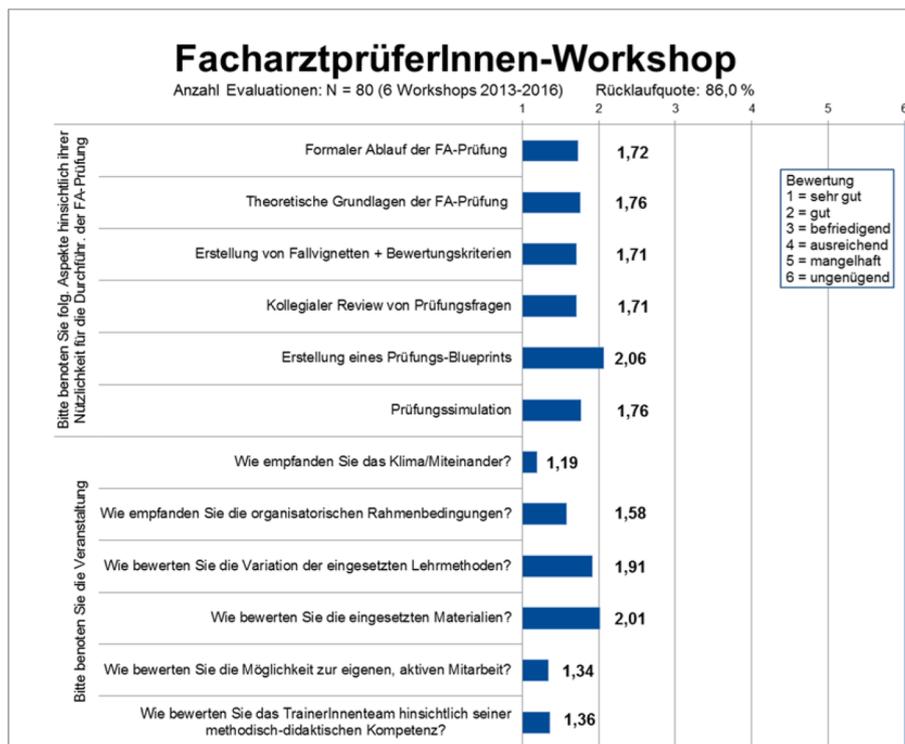
Von den 449 eingetragenen PrüferInnen der Ärztekammer Südbaden sind nur ca. 250 aktiv, von den 33 Prüfungsvorsitzenden nur 10. Von diesen 260 aktiven PrüferInnen und Prüfungsvorsitzenden haben bisher 93 (36%) an der Schulung teilgenommen. Davon gaben 42 TeilnehmerInnen (52,5%) mehr als 5 Jahre Prüfungserfahrung an, 10 (12,5%) zwischen 3 und 5 Jahren und 12 (15%) zwischen 1 und 2 Jahren. 11 TeilnehmerInnen (13,7%) hatten noch keine Prüfung durchgeführt und 5 (6,3%) machten dazu keine Angaben.

Die fachspezifische Zusammensetzung der PrüferInnen und Prüfungsvorsitzenden in den sechs Schulungsveranstaltungen ist in Tabelle 2 dargestellt. Vor allem beteiligten sich PrüferInnen und Vorsitzende aus den Fächern Innere Medizin, Chirurgie, Allgemeinmedizin, Gynäkologie sowie aus der Notfallmedizin.

Die nach den einzelnen Schulungen durchgeführten Evaluationen (Rücklaufquote: 86,0%) ergaben eine hohe Akzeptanz des Schulungskonzeptes, vor allem die Möglichkeit der aktiven Mitarbeit während des Workshops

**Tabelle 2: Fächerverteilung bei den ausgewerteten Evaluationsbögen (N=80)**

Gebietsbezeichnungen	Anzahl der Evaluationsbögen	Zusatzbezeichnungen	Anzahl der Evaluationsbögen
Innere Medizin (Kardiologie/Angiologie/Phlebologie/Onkologie)	N = 12	Notfallmedizin	N = 8
Chirurgie (Orthopädie/Unfall-/Gefäß-/Allgemeinchirurgie)	N = 11	Akupunktur	N = 4
Allgemeinmedizin	N = 9	Physikalische Therapie	N = 2
Gynäkologie	N = 7	Suchtmedizin	N = 1
Anästhesiologie	N = 5		
Neurologie	N = 5		
Psychiatrie	N = 4		
Mund-, Kiefer-, Gesichtschirurgie	N = 3		
HNO	N = 3		
Dermatologie	N = 1		
Augenheilkunde	N = 1		
Psychosomatik	N = 1		
Keine Angaben	N = 3		



**Abbildung 1: Evaluation der FacharztprüferInnen-Workshops 2013 bis 2016 (N=80)**

sowie das gute kollegiale Klima wurden hervorgehoben. Besonders nützlich für die Facharztprüfung wurden neben dem formalen Ablauf und den theoretischen Grundlagen zur Notwendigkeit einer strukturierten mündlichen Prüfung die Erstellung von Fallvignetten, der kollegiale Review der Prüfungsfragen sowie die Prüfungssimulation angesehen (siehe Abbildung 1).

Als bedeutendstes Ergebnis bei der Auswertung der Freitextkommentare zeigte sich bei der Frage nach der wichtigsten Erkenntnis aus der Veranstaltung, dass die TeilnehmerInnen vor allem Anregungen zur Standardisie-

rung und Strukturierung der Prüfung als sinnvoll erachteten. Ebenso waren Austausch und Abstimmung zwischen den PrüferInnen eines Faches von großer Bedeutung. Explizit wurden z.B. eine nachprüfbar und vergleichbare Strukturierung der Facharztprüfungen genannt, eine Standardisierung der Bewertung der Prüfungsleistung, eine vorherige Abstimmung mit den MitprüferInnen und die Erstellung eines internen Fallvignettenkataloges für das jeweilige Prüfungsfach. Außerdem waren den PrüferInnen und Prüfungsvorsitzenden juristische und formale Aspekte der Facharztprüfung klarer geworden. Viele

langjährige PrüferInnen fühlten sich zudem in ihrem bisherigen Vorgehen bestätigt, konnten aber zusätzliche Anregungen aus dem Workshop für ihre PrüferInnen-Tätigkeit mitnehmen.

Einige Freitextkommentare zu den wichtigsten persönlichen Erkenntnissen aus dem Workshop-Angebot sind in Abbildung 2 nach den drei am häufigsten genannten Themenbereichen (Standardisierung, Strukturierung und Abstimmung bzw. Austausch mit den anderen PrüferInnen eines Faches) exemplarisch aufgeführt.

## Diskussion

Die Evaluationen am Ende der Facharztprüfer-Workshops ergaben, dass die FacharztprüferInnen und Prüfungsvorsitzende/n vor allem die Notwendigkeit sehen, gemeinsame Minimalstandards bzw. ein einheitliches Anforderungsniveau für die jeweiligen Weiterbildungsfächer und damit für die Facharztprüfungen zu entwickeln und vor allem auch neuen FacharztprüferInnen Instrumente an die Hand zu geben, die es ermöglichen, die Objektivität, die Reliabilität und damit auch die Validität der mündlichen Prüfungen zu verbessern [1], [2], [4]. Besonders wichtig erscheinen den PrüferInnen im Hinblick auf die mündliche Prüfung die Notwendigkeit einer guten Strukturierung und Standardisierung. Dies ist vor allem abzulesen an den subjektiven Einschätzungen der Nützlichkeit von Fallvignettenerstellung und kollegialem Review sowie an den vielfachen Freitextkommentaren zu der Frage nach der wichtigsten Erkenntnis aus der Veranstaltung (siehe Abbildung 2).

Zu einem ähnlichen Ergebnis kam eine Ulmer Studie von Öchsner et. al zu Effekten und Nachhaltigkeit des Trainingsworkshops zur M2- Prüfung [10]. Hier gaben 86% der TeilnehmerInnen dieses Workshops an, das Konzept des „strukturierten mündlichen Prüfens“ anzustreben. Auch die Prüfungssimulation als praktische Umsetzung der vorangegangenen Schulungsinhalte und damit ein zentrales Element des Facharztprüferworkshops wird von den PrüferInnen als äußerst nützlich für die Durchführung der Facharztprüfung angesehen.

Eine zweite Evaluation der Facharztprüfungen und die damit verbundene Befragung von PrüferInnen, Prüfungsvorsitzenden und Prüflingen ist geplant, vor allem um folgende Fragen zu beantworten:

- Wurden durch den Workshop Änderungen im Prüfungsverhalten der geschulten PrüferInnen initiiert und wenn ja, welche?

Hier wären als mögliche Kriterien die Erstellung eines gemeinschaftlichen, fachinternen Blueprints sowie die vorherige Festlegung von Bestehenskriterien und -grenzen zu nennen.

- Gibt es Unterschiede zwischen geschulten und nicht-geschulten PrüferInnen in den Facharztprüfungen?

Mögliche Kriterien für hierfür wären, ob vor der Prüfung schriftliche Fallvignetten einschließlich eines zuvor fest-

gesetzten Erwartungshorizontes erstellt wurden, ob diese in der Prüfung auch genutzt wurden und ob die formalen Kriterien für eine gute Fallvignette erfüllt sind.

Die vor Einführung der FacharztprüferInnen-Workshops erhobenen Nichtbestehensquoten in den Prüfungen von <3% zeigten in den letzten 3 Jahren keine wesentlichen Änderungen, so dass geprüft werden muss, ob dieses Kriterium als Qualitätsmarker für eine veränderte Prüfungskultur herangezogen werden kann.

Schaut man sich die veröffentlichten Nichtbestehensquoten bei Prüfungen für Gebiets-, Schwerpunkts- und Zusatzbezeichnungen in den verschiedenen Landesärztekammern im Bundesgebiet an, so zeigt sich, dass diese erheblich divergieren. Nicht alle Landesärztekammern machen diese Zahlen jährlich transparent und die veröffentlichten Nichtbestehensquoten zeigen eine Spannweite von z.B. 3,5% in Hamburg [11] und 6,4% in Westfalen-Lippe [12] bei allen Prüfungen (Gebiets-, Schwerpunkts- und Zusatzbezeichnungen) bzw. von z.B. 2,5% in Hamburg [11] und 7,2% in Westfalen-Lippe [12] bei den Prüfungen ausschließlich für die Gebietsbezeichnung.

## Stärken-Schwächenanalyse

Die Stärken des Projektes liegen in der Tatsache, dass das im Projekt vorgestellte FacharztprüferInnen-Angebot bundesweit das erste überhaupt ist und damit Vorbildcharakter aufweist. So hat die VertreterInnenversammlung der Bezirksärztekammer Südbaden das Thema einer Qualitätsverbesserung in Facharztprüfungen früher als andere Ärztekammern nicht nur angedacht, sondern auch beschlossen, entsprechende Qualifizierungsmaßnahmen für PrüferInnen und Prüfungsvorsitzende anzubieten und diese konkret umzusetzen. Außerdem konnte ein erfolgreiches Workshop-Konzept erarbeitet werden, dass von den TeilnehmerInnen äußerst positiv angenommen wurde.

Eine der Schwierigkeiten dieses Projektes bestand darin, dass die TeilnehmerInnen des Workshops sich aus einer Vielzahl von PrüferInnen aus den unterschiedlichsten Fachgruppen zusammensetzten, so dass es nicht in allen Schulungen möglich war, mindestens 4 PrüferInnen aus den einzelnen Fachgruppen zusammenzubringen (siehe Tabelle 2), damit ein kollegialer Austausch über die Prüfungsziele und -inhalte sowie eine Simulationsprüfung im Rahmen des Workshops möglich war. In den ersten Veranstaltungen war dies noch kein Problem, es konnten immer eine genügende Anzahl von VertreterInnen aus 3 - 4 Fachgruppen eingeladen werden. In den letzten Veranstaltungen gab es häufiger das Problem, dass nur einzelne FachvertreterInnen am Workshop teilnahmen. In diesen Fällen war es notwendig, dass die FachvertreterInnen in der Reviewphase intensiv durch das Trainerteam betreut wurden. In der Simulationsphase konnten sie auch nur eine Beobachtungsfunktion wahrzunehmen.

Eine weiterer Einschränkung besteht darin, dass bisher in Bezug auf die Schulungen noch kein Outcome geprüft werden konnte. Dies ist in der Folge angedacht.

Was ist für Sie die wichtigste Erkenntnis aus dieser Veranstaltung?	
Freitextkommentare zu den häufigsten Kategorien	
Standardisierung der Prüfungen	"Standardisierung ist sinnvoll!" "Es sollen Anstrengungen unternommen werden, um die Bewertung der Prüfungsleistung mehr zu standardisieren" "Festlegung von Muss-Kriterien erleichtert die Beurteilung"
Strukturierung der Prüfungen	"Blueprint - Erstellung sehr nützlich zur Strukturierung und Bewertung" "Instrumente der Strukturierung sind extrem hilfreich für einen unkomplizierten fachspezifischen Dialog" "Ich werde meinen Prüflingspart besser und knapper strukturieren"
Abstimmung/Austausch unter den KollegInnen	"Abstimmung mit dem Co-Prüfer vor der Prüfung über Prüfungsfragen" "Erfahrungsaustausch mit anderen Prüfern" "Kollegialer Konsens erforderlich"

Abbildung 2: Ausgewählte Freitextkommentare zur Frage nach den wichtigsten persönlichen Erkenntnissen aus dem Workshop-Angebot.

## Schlussfolgerung

Ein Angebot von Schulungsprogrammen für FacharztprüferInnen von Gebiets-, Schwerpunkts- und Zusatzbezeichnungen erscheint trotz oder auch gerade wegen der geringen Nichtbestehensquoten sinnvoll und notwendig. Da die Nichtbestehensquoten in den Prüfungen zur Erlangung von Gebiets-, Schwerpunkts- und Zusatzbezeichnungen bundesweit erheblich divergieren, wäre zu überlegen, nicht nur in Südbaden, sondern auch in anderen Kammerbezirken eine Qualifizierung der FacharztprüferInnen anzustreben, so wie es in anderen europäischen Ländern bereits Standard ist. Für eine gemeinschaftliche Qualitätsverbesserung erscheint es sinnvoll, ein einheitliches Konzept für das gesamte Bundesgebiet zu entwickeln und einzuführen.

Unabhängig davon sollten in Deutschland auch Überlegungen angestellt werden, andere Prüfungsverfahren in der ärztlichen Facharztweiterbildung zu etablieren. Um die für den Facharzt notwendigen Kompetenzen zu erfassen, reichen reine wissensbasierte Prüfungen nicht aus. Vielmehr ist die Prüfung des Transfers von Fachwissen in die Praxis von entscheidender Bedeutung [4]. Beispiele aus anderen europäischen Ländern weisen auf eine bessere Kompetenzmessung durch den Einsatz von longitudinalen Prüfungsformaten hin [6]. Daher wäre es sinnvoll, auch für Deutschland ein neues Prüfungsformat für die Facharztprüfungen anzudenken und eine Kombination aus mehreren longitudinalen Prüfungen sowie einer Abschlussprüfung zu implementieren. Ähnliche Überlegungen wurden auch im Rahmen des 2013 veröffentlichten Positionspapieres des Ausschusses Weiterbildung der Gesellschaft für Medizinische Ausbildung angestellt [13]. Zusätzlich könnte auch eine mindestens 90 Minuten dauernde Prüfung am Ende der Facharztweiterbildung, wie in Thüringen bereits eingeführt [7] – angedacht werden.

Die neuen Prüfungsformate könnten die bisher praktizierte minimal halbstündige bis maximal einstündige mündliche Prüfung durch zwei FachärztInnen und eine/n Prüfungsvorsitzende/n ersetzen. Allerdings muss bei allen Veränderungen des Prüfungsformates mit einer Erhöhung der Kosten gerechnet werden, wie in dem Artikel von Flum et. al, 2015 ausdrücklich erwähnt wird [6].

## Anmerkung

Dr. med. Irmgard Streitlein-Böhme, Dr. med. Götz Fabry und Dr. med. Klaus Böhme haben für die Konzeption und Durchführung der FacharztprüferInnen-Workshops Aufwandsentschädigungen von der Bezirksärztekammer Südbaden erhalten.

## Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

## Literatur

1. Van der Vleuten CP. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract.* 1996;1(1):41-47. DOI: 10.1007/BF00596229
2. Wakeford R, Southgate L, Wass V. Improving oral examinations: selecting, training and monitoring examiners for the MRCGP. *BMJ.* 1995;311(7010):931-935. DOI: 10.1136/bmj.311.7010.931
3. Jefferies A, Simmons B, Ng E, Skidmore M. Assessment of multiple physician competencies in postgraduate Training: utility of the structured oral examination. *Adv Health Sci Educ Theory Pract.* 2011;16(5):569-577. DOI: 10.1007/s10459-011-9275-6
4. Bloch R, Hofer D, Krebs R, Schläppi P, Weis S, Westkämper R. *Kompetent prüfen. Handbuch zur Planung, Durchführung und Auswertung von Facharztprüfungen.* Bern, Wien: Universität Bern und Med. Universität Wien; 1999.
5. Landesärztekammer Thüringen. *Tätigkeitsbericht der Landesärztekammer 2014.* Jena: Landesärztekammer Thüringen; 2014. S.23-31. Zugänglich unter/available from: <http://www.laek-thueringen.de>
6. Flum E, Maagard R, Godycki-Cwirko M, Scarborough N, Scherpbier N, Ledig T, Roos M, Steinhäuser J. *Facharztprüfung Allgemeinmedizin – was können wir von den europäischen Nachbarn lernen?* *GMS Z Med Ausbild.* 2015;32(2):Doc21. DOI: 10.3205/zma000963
7. Schickler A, Brüstle P, Biller S. *Mündlich-praktischer Teil des Zweiten Abschnitts der Ärztlichen Prüfung in Freiburg 2012 - Analyse der Notenvergabe zur Überprüfung von Qualitätssicherungsmaßnahmen.* *GMS Z Med Ausbild.* 2015;32(4):Doc39. DOI: 10.3205/zma000981

8. Streitlein-Böhme I, Giesler M, Niebling W, Böhme K, Fabry. Facharztprüferworkshop in Südbaden. Strukturierte Prüfungen nach der Facharztweiterbildung. Ärztebl Baden-Württemberg. 2014;69(4 Suppl):142.
9. Kompetenznetz "Lehre in der Medizin" in Baden-Württemberg. Handbuch zur Vorbereitung der Prüfer auf M3. Freiburg: Kompetenznetz "Lehre in der Medizin" Baden-Württemberg; 2014.
10. Öchsner W, Geiler S, Huber-Lang M. Effekte und Nachhaltigkeit von Trainingsworkshops für den mündlich-praktischen Teil des M2-Examens. GMS Z Med Ausbild. 2013;30(3):Doc36. DOI: 10.3205/zma000879
11. Ärztekammer Hamburg. Tätigkeitsbericht. Hamburg: Ärztekammer Hamburg; 2014. S.21-29. Zugänglich unter/available from: <http://www-aerztekammer-hamburg.org>
12. Ärztekammer Westfalen-Lippe. Bericht des Vorstandes der Ärztekammer Westfalen-Lippe. Münster: Ärztekammer Westfalen-Lippe; 2015. S.48. Zugänglich unter/available from: <http://www.aekwl.de>
13. David DM, Euteneier A, Fischer MR, Hahn EG, Johannink J, Kulike K, Lauch R, Lindhorst E, Noll-Hussong M, Pinilla S, Weih M, Wenneke V. Die Zukunft der ärztlichen Weiterbildung in Deutschland – Positionspapier des Ausschusses Weiterbildung der Gesellschaft für Medizinische Ausbildung (GMA). GMS Z Med Ausbild. 2013;30(2):Doc26. DOI: 10.3205/zma000869

**Korrespondenzadresse:**

Dr. med. Irmgard Streitlein-Böhme  
 Universität Freiburg, Lehrbereich Allgemeinmedizin,  
 Elsässerstr. 2 m, 79110 Freiburg, Deutschland, Tel.: +49  
 (0)761/270-72461, Fax: +49 (0)761/270-72480  
[irmgard.streitlein-boehme@uniklinik-freiburg.de](mailto:irmgard.streitlein-boehme@uniklinik-freiburg.de)

**Bitte zitieren als**

*Streitlein-Böhme I, Niebling W, Fabry G, Böhme K. Three years of experience with a workshop for medical specialty examiners in South Baden: A project report and initial evaluation results. GMS J Med Educ. 2017;34(5):Doc55.*  
 DOI: 10.3205/zma001132, URN: [urn:nbn:de:0183-zma0011323](http://nbn:de:0183-zma0011323)

**Artikel online frei zugänglich unter**

<http://www.egms.de/en/journals/zma/2017-34/zma001132.shtml>

**Eingereicht:** 15.11.2016

**Überarbeitet:** 19.06.2017

**Angenommen:** 07.08.2017

**Veröffentlicht:** 15.11.2017

**Copyright**

©2017 Streitlein-Böhme et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.