

The prognostic validity of the formative for the summative MEQ (Modified Essay Questions)

Abstract

Objective: The purpose of formative examinations is that students and lecturers receive an early feedback regarding the success of learning behavior and teaching methods. These also serve as practice for later summative exams. The aim of this paper is to investigate to what extent the result of the formative MEQ* at the end of the first semester at Witten/Herdecke University (UW/H) in the study program human medicine can be used as a predictor for the summative MEQ-1 at the end of the second semester which is part of the equivalence examination replacing the state examination.

Methodology: The predictive value of the score achieved in the MEQ* on the MEQ-1 score, as well as the potential influence of the variables gender, age, high school graduation grade (German Abiturnote), professional background, and self-efficacy expectancy, was determined for students of human medicine.

Results: Data from two cohorts of UW/H with a total of 88 students were included. Scores on the formative MEQ* correlate with those on the summative MEQ-1 in both cohorts. In regression analyses, only the score on the MEQ* proves to be a significant predictor of performance on the MEQ-1 (40.5% variance explanation). Particularly significant predictors are the scores in the subjects anatomy and clinical reasoning. Vocational training or pre-study only appear to contribute to higher scores in the MEQ* after the first semester, but have no further significance in predicting scores in the MEQ-1.

Conclusion: The MEQ* was confirmed to be a good predictor of the MEQ-1. Thus, it serves as a formative exam to inform students about their current state of knowledge with regard to the summative exam MEQ-1, so that they can adequately adapt their learning strategies in the course of the second semester.

Keywords: Modified Essay Questions (MEQ), formative testing, summative testing, predictors of academic success, constructive alignment

Oliver Büssing¹

Jan P. Ehlers²

Michaela Zupanic³

1 Klinikum Westfalen, Hellmig Hospital Kamen, Medical Clinic I - Clinic for Angiology, Cardiology and Intensive Care Medicine, Kamen, Germany

2 Witten/Herdecke University, Faculty of Health, Chair Didactics and Educational Research in Health Care, Witten, Germany

3 Witten/Herdecke University, Faculty of Health, Interprofessional and Collaborative Didactics in Medical and Health Professions, Witten, Germany

Introduction

Since a university course of study is associated with special challenges, it is of particular interest to students and lecturers to find out at an early stage whether learning behavior and teaching methods are goal-oriented. According to the constructive alignment concept, teaching content, learning outcomes and examination formats must be coherently related to each other [1], [2]. Besides the quality criteria of objectivity and reliability, the basic prerequisite for a valid examination is therefore testing of the content that has previously been defined in the learning outcomes [3]. Through examinations during the course of study, students receive regular feedback on the current state of their knowledge. In this context, formative testing aims to monitor and test a (learning) program that is still in progress [4], [5]. The individual feedback provides valuable guidance for the students'

learning plan and the opportunity to reflect on their learning strategies. Formative testing can thus support intrinsic motivation [6]. Summative tests, on the other hand, are used to assess competence, evaluate an already completed (learning) program, and entitle students to more advanced educational segments [4]. However, they may not provide sufficient feedback to specifically support student's learning, but can influence and direct it, in the sense of extrinsic learning motivation. This is often socially reinforced by fellow students [7]. An overview of the examinations used in medical education can be found in Epstein [8] and Schuwirth and van der Vleuten [9] and for the DACH countries in Thiessen et al. [10]. Written exams can be roughly divided into two groups: closed-ended response, such as multiple choice questions (MCQ), and open-ended response, such as free-response essays [11]. As a synthesis of these two basic examination formats, the first MEQ (Modified Essay

Questions) was developed as an examination for general practitioners in the UK in 1971 [12], [13]. MEQ (Modified Essay Questions) is intended to combine the reliability and objectivity of an MCQ examination with the validity of an essay. The tasks in the MEQ consist of a clinical case example, which is first described using a brief outline with a reason for treatment. The following questions, which build upon each other, have to be answered using short free texts [11]. This shall provide both a high level of cognitive challenge and a largely standardized correction of the questions. Free-text tasks are well suited for testing clinical reasoning, but at the same time require a relatively high effort in the formulation of the answer horizons and score distribution and in the correction [14], [15]. The content of the tasks relates to the medical activities

1. history taking, diagnosis, and therapy,
2. differential diagnostic thinking and problem-solving strategies, and
3. holistic thinking and judgment [16].

In order to adequately master the clinical decision-making process simulated in the case study of a real situation, students must actively reproduce or apply their knowledge [17]. An MEQ can test knowledge with regard to the modified competence levels according to Bloom [18] not only at level I (factual knowledge), but also at levels II (conceptual knowledge) and III (procedural knowledge). At Witten/Herdecke University (UW/H), problem-based learning (POL) was introduced in 1992 in the then reformed course of study in human medicine and established in 2000 as an interdisciplinary concept in the first four semesters [16]. The musculoskeletal system is the superordinate topic in the first semester, followed by internal organs (metabolism, cardiovascular system, respiration, fluid and electrolyte balance, hormones) in the second semester, nervous and sensory systems in the third semester, and finally reproduction, blood and immune systems in the fourth semester. Working on patient cases together in the POL tutorials with six students, one medical tutor and one student co-tutor is useful for the acquisition of basic science and clinical knowledge and problem solving skills [19]. This method is considered both motivating [20] and supportive to establish interdisciplinary thinking [21], communication skills, independent sustained learning and understanding of ethical aspects of the healthcare system [22].

§ 5 of the examination regulations of the UW/H for the model course of studies in human medicine [23] provides for summative equivalence examinations for the first section of the medical examination (preliminary medical exam, M1) according to § 41 para. 2 no. 3 of the Regulations for the Licensing of Physicians (ÄAppO) [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] within the framework of a special regulation for model courses of studies. For this purpose, three written and two combined examinations are taken. The two combined examinations are Objective Structured Clinical Examinations (OSCE) [24], while the three written

examinations comprise tasks in the free-text format MEQ at the end of the second semester (MEQ-1), third semester (MEQ-2) and fourth semester (MEQ-3). The choice of these examination formats for equivalence exams reflects the central significance of the POL learning format in accordance with the concept of constructive alignment [10], [17], [25]. To prepare for these summative exams, a formative exam is taken, i.e. the MEQ*. As known from informal discussions, many students do not specifically prepare for this, since the MEQ* is not part of the overall grade of the equivalence examination.

Therefore, the aim of this study was to clarify to what extent the formative examination MEQ* at the end of the 1st semester in the study program of human medicine at UW/H can serve as a predictor for the summative examination MEQ-1, which is part of the equivalence examination replacing the state examination at the end of the second semester. Here, the following potential influencing variables of the exam performance were considered: Age, gender, high school graduation grade, professional background, and self-efficacy expectancy. Self-efficacy from Bandura's social cognitive theory of human agency represents an essential motivational component, namely a person's internal personal belief that he/she can make substantial contributions [26], [27]. Accordingly, self-efficacy also plays an important role in learning and developing new competencies [28], especially in problem-based learning [29].

It can be assumed that, on the one hand, people with good problem-solving and learning strategies, mapped by their high school graduation grade [30], [31], [32] and their own self-efficacy expectations [26], [33] are more likely to achieve better results and, on the other hand, people with subject-specific prior knowledge, in the sense of previous professional knowledge and associated higher age [34], might also have an advantage.

Participants and methods

Participants

We investigated students of human medicine at Witten/Herdecke University in the fall semester 2017/18 (cohort 45, N=44) and the spring semester 2018 (cohort 46, N=44). They were informed about the background of the study and provided their written consented. Data were used in anonymous form in accordance with the provisions of the Federal Data Protection Act (DSGVO) and the ethical standards of the Declaration of Helsinki [35]. The procedure was approved by the Ethics Committee of the UW/H (No. 39/2018). The personal data available in the electronic campus management system of the university (gender, age, high school graduation grade, and professional background in the healthcare sector, operationalized as professional training, previous studies in biology or biochemistry, number of internships, and/or a voluntary social year) were used with the consent of the students.

Table 1: Distribution of the points that can be achieved in the formative MEQ* and summative MEQ-1 by subjects (score, percentage)

	MEQ* fall semester 2017/18 (N=44) or spring semester 2018 (N=44)	MEQ-1 year 45 (N=44) fall semester 2017/18	MEQ-1 year 46 (N=44) spring semester 2018
Anatomy	46 (39.7%)	76 (32.3%)	81 (31.7%)
Physiology	36 (31.0%)	67 (28.5%)	74 (29.0%)
Clinical reasoning	13 (11.2%)	39 (16.6%)	45 (17.6%)
Biology	10 (8.6%)	7 (3.0%)	-
Biochemistry	6 (5.2%)	36 (15.3%)	30 (11.8%)
Chemistry	3 (2.6%)	-	-
Radiology	2 (1.7%)	8 (3.4%)	16 (6.3%)
Pharmacology	-	2 (0.8%)	5 (2.0%)
Immunology	-	-	4 (1.6%)
Maximum score	116	235	255
Passing score (60%) ¹	70	141	153

¹ Passing score which is relevant for summative MEQ-1 and serves for orientation for MEQ*.

Survey methodology

The tasks of the Modified Essay Questions (MEQ) comprise a case study with several related, sequential questions, which have to be answered in a structured way in short free texts [11]. As students work through the case, they receive new information on each page of the exam that might answer the questions on the previous page. Therefore, turning back pages is not allowed in the paper-pencil version of the MEQ [17], [36]. Edited answer sheets must be placed in a separate envelope so that they cannot be subsequently corrected.

Medical students of the UW/H take the formative MEQ* at the end of the first semester under examination conditions. It is adapted to the level of knowledge and structured like a summative MEQ. The content of the formative MEQ* was the same in the fall semester 2017/18 and the spring semester 2018 and the scope was reduced to two patient cases and a few additional free questions. It consisted of two case histories on acute cholecystitis and traumatic shoulder dislocation with a total of 17 case-related questions on specific subjects and 5 free questions (#22 questions). A total of 116 points could be obtained. The students had 90 minutes to complete the test. Immediately afterwards, the exam was corrected within 75 minutes using the student peer review process based on an answer horizon provided by the Examination Office of the Dean of Students Office. Thus, students receive immediate feedback on their performance and can align their individual learning strategies on the assessment of the learning outcome in accordance with constructive alignment [2]. To verify the points assigned by the fellow students and the quality of this feedback, all students' answer sheets were again evaluated by a qualified author (OB) and the scores from the peer review process adjusted accordingly.

Due to an increased number of questions on the topics internal organs and musculoskeletal system, students

had six hours to complete the summative MEQ-1. MEQ-1 for year 45 in the fall semester 2017/18 comprised five case histories with 31 case-related questions and 13 free-response questions (#44 questions), while year 46 in the spring semester 2018 comprised five case histories with 36 case-related questions and 9 free questions (#45 questions). Students' scores for each task in the summative MEQ-1 were provided by the UW/H Dean of Human Medicine's Office of Examinations. Overall, students were able to score a maximum of 235 (#45) or 255 (#46), respectively (see table 1).

In the present study, students completed the General Self-Efficacy Expectation (SWE) scale according to Schwarzer and Jerusalem before starting the exam with the formative MEQ* to assess the influence of the students' own competence expectation to deal with difficult situations [37]. The 10 items of the four-point Likert scale with the same polarity were answered with the answer options (1) not true, (2) hardly true, (3) rather true, and (4) true exactly, and added up for the sum value. For example, one of the SWE items is "I always succeed in solving difficult problems when I try." The SWE scale has good internal consistency in German samples, ranging from Cronbach's alpha=.80 to .90 [38]. Results on the validity are available from empirical studies showing theory-consistent positive correlations with dispositional optimism and job satisfaction, and close negative correlations with anxiety, depression, and burnout [33].

Statistical analyses

During the preliminary analyses, the sociodemographic variables age at baseline, gender and high school graduation grade, as well as the dependent variables total scores in the formative MEQ* and summative MEQ-1 for years 45 and 46, and self-efficacy expectancy were tested for normal distribution using the Kolmogorov-Smirnov test. As the assumption of normal distribution

Table 2: Comparison of results in formative MEQ* (mean, standard deviation) by peer review of medical students and professional evaluation by an author (OB)

	MEQ* peer review year 45 (N=44)	Technical evaluation author (OB)	MEQ* peer review year 46 (N=44)	Technical evaluation author (OB)
	M ± SD	M ± SD	M ± SD	M ± SD
Anatomy	31.48 ± 6.72	29.78 ± 6.88	31.06 ± 6.27	29.73 ± 6.40
Physiology	21.48 ± 7.09	20.07 ± 6.98	20.99 ± 6.29	19.05 ± 6.13
Clinical reasoning	10.14 ± 1.90	9.57 ± 1.75	10.81 ± 1.89	10.61 ± 1.72
Biochemistry	3.99 ± 1.74	3.91 ± 1.71	3.07 ± 1.52	2.80 ± 1.36
Radiology	1.36 ± .55	1.26 ± .55	1.41 ± .52	1.29 ± .55
Total	70.87 ± 13.65	66.89 ± 13.92	69.77 ± 12.44	65.37 ± 13.11

could not be confirmed, non-parametric testing was performed. For comparisons between the independent groups of students (year 45 vs 46), Mann-Whitney U tests were performed. Test size was analogously converted to Cohen's d effect size [39], [40]. Associations between variables were calculated by means of correlation analyses (Spearman rho) with correlation coefficient r as effect size. Multiple regression analyses were used to predict the score obtained on the MEQ-1 by the variables age, high school graduation grade, prior knowledge (occupation, studies), self-efficacy expectation, and MEQ* score. The significance level was set at $p < .05$ for these correlation and regression analyses performed with SPSS 26, and at $p < .01$ for the Mann-Whitney U tests after Bonferroni correction of the α -error [41].

Results

Description of the sample and group comparison

88 medical students from two cohorts of the UW/H (45 and 46) were included in this study. With regard to the variables age at beginning of the study, gender distribution, high school graduation grade as well as completed vocational training, pre-study, internships or voluntary social year and self-efficacy expectations, the two cohorts did not differ significantly and could therefore be considered as one sample for further analysis.

On average, students (46 females, 42 males) were 22.3 ± 2.5 years old (ranging from 19 to 29 years) at the beginning of the study. The high school graduation grade was 2.1 ± 0.4 (ranging from 1.2 to 3.2), with females having a better high school graduation grade than males (1.96 ± 0.42 versus 2.15 ± 0.39 ; Mann-Whitney-U=700, $p=.026$, $d=.488$), which was not significant after α -correction. Prior to the start of the study, 36 students (40.9%) had completed a voluntary social year or learned a profession in the healthcare sector, e.g. general healthcare and nursing, physiotherapy or paramedic. Overall, 80% of the professions in this sample were in the medical field. 15 students (17%) had been engaged in a pre-study program, e.g., biology, nursing, or dentistry. In addition

to the six-month nursing internship required to enter the study program, 67% (N=59) of all participants had completed other internships. The average self-efficacy expectancy of the students of 30.2 ± 3.4 corresponds to the reference mean of 29 points [37].

Demographic characteristics and MEQ results

Students scored a total of 66.9 ± 13.9 (year 45 in the fall semester 2017/18) and 65.4 ± 13.1 (year 46 in the spring semester 2018) on the formative MEQ* with a maximum score of 116. This small difference between the two cohorts is not significant. The points awarded by medical students in the peer-reviewed formative MEQ* are not significantly higher than in the objectified post-evaluation by a qualified author (OB) (see table 2). Both total scores are highly correlated with each other (Spearman: $r=.837$, $p=.000$). The closest match is found in the subject anatomy ($r=.933$), followed by biochemistry ($r=.904$), physiology ($r=.834$), radiology ($r=.775$) and finally clinical reasoning ($r=.463$) (results of all correlation analyses $p=.000$). For further analyses, the results in the formative MEQ* as determined by the qualified author (OB) were used.

In the summative MEQ-1, year 45 achieved an average of 160.8 ± 29.3 points out of 235 maximum possible points (corresponding to 68.4%) and year 46 achieved an average of 193.8 ± 29.6 points out of 255 possible points (corresponding to 76%). There are no significant differences between the cohorts in the total points achieved after their z-transformation (U=909, $p=.753$, $d=.105$).

The assessment of a potential relationship between age at the beginning of the study and MEQ results revealed that there is a moderate positive correlation ($r=.306$, $p=.004$) to the formative MEQ* and to the same extent one semester later to the summative MEQ-1 ($r=.307$, $p=.004$). In addition, there is a positive correlation between age at the beginning of the study as well as the high school graduation grade ($r=.341$, $p=.001$) which is due to the waiting semesters until the beginning of the study. There are no statistically significant gender-specific differences in the total score achieved on the MEQ* and MEQ-1. Moreover, high school graduation grades and

Table 3: Group comparison on professional knowledge (undergraduate or professional) and results in formative MEQ* and summative MEQ-1 (mean, standard deviation)

	MEQ* fall semester 2017/18 or spring semester 2018	MEQ-1 year 45 fall semester 2017/18	MEQ-1 year 46 spring semester 2018
	M ± SD	M ± SD	M ± SD
Pre-study: yes (N=15)	73.41 ± 9.46	176.56 ± 13.82	201.38 ± 15.99
no (N=72)	64.51 ± 13.71	157.16 ± 30.77	192.10 ± 31.82
Profession: yes (N=36)	70.32 ± 10.54	168.87 ± 22.98	203.40 ± 17.52
no (N=52)	63.23 ± 14.54	156.43 ± 31.67	185.00 ± 35.64

Table 4: Predicting outcomes in the formative MEQ* (influencing variables in regression analysis: Age, high school graduation grade, self-efficacy expectancy, vocational training, and pre-study) (N=86) (regression coefficient beta, test value T, significance)

	Beta	T	p
(Constant)		3.649	.000
Age	.230	2.008	.048
High school graduation grade	-.232	-2.264	.026
SWE	-.117	-1.186	.239
Profession	.223	2.008	.048
Study	.216	2.069	.042

self-efficacy expectations show no significant correlation to the score in the formative MEQ* or summative MEQ-1 (results of all correlation analyses $p > .050$).

Consideration of the variables related to students' professional experience did not reveal any significant differences for internships or a voluntary social year, respectively, with respect to the total score in the formative MEQ* and summative MEQ-1. However, in case of a completed pre-study (N=15), the points achieved in the formative MEQ* were higher than without pre-study ($U=363$, $p=.021$, $d=.437$). However, this difference is not significant, nor is it significant one semester later for scores on the summative MEQ-1. Students with a completed vocational training (N=36) scored higher on the MEQ* ($U=685$, $p=.033$, $d=.466$), as well as one semester later with significant difference on the MEQ-1 ($U=615$, $p=.009$, $d=.607$) (see table 3).

Predictors of MEQ results

In multiple regression analyses, all influencing factors were considered together as potential predictors of scores achieved on the formative MEQ* and summative MEQ-1 (each as a dependent variable). Assessment of the predictors for the MEQ* revealed that the variables age, high school graduation grade, vocational training, and pre-study recorded during the admission process were significant predictors of performance on the MEQ*, while self-efficacy expectancy was not (see table 4). However, this model explains only 23.1% of the variance, leaving other unknown variables to account for performance. The common variance components of the variables age, vocational training, and pre-study were differentiated in a stepwise regression analysis. In this analysis, age alone

($\beta=1.74$, $T=3.23$, $p=.002$) accounted for 10.8% of the variance.

The consideration of age, high school graduation grade, vocational training, pre-study, and self-efficacy expectations in the regression analysis revealed that these were not significant predictors of the total score on the MEQ-1. In contrast, the score from the formative MEQ* alone was able to resolve 40.5% of the variance ($\beta=1.58$, $T=7.61$, $p=.000$), and 44.4% in the joint model with the aforementioned variables (see table 5). The significant correlation of positive proportionality between the formative MEQ* and the summative MEQ-1 in year 45 ($r=.769$, $p=.001$) and year 46 ($r=.684$, $p=.001$) is a proof of content validity [42].

For a differentiated analysis, the five medical subjects identified in both the formative MEQ* and the summative MEQ-1 (see table 1) are considered as independent variables in the regression analyses. Scores in the subjects of anatomy, physiology, clinical reasoning, biochemistry, and radiology in the MEQ* explained 53.5% of the variance in the MEQ-1 (see table 6). Except for radiology, all proved to be significant predictors. Results in anatomy alone explained 36.2% of the variance and together with clinical reasoning explained another 11.2% (47.4% total). In contrast, physiology and biochemistry only have a weak predictive effect.

Table 5: Predicting outcomes in the summative MEQ-1 (influencing variables in regression analysis: Age, high school graduation grade, self-efficacy expectancy, vocational training, pre-study and formative MEQ*) (N=88) (regression coefficient beta, test value T, significance)

		Beta	T	p
	(Constant)		.335	.739
	Age	.005	.051	.960
	High school graduation grade	.081	.888	.377
	SWE	.128	1.478	.143
	Profession	.110	1.120	.266
	Study	.019	.202	.841
	MEQ*	.636	6.627	.000

Table 6: Predicting outcomes in the summative MEQ-1 (influencing variables in regression analysis: subjects of the formative MEQ*) (N=88) (regression coefficient beta, test value T, significance)

		Beta	T	p
	(Constant)		1.870	.065
	MEQ* - Anatomy	.442	4.704	.000
	MEQ* - Physiology	.199	2.201	.031
	MEQ* - Clinical reasoning	.321	4.048	.000
	MEQ* - Biochemistry	-.179	-2.056	.043
	MEQ* - Radiology	.155	1.840	.069

Discussion

The purpose of this study was to determine whether a formative examination such as the MEQ* in the model course of study in human medicine at Witten/Herdecke University can serve as a predictor for the summative examination MEQ-1 one semester later. In addition, high school graduation grade, professional experience, self-efficacy expectancy, age, and gender were considered as possible influencing variables.

Gender, high school graduation grade, and self-efficacy expectancy

Female students did not perform better than male students on both the MEQ* and MEQ-1, despite a gender-related difference in the high school graduation grade that narrowly failed to demonstrate statistical significance. This observation is similar to that among medical students at Heidelberg University [32]. The second state exam in fall 2018 also showed comparable exam performance for all medical students in Germany with female students achieving 79.0% versus male students achieving 79.2% of the total score [42]. In their meta-analysis on the prediction of academic success, Trapmann et al. [30] reported that in dental and veterinary medicine programs, the validity of the high school graduation grade was higher for pre-clinical than for clinical semesters. The same applies for human medicine with some 23% clarification of the variance in performance by previous academic performance at the beginning of the study and a total of about 9%, as demonstrated by Ferguson et al. [43] in a systematic literature review. Thus, previous academic experience is a good but not perfect predictor of perform-

ance in medical education. Moreover, gender-specific performance differences were rather small and only reach statistical significance in large cohorts. Thus, even in the present study, with moderate prediction for the formative MEQ*, there was no significant relationship between the high school graduation grade and performance in the formative MEQ* or in the summative MEQ-1. The same applies for the students' self-efficacy expectations [38], thus not confirming the findings of Klassen & Klassen [28] and especially Demirören et al. [29] on the effect in the context of problem-based learning.

Age and previous professional training

The total score in the formative MEQ* shows a moderate positive correlation with age for year 46 (spring semester 2018), which is also positive for year 45 (fall semester 2017/18), but fails to reach statistical significance. The positive age effect is also still detectable one semester later for the MEQ-1. Older students had acquired professional experience in their waiting semesters through a pre-study, vocational training in the healthcare sector, specific internships or a voluntary social year (FSJ). However, internships and FSJ had no relevant influence on the results in the formative MEQ* or summative MEQ-1, as it seems that the acquired practical knowledge without a structuring theoretical foundation does not provide an advantage for the academic performance required in the exams. In contrast, students with pre-study (17% of the sample) performed better in the formative MEQ* in group comparison and achieved higher scores in physiology. Students with previous professional training (40.9% of the sample) also scored higher on the formative MEQ*, especially in the subject of anatomy. Medical students appear to benefit from both types of previous

education: firstly from the structural, subject-independent previous education of a pre-study, in which independent university learning is learned, and secondly from the theoretical-practical previous education of a vocational training in the healthcare sector (80% of the sample) associated with content knowledge from vocational school and practical experience. However, this advantage seems to apply only until the end of the first semester for the formative MEQ*, not until the summative MEQ-1 that is taken one semester later. In terms of previous academic experience, this corresponds to the results of Ferguson et al. [43]. Parallels can also be found in a study by Grendel et al. [34] examining the effects of professional experience of vocationally qualified students. The duration and relevance of previous work experience seem to have a significant influence on the academic success.

Formative MEQ* and peer review process

The predictive significance of the formative MEQ* score for subsequent performance on the summative MEQ-1 was confirmed in the regression model with a 40.5% variance explanation. Moreover, correlation analyses showed a high content validity. Regarding the significance of specific subjects in the formative MEQ*, the particular importance of anatomy and clinical reasoning becomes clear. Traditionally, anatomy is the most challenging subject to learn at the beginning of medical studies, and this is also the case with the superordinate topic musculoskeletal system in the first semester of the model course in human medicine at Witten/Herdecke University. Clinical reasoning is trained by means of Problem-Oriented Learning (POL) and can be best tested using MEQ free-text questions according to constructive alignment [17], [18]. The points awarded in the formative MEQ* in peer review are not significantly higher as compared to those in the subsequent correction by the professionally qualified author (OB). Thus, despite the rather favorable evaluation by fellow students, student peer review seems to be an efficient method to provide quick feedback in a formative examination. The high correlation between peer review and objectified post-evaluation proves the reliability of the MEQ*, which is highest for anatomy and – as expected – lowest for clinical reasoning [14], [15], [44]. The problem, however, is that some students do not prepare for the formative exam. If the formative MEQ* had a pass mark of 60% [45], as is the case for the first state examination according to § 14 of the Regulations for the Licensing of Physicians (ÄAppO, 2002), more than half of the students would not pass. The average score achieved corresponds to 57% and is thus clearly above the average 30% score achieved by psychology students at the UW/H in the formative Progress Test Psychology [46], but demonstrates that students do not yet sufficiently realize the purpose and benefit of formative tests as feedback about their own level of knowledge, for learning motivation and reduction of test anxiety [47].

Limitations

The small sample size of 88 medical students from two semesters should be taken into account. Moreover, they represent a heterogeneous group with regard to age and previous professional experience. In this respect, our findings can probably not be generalized, since they could be a matter of random, cohort-specific effects. Although the peer-reviewed results of the formative MEQ* were not significantly different in the post-correction by the professionally qualified author (OB), post-correction was difficult, as the students' responses did often not correspond to the response horizon. Thus, it was at the discretion of the post-correction staff how to score the answers given. The risk here is that different reviewers will arrive at different assessments of the answers which will reduce the reliability of the formative MEQ*. This is a fundamental problem of free text formats and requires a relatively high effort, both in the formulation of the response horizons and in the review process and correction [14], [15].

Conclusions

The present investigation showed that there was a significant increase in knowledge during the second semester with year 45 (fall semester 2017/18) achieving an average score of 68% in the summative MEQ-1 and year 46 (spring semester 2018) 76%, respectively. The performance increase as compared to the result in the formative MEQ* is probably due to the systematic exam preparation, with approximately the same learning concept of POL tutorials during the first and second semester. The MEQ* as a formative test seems to give students feedback on their current knowledge level. Failure to pass could serve as a “wake-up call” for many students to intensify their learning efforts for the summative MEQ-1. Parallels can be found in Heeneman's interview study [48], in which students confirm that they use the results analysis of the formative progress test to adjust their learning strategies.

The fact students find out about the free text format in the formative MEQ* might also help them to adapt their learning style and exam preparation in accordance with the concept of constructive alignment. Thus, the goal of formative testing would be achieved by the MEQ* and the effort associated with the development and implementation of this type of formative testing would be justified for both lecturers and dean's office at Witten/Herdecke University.

Outlook

After it was demonstrated that the outcome of the formative MEQ* at the end of the first semester is a significant predictor for the outcome of the summative MEQ-1 at the end of the second semester, one relevant objective for further investigations would be to investigate for which

period in the course of the study the MEQ* retains its function as a predictor. Does the performance in the formative MEQ* exam also say something about the overall summative equivalence exam (M1, preliminary medical exam)? It is quite interesting that – with respect to the subjects tested in the MEQ* – anatomy and clinical reasoning are paramount for the score in the summative MEQ-1. Since the performance in the subject of clinical reasoning seems to be unrelated to the overall scores of the other subjects, the teaching concept of problem-based learning has a particularly important role to play in this acquisition of competencies. From the low average score and the individual student statements it can be deduced that some students probably only prepare specifically for the formative MEQ*. In order to increase the benefit of this formative examination in its function as a useful feedback, a thorough exam preparation by of the students would be desirable. It would be important to investigate in further studies to what extent feedback from the formative MEQ* is actually used by the students to change their learning strategies.

Competing interests

The authors declare that they have no competing interests.

References

- Biggs J. Enhancing teaching through constructive alignment. *High Educ.* 1996;32:347-364. DOI: 10.1007/BF00138871
- Tang C, Biggs J. Teaching for Quality Learning at University. 4th Ed. Buckingham: Open University Press/McGraw Hill; 2011.
- Wildt B, Wildt J. Lernprozessorientiertes Prüfen im "Constructive Alignment". HRK Projekt nexus. Berlin, Bonn: HRK; 2011. Zugänglich unter/available from: <https://www.hrk-nexus.de/fileadmin/redaktion/hrk-nexus/07-Downloads/07-03-Material/pruefen.pdf>
- Maier U. Vergleichsarbeiten im Spannungsfeld zwischen formativer und summativer Leistungsmessung. *Dtsch Schule.* 2010;102:60-69.
- Black P, William D. Assessment and Classroom Learning. *Assessment in Education: Principles. Policy Pract.* 1998;5:7-74. DOI: 10.1080/0969595980050102
- Ben-David MF. The role of assessment in expanding professional horizons. *Med Teach.* 2000;22(5):472-477. DOI: 10.1080/01421590050110731
- Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract.* 2010;15(5):695-715. DOI: 10.1007/s10459-010-9232-9
- Epstein RM. Assessment in medical education. *New Engl J Med.* 2007;356(4):387-396. DOI: 10.1056/NEJMr054784
- Schuwirth LW, van der Vleuten CM. Current Assessment in Medical Education: Programmatic Assessment. *J Appl Test Technol.* 2019;Vol 20(S2):2-10.
- Thiessen N, Fischer MR, Huwendiek S. Assessment methods in medical specialist assessments in the DACH region - overview, critical examination and recommendations for further development. *GMS J Med Educ.* 2019;36(6):Doc78. DOI: 10.3205/zma001286
- Hift RJ. Should essays and other open-ended-Type questions retain a place in written summative assessment in clinical medicine? *BMC Med Educ.* 2014;14:249. DOI: 10.1186/s12909-014-0249-2
- Knox JD. What is a Modified Essay Question? *Med Teach.* 1989;1:51-s7. DOI: 10.3109/01421598909146276
- Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ.* 2007;7:49. DOI: 10.1186/1472-6920-7-49
- Feletti GI, Smith EK. Modified Essay Questions: Are they Worth the Effort? *Med Educ.* 1986;20:126-132. DOI: 10.1111/j.1365-2923.1986.tb01059.x
- Nendaz M, Tekian A. Assessment in problem-based learning medical schools: A literature review. *Teach Learn Med.* 1999;11(4):232-243. DOI: 10.1207/S15328015TLM110408
- Butzlaff M, Hofmann M, Brunk C, Edelhäuser F, Lutz G, Reißweber J, Scheffer C, Tauschel D, Thiele S, Wirth D, Zupanec M. Der Modellstudiengang Medizin an der Universität Witten/Herdecke - auf dem Weg zur lebenslang lernfähigen Arztpersönlichkeit. In: Fuhrmann M, Gädler J, Kohler J, Schmidt U, editors. *Handbuch Qualität in Studium, Lehre und Forschung.* Berlin: DUZ Verlags- und Medienhaus GmbH; 2014. p.65-104.
- Wild D, Rützler M, Haarhaus M, Peters K. The Modified Essay Question (MEQ) Test at the Medical School of the University Witten/Herdecke. *Gesundheitswesen (Suppl Med Ausbild).* 1998;15(Suppl2):65-69.
- Anderson L, Krathwohl D. A taxonomy for teaching, learning and assessing: A revision of Bloom's taxonomy of educational objectives. London: Pearson Education; 2000.
- Albanese MA, Mitchell S. Problem-based Learning: A Review of Literature on its Outcomes and Implementation Issues. *Acad Med.* 1993;68(1):52-81. DOI: 10.1097/00001888-199301000-00012
- Colliver JA. Effectiveness of Problem-based Learning Curricula: Re-search and Theory. *Acad Med.* 2000;75(3):259-266. DOI: 10.1097/00001888-200003000-00017
- Schlett CL, Doll H, Dahmen J, Polacsek O, Federkeil G, Fischer MR, Bamberg F, Butzlaff M. Job requirements compared to medical school education: differences between graduates from problem-based learning and conventional curricula. *BMC Med Educ.* 2010;10:1. DOI: 10.1186/1472-6920-10-1
- Koh GC, Khoo HE, Wong ML, Koh D. The effects of problem-based learning during medical school on physician competency: a systematic review. *CMAJ.* 2008;178(1):34-41. DOI: 10.1503/cmaj.070565
- Universität Witten/Herdecke. Prüfungsordnung der Universität Witten/Herdecke für den Modellstudiengang Medizin. Witten: Universität Witten/Herdecke; 2018. Zugänglich unter/available from: https://www.uni-wh.de/fileadmin/user_upload/03_G/07_Humanmedizin/Studiengang_HM/links_und_downloads/Pruefungsordnung_MSG.pdf
- Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of Clinical Competence using Objective Structured Examination. *Br Med J.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447

25. Zupanic M, Ehlers S, Bauer D, Hofmann M, Ehlers JP. Was prüft der Modified Essay Questions Test an der Universität Witten/Herdecke? In: Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Wien, 19.-22.09.2018. Düsseldorf: German Medical Science GMS Publishing House; 2018. Doc28.2. DOI: 10.3205/18gma116
26. Bandura A. Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychol Rev.* 1977;84:191-215. DOI: 10.1037/0033-295X.84.2.191
27. Bandura A. Social cognitive theory: an agentic perspective. *Annu Rev Psychol.* 2001;52:1-26. DOI: 10.1146/annurev.psych.52.1.1
28. Klassen RM, Klassen JRL. Self-efficacy beliefs of medical students: a critical review. *Perspect Med Educ.* 2018;7(2):76-82. DOI: 10.1007/s40037-018-0411-3
29. Demirören M, Turan S, Öztuna D. Medical students' self-efficacy in problem-based learning and its relationship with self-regulated learning. *Med Educ Online.* 2016;21:30049. DOI: 10.3402/meo.v21.30049
30. Trapmann S, Hell B, Weigand S, Schuler H. Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Z Päd Psychol.* 2007;21:11-27. DOI: 10.1024/1010-0652.21.1.11
31. McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *BMJ.* 2003;327(7407):139-142. DOI: 10.1136/bmj.327.7407.139
32. Kadmon G, Resch F, Duelli R, Kadmon M. Predictive Value of the School-leaving Grade and Prognosis of Different Admission Groups for Academic Performance and Continuity in the Medical Course - a Longitudinal Study. *GMS Z Med Ausbild.* 2014;31(2):Doc21. DOI: 10.3205/zma000913
33. Schwarzer R, Jerusalem M. Das Konzept der Selbstwirksamkeit. In: Jerusalem M, Hopf D, editors. *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen.* Weinheim: Beltz; 2002. p.28-53.
34. Grendel T, Lübke H, Haußmann I. Effekte der Dauer und der Qualität berufspraktischer Vorerfahrungen auf den Studienerfolg beruflich Qualifizierter. *Beitr Hochschulforsch.* 2014;36:40-63.
35. WMA. Deklaration von Helsinki, Version 2008. Ferner-Voltaire: WMA; 2008. Zugänglich unter/available from: https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/International/Deklaration_von_Helsinki_2013_20190905.pdf
36. Gößling S. Systematische Übersichtsarbeit zur Anwendung schriftlicher Formate, Wissen kompetenzorientiert abzufragen. Dissertation. Witten: Universität Witten/Herdecke, Fakultät für Gesundheit; 2017.
37. Schwarzer R, Jerusalem M. Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Freie Universität Berlin; 1999.
38. Hinz A, Schumacher J, Albani C, Schmid G, Brähler E. Bevölkerungsrepräsentative Normierung der Skala zur Allgemeinen Selbstwirksamkeitserwartung. *Diagnostika.* 2006;52:26-32. DOI: 10.1026/0012-1924.52.1.26
39. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. *J Experim Psychol Gen.* 2012;141(1):30. DOI: 10.1037/a0026092
40. Lenhard W, Lenhard A. Berechnung von Effektstärken. Dettelbach: Psychometrica; 2016. Zugänglich unter/available from: <https://www.psychometrica.de/effektstaerke.html>
41. Victor A, Elsässer A, Hommel G, Blettner M. Wie bewertet man die p-Wert-Flut? Hinweise zum Umgang mit dem multiplen Testen. Teil 10 der Serie zur Bewertung wissenschaftlicher Publikationen. *Dtsch Arztebl Int.* 2010;107(4):50-56. DOI: 10.3238/arztebl.2010.0050
42. Institut für medizinische und pharmazeutische Prüfungsfragen. Ergebnisse des Zweiten Abschnitts der Ärztlichen Prüfung - Herbst 2018. Mainz: IMPP; 2018. Zugänglich unter/available from: <https://www.impp.de/informationen/berichte/ergebnisberichte/medizin.html>
43. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ.* 2002;324(7343):952-957. DOI: 10.1136/bmj.324.7343.952
44. von Bergmann H, Dalrymple KR, Wong S, Shuler CF. Investigating the Relationship between PBL Process Grades and Content Acquisition Performance in a PBL Dental Program. *J Dent Educ.* 2007;71(9):1160-1170. DOI: 10.1002/j.0022-0337.2007.71.9.tb04380.x
45. Möltner A, Schellberg D, Jünger J. Grundlegende quantitative Analysen medizinischer Prüfungen. *GMS Z Med Ausbild.* 2006;23(3):Doc 53. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000272.shtml>
46. Zupanic M, Troche SJ, Ehlers JP. Absolvierendenniveau im formativen Progress Test Psychologie: Anspruch oder Wirklichkeit? In: Krämer M, Preiser S, Brusdeylins K, editors. *Psychologiedidaktik und Evaluation XII.* Aachen: Shaker; 2018. p.359-368.
47. Schüttelpelz-Brauns K, Karay Y, Arias J, Gehlhar K, Zupanic M. Comparison of the evaluation of formative assessment at two medical faculties with different conditions of undergraduate training, assessment and feedback. *GMS J Med Educ.* 2020;37(4):Doc41. DOI: 10.3205/zma001334
48. Heeneman S, Schut S, Donkers J, Vleuten van der C, Muijtjens A. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Med Teach.* 2017;39(1):44-52. DOI: 10.1080/0142159X.2016.1230183
49. Moosbrugger H, Kelava A, editors. *Testtheorie und Fragebogenkonstruktion.* 2. Aufl. Berlin: Springer Verlag; 2012. DOI: 10.1007/978-3-642-20072-4

Corresponding author:

Jun.-Prof. Dr. phil. Michaela Zupanic
Witten/Herdecke University, Faculty of Health,
Interprofessional and Collaborative Didactics in Medical
and Health Professions, Witten, Germany
Michaela.Zupanic@uni-wh.de

Please cite as

Büssing O, Ehlers JP, Zupanic M. The prognostic validity of the formative for the summative MEQ (Modified Essay Questions). *GMS J Med Educ.* 2021;38(6):Doc99. DOI: 10.3205/zma001495, URN: urn:nbn:de:0183-zma001495

This article is freely available from

<https://www.egms.de/en/journals/zma/2021-38/zma001495.shtml>

Received: 2021-03-01

Revised: 2021-06-14

Accepted: 2021-06-28

Published: 2021-09-15

Copyright

©2021 Büssing 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/>.

Die prognostische Validität des formativen für den summativen MEQ (Modified Essay Questions)

Zusammenfassung

Zielsetzung: Damit Studierende und Dozierende frühzeitig ein Feedback bezüglich des Erfolgs von Lernverhalten und Lehrmethoden bekommen, werden formative Prüfungen geschrieben. Diese dienen auch als Übung für spätere summative Prüfungen. Ziel dieser Arbeit ist die Untersuchung, in wie weit das Ergebnis des formativen MEQ* am Ende des 1. Semesters an der Universität Witten/Herdecke (UW/H) im Studiengang Humanmedizin als Prädiktor für den summativen MEQ-1 am Ende des 2. Semesters, der Teil der Staatsexamens ersetzenden Äquivalenzprüfung ist, dienen kann.

Methodik: Der prädiktive Wert der Punktzahl aus dem MEQ* auf die Punktzahl des MEQ-1 sowie der mögliche Einfluss der Variablen Geschlecht, Alter, Abiturnote, berufliches Vorwissen und Selbstwirksamkeitserwartung wurde für die Studierenden der Humanmedizin ermittelt.

Ergebnisse: Einbezogen wurden Daten von zwei Jahrgängen der UW/H mit insgesamt 88 Studierenden. Die Punktzahlen im formativen MEQ* korrelieren für beide Jahrgänge mit denen des summativen MEQ-1. In Regressionsanalysen erweist sich lediglich die Punktzahl im MEQ* als signifikanter Prädiktor der Leistung im MEQ-1 (40.5% Varianzerklärung). Besonders bedeutsame Prädiktoren sind die Ergebnisse in den Fächern Anatomie und Klinisches Denken. Eine Berufsausbildung oder ein Vorstudium scheinen nur zu höheren Punktzahlen im MEQ* nach dem ersten Semester beizutragen, haben aber für die Prädiktion der Punktzahl im MEQ-1 keine weitere Bedeutung.

Schlussfolgerung: Der MEQ* konnte als guter Prädiktor für den MEQ-1 bestätigt werden. Somit kommt dieser seiner Funktion als formative Prüfung nach, die Studierenden über ihren aktuellen Wissensstand im Hinblick auf die summative Prüfung MEQ-1 zu informieren, damit diese ihre Lernstrategien im Verlauf des 2. Semesters adäquat anpassen können.

Schlüsselwörter: Modified Essay Questions (MEQ), formative Testung, summative Testung, Prädiktoren für Studienerfolg, constructive alignment

Oliver Büssing¹

Jan P. Ehlers²

Michaela Zupanic³

1 Klinikum Westfalen, Hellmig-Krankenhaus Kamen, Medizinische Klinik I - Klinik für Angiologie, Diabetologie, Kardiologie und Intensivmedizin, Kamen, Deutschland

2 Universität Witten/Herdecke, Fakultät für Gesundheit, Lehrstuhl Didaktik und Bildungsforschung im Gesundheitswesen, Witten, Deutschland

3 Universität Witten/Herdecke, Fakultät für Gesundheit, Interprofessionelle und kollaborative Didaktik in Medizin- und Gesundheitsberufen, Witten, Deutschland

Einleitung

Da ein Hochschulstudium besondere Herausforderungen mit sich bringt, ist es für Studierende und Dozierende von besonderem Interesse, frühzeitig zu erfahren, ob Lernverhalten und Lehrmethoden zielführend sind. Entsprechend dem „Constructive Alignment“ müssen die Lehrinhalte, Lernergebnisse und Prüfungsformate kohärent aufeinander bezogen werden [1], [2]. Die Grundvoraussetzung für eine valide Prüfung ist demnach, neben den zu erfüllenden Gütekriterien Objektivität und Reliabilität, das Prüfen der zuvor in den Lernergebnissen festgelegten Inhalte [3]. Durch die Prüfungen im Studienverlauf erhalten Studierende eine regelmäßige Rückmeldung über den aktuellen Stand ihres Wissens. Dabei haben formative

Testungen zum Ziel, ein noch laufendes (Lern-) Programm zu begleiten und zu prüfen [4], [5]. Das individuelle Feedback bietet eine Anleitung für die Lernplanung der Studierenden und zugleich eine Möglichkeit zur Reflexion ihrer Lernstrategien. Formative Prüfungen können somit die intrinsische Motivation unterstützen [6]. Summative Testungen hingegen dienen der Beurteilung der Kompetenz, bewerten ein abgeschlossenes (Lern-) Programm und berechtigen zu weitergehenden Bildungsabschnitten [4]. Möglicherweise liefern sie nicht genügend Feedback, um das Lernen von Studierenden gezielt zu unterstützen, können es aber beeinflussen und lenken, i. S. einer extrinsischen Lernmotivation. Diese wird zudem oftmals durch Kommiliton*innen sozial verstärkt [7].

Ein Überblick über die in der medizinischen Ausbildung verwendeten Prüfungen findet sich bei Epstein [8] sowie Schuwirth und van der Vleuten [9] und für die DACH-

Länder bei Thiessen et al. [10]. Schriftliche Prüfungen lassen sich dabei grob in zwei Gruppen einteilen: Closed-Ended Response, wie z. B. Multiple Choice Questions (MCQ), und Open-Ended Response, wie z. B. Freitexte in Form von Essays [11]. Als Synthese dieser beiden grundlegenden Prüfungsformate wurde 1971 der erste MEQ (Modified Essay Questions) als Prüfung für Allgemeinärzt*innen in Großbritannien entwickelt [12], [13]. MEQ (Modified Essay Questions) soll die Kombination der Reliabilität und Objektivität einer MCQ-Prüfung mit der Validität eines Essays ermöglichen. Die Aufgaben im MEQ (Modified Essay Questions) bestehen aus einem klinischen Fallbeispiel, das zunächst anhand einer Vignette mit Behandlungsanlass beschrieben wird. Die folgenden hierauf bezogenen, aufeinander aufbauenden Fragen, müssen in kurzen Freitexten strukturiert beantwortet werden [11]. Dies soll sowohl eine hohe kognitive Herausforderung, als auch die weitgehend standardisierte Korrektur der Fragen ermöglichen. Freitextaufgaben sind für das Prüfen klinischen Denkens gut geeignet, erfordern aber zugleich einen relativ hohen Aufwand bei der Formulierung der Antworthorizonte und Punkteverteilung sowie bei der Korrektur [14], [15]. Inhaltlich beziehen sich die Aufgaben auf die ärztlichen Tätigkeiten

1. Anamneseerhebung, Diagnostik und Therapie,
2. differenzialdiagnostisches Denken und Problemlösestrategien sowie
3. ganzheitlichen Denken und Urteilsvermögen [16].

Um den im Fallbeispiel einer realen Situation nachgebildeten klinischen Entscheidungsprozess adäquat zu bewältigen, müssen die Studierenden demnach ihr Wissen aktiv reproduzieren bzw. anwenden [17]. Ein MEQ (Modified Essay Questions) kann Wissen hinsichtlich der modifizierten Kompetenzniveaustufen nach Bloom [18] nicht nur auf dem Level I (Faktenwissen) abfragen, sondern auch die Niveaustufen II (Konzeptuelles Wissen) und III (Prozedurales Wissen) erfassen.

An der Universität Witten/Herdecke (UW/H) wurde 1992 das Problemorientierte Lernen (POL) in den damaligen Reformstudiengang Humanmedizin eingeführt und 2000 als fächerübergreifendes Konzept in den ersten vier Semestern etabliert [16]. Übergeordneter Themenbereich ist dabei im 1. Semester der Bewegungsapparat, im 2. Semester die inneren Organe (Stoffwechsel, Herz-Kreislauf, Atmung, Flüssigkeits- und Elektrolythaushalt, Hormone), im 3. Semester das Nerven- und Sinnessystem und im 4. Semester Fortpflanzung, Blut- und Immunsystem. Die gemeinsame Bearbeitung von Patient*innenfällen in den POL-Tutorien mit sechs Studierenden, einer/einem ärztlichen Tutor*in und einer/einem studentischen Co-Tutor*in wird dazu genutzt, grundlagenwissenschaftliches und klinisches Wissen sowie die Fähigkeit zur Problemlösung zu erwerben [19]. Diese Methode gilt als motivierend [20] sowie unterstützend bei der Ausbildung von interdisziplinärem Denken [21], kommunikativen Fähigkeiten, selbstständigem anhaltenden Lernen und dem Verständnis von ethischen Aspekten des Gesundheitssystems [22]. Die Prüfungsordnung der UW/H für den Modellstu-

diengang Medizin [23] sieht im Rahmen einer Sonderregelung für Modellstudiengänge in § 5 Summative Äquivalenzprüfungen zum Ersten Abschnitt der ärztlichen Prüfung (Physikum, M1) gemäß § 41 Abs. 2 Nr. 3 der Approbationsordnung für Ärzte (ÄAppO) [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] vor. Dazu werden drei schriftliche und zwei kombinierte Prüfungen durchgeführt. Die zwei kombinierten Prüfungen erfolgen im Format Objective Structured Clinical Examination (OSCE) [24]. Die drei schriftlichen Prüfungen erfolgen mit Aufgaben im Freitextformat MEQ am Ende des 2. Semesters (MEQ-1), des 3. Semesters (MEQ-2) und des 4. Semesters (MEQ-3). Mit der Wahl dieser Prüfungsformate für Äquivalenzprüfungen wird der zentralen Bedeutung des Lernformats POL entsprechend des Constructive Alignment Rechnung getragen [10], [17], [25]. Zur Vorbereitung auf diese summativen Prüfungen wird eine formative Prüfung geschrieben, der MEQ*. Darauf bereiten sich, wie aus informellen Gesprächen bekannt, viele Studierende nicht gezielt vor, da der MEQ* nicht in die Gesamtnote der Äquivalenzprüfung einfließt.

Ziel dieser Arbeit war es daher, zu klären, in wie weit die formative Prüfung MEQ* am Ende des 1. Semesters im Studiengangs Humanmedizin der UW/H als Prädiktor für die summative Prüfung MEQ-1, welche Teil des Staatsexamens ersetzenden Äquivalenzprüfung ist, am Ende des 2. Semesters dienen kann. Hierbei sollten folgende mögliche Einflussvariablen der Prüfungsleistung berücksichtigt werden: Alter, Geschlecht, Abiturnote, berufliches Vorwissen und Selbstwirksamkeitserwartung. Die Selbstwirksamkeit aus Banduras sozialkognitiver Theorie der menschlichen Handlungsfähigkeit stellt eine wesentliche motivationale Komponente dar, nämlich die interne persönliche Überzeugung von Menschen, dass sie substantielle Beiträge leisten können [26], [27]. Demzufolge kommt der Selbstwirksamkeit auch beim Lernen und Entwickeln neuer Kompetenzen eine wichtige Rolle zu [28], insbesondere beim problemorientierten Lernen [29]. Es kann vermutet werden, dass einerseits Personen mit guten Lösungs- und Lernstrategien, abgebildet über die Abiturnote [30], [31], [32] und die eigene Selbstwirksamkeitserwartung [26], [33] eher bessere Ergebnisse erzielen und andererseits auch Personen mit fachspezifischem Vorwissen, i. S. von beruflichem Vorwissen und damit assoziiert höherem Alter [34], ebenfalls im Vorteil sein könnten.

Studienteilnehmende und Methodik

Studienteilnehmende

Untersucht wurden Studierende der Humanmedizin der Universität Witten/Herdecke im Wintersemester 2017/18 (Jahrgang 45, N=44) und Sommersemester 2018 (Jahrgang 46, N=44). Sie wurden über den Hintergrund der Untersuchung informiert und willigten schriftlich ein. Alle Daten wurden, den Bestimmungen des Bundesdaten-

schutzgesetzes und den ethischen Normen der Deklaration von Helsinki [35] entsprechend, in anonymisierter Form verwendet. Das Procedere der Arbeit wurde von der Ethikkommission der UW/H genehmigt (Nr. 39/2018). Die im elektronischen Campus Managementsystem der Universität vorliegenden persönlichen Daten (Geschlecht, Alter, Abiturnote und berufliches Vorwissen im Gesundheitsbereich, operationalisiert als Berufsausbildung, Vorstudium der Biologie oder Biochemie, Anzahl der Praktika und/oder ein Freiwilliges Soziales Jahr) wurden mit dem Einverständnis der Studierenden verwendet.

Erhebungsmethoden

Die Aufgaben des MEQ (Modified Essay Questions) bestehen aus einem Fallbeispiel mit mehreren hierauf bezogenen, aufeinander aufbauenden Fragen, die in kurzen Freitexten strukturiert beantwortet werden müssen [11]. Während die Studierenden den Fall bearbeiten, erhalten sie auf jeder Seite der Klausur neue Informationen, die u. U. die Fragen auf der vorherigen Seite beantworten. Deshalb ist ein Zurückblättern während des MEQ in der paper-pencil-Version nicht gestattet [17], [36]. Bearbeitete Antwortbogen müssen in einen separaten Briefumschlag gelegt werden, so dass sie nicht nachträglich korrigiert werden können.

Den formativen MEQ* schreiben die Medizinstudierenden der UW/H am Ende des 1. Semesters unter Prüfungsbedingungen. Er ist dem Wissensstand angepasst und von der Struktur wie ein summativer MEQ aufgebaut. Der formative MEQ* im Wintersemester 2017/18 und Sommersemester 2018 war inhaltlich identisch und in seinem Umfang auf zwei Patientenfälle und wenige zusätzliche freie Fragen reduziert. Er bestand aus zwei Fallgeschichten zum Thema akute Cholezystitis und traumatische Schulterluxation mit insgesamt 17 fallbezogenen Fragen zu spezifischen Fächern sowie 5 freien Fragen (#22 Fragen). Insgesamt konnten 116 Punkte erzielt werden. Zur Bearbeitung hatten die Studierenden 90 Minuten Zeit. Direkt im Anschluss erfolgte die Korrektur innerhalb von 75 Minuten im studentischen Peer-Review-Verfahren anhand eines vom Prüfungswesen des Studiendekanats zur Verfügung gestellten Antworthorizonts. So erhalten die Studierenden ein unmittelbares Feedback bezüglich ihrer Leistungen und können ihre individuellen Lernstrategien gemäß des Constructive Alignments [2] mit dem Assessment des Learning Outcomes abstimmen. Zur Überprüfung der hier von den Kommiliton*innen vergebenen Punkte und der Qualität dieses Feedbacks wurden die Antwortbögen aller Studierenden nochmals durch einen fachlich qualifizierten Autor (OB) nachträglich ausgewertet und die Punktzahl aus dem Peer-Review-Verfahren angepasst.

Im summativen MEQ-1 hatten die Studierenden aufgrund einer erhöhten Fragenzahl zu den Themenbereichen Innere Organe und Bewegungsapparat sechs Stunden zur Bearbeitungszeit. Der MEQ-1 des Jahrgangs 45 im Wintersemester 2017/18 bestand aus fünf Fallgeschichten mit 31 fallbezogenen Fragen sowie 13 freien Fragen

(#44 Fragen), während Jahrgang 46 im Sommersemester 2018 fünf Fallgeschichten mit 36 fallbezogenen und 9 freien Fragen (#45 Fragen) bearbeitete. Die Ergebnisse der Studierenden für jede Aufgabe im summativen MEQ-1 wurden durch das Prüfungswesen des Studiendekanats Humanmedizin der UW/H zur Verfügung gestellt. Insgesamt konnten die Studierenden maximal 235 (Jahrgang 45) bzw. 255 Punkte (Jahrgang 46) erzielen (siehe Tabelle 1).

In der vorliegenden Studie wurde von den Studierenden zur Erhebung des Einflusses der eigenen Kompetenzerwartung zum Umgang mit schwierigen Situationen vor Beginn der Klausur mit dem formativen MEQ* die Skala zur Allgemeinen Selbstwirksamkeitserwartung (SWE) nach Schwarzer und Jerusalem [37] bearbeitet. Die 10 gleichsinnig gepolten Items der vierstufigen Likert-Skala werden mit den Antwortmöglichkeiten von (1) stimmt nicht, (2) stimmt kaum, (3) stimmt eher und (4) stimmt genau beantwortet und zum Summenwert aufaddiert. Eines der SWE-Items lautet z. B. „Die Lösung schwieriger Probleme gelingt mir immer, wenn ich mich darum bemühe.“ Die SWE-Skala hat in deutschen Samples eine gute interne Konsistenz, die von Cronbachs $\alpha = .80$ bis $.90$ reicht [38]. Zur Validität liegen Ergebnisse aus empirischen Studien vor, die theoriekonforme positive Zusammenhänge zum dispositionalen Optimismus und zur Arbeitszufriedenheit sowie enge negative Zusammenhänge zu Ängstlichkeit, Depressivität und Burnout belegen [33].

Statistische Analysen

In Voranalysen wurden die soziodemografischen Variablen Alter bei Studienbeginn, Geschlecht und Abiturnote sowie die abhängigen Variablen Gesamtpunktzahlen des formativen MEQ* und summativen MEQ-1 für die zwei Jahrgänge 45 und 46 sowie die Selbstwirksamkeitserwartung mittels Kolmogorov-Smirnow-Tests auf Normalverteilung geprüft. Da die Annahme der Normalverteilung nicht bestätigt werden konnte, wurde nicht-parametrisch getestet. Für Vergleiche zwischen den unabhängigen Gruppen von Studierenden (Jahrgang 45 versus 46) wurden Mann-Whitney-U-Tests durchgeführt. Die Testgröße wurde analog in die Effektstärke Cohens d umgewandelt [39], [40]. Zusammenhänge zwischen Variablen wurden mit Korrelationsanalysen (Spearman rho) berechnet mit dem Korrelationskoeffizienten r als Effektstärke. Die Prädiktion der erzielten Punktzahl im MEQ-1 durch die Variablen Alter, Abiturnote, Vorwissen (Beruf, Studium), Selbstwirksamkeitserwartung und das MEQ* Ergebnis erfolgte mit multiplen Regressionsanalysen. Das Signifikanzniveau wurde bei diesen mit SPSS 26 durchgeführten Korrelations- und Regressionsanalysen auf $p < .05$ gesetzt, für die Mann-Whitney-U-Tests nach Bonferroni-Korrektur des α -Fehlers auf $p < .01$ [41].

Tabelle 1: Verteilung der zu erzielenden Punkte im formativen MEQ* und im summativen MEQ-1 nach Fächern (Punktzahl, Prozent)

	MEQ* Wintersemester 2017/18 (N=44) oder Sommersemester 2018 (N=44)	MEQ-1 Jahrgang 45 (N=44) Wintersemester 2017/18	MEQ-1 Jahrgang 46 (N=44) Sommersemester 2018
Anatomie	46 (39.7 %)	76 (32.3 %)	81 (31.7 %)
Physiologie	36 (31.0 %)	67 (28.5 %)	74 (29.0 %)
Klinisches Denken	13 (11.2 %)	39 (16.6 %)	45 (17.6 %)
Biologie	10 (8.6 %)	7 (3.0 %)	-
Biochemie	6 (5.2 %)	36 (15.3 %)	30 (11.8 %)
Chemie	3 (2.6 %)	-	-
Radiologie	2 (1.7 %)	8 (3.4 %)	16 (6.3 %)
Pharmakologie	-	2 (0.8 %)	5 (2.0 %)
Immunologie	-	-	4 (1.6 %)
Maximale Punktzahl	116	235	255
Bestehensgrenze (60 %)¹	70	141	153

¹ Punktzahl der Bestehensgrenze für den summativen MEQ-1 relevant, für den MEQ* zur Orientierung.

Ergebnisse

Beschreibung der Stichprobe und Gruppenvergleich

In dieser Untersuchung wurden 88 Medizinstudierende aus zwei Jahrgängen der UW/H (45 und 46) einbezogen. Hinsichtlich der Variablen Alter bei Studienbeginn, Geschlechtsverteilung, Abiturnote sowie geleisteter Berufsausbildung, Vorstudium, Praktika oder Freiwilliges Soziales Jahr und der Selbstwirksamkeitserwartung unterschieden sich die beiden Jahrgänge nicht signifikant voneinander und konnten deshalb für die weitere Auswertung als eine Stichprobe betrachtet werden.

Im Durchschnitt waren die Studierenden (46 Frauen, 42 Männer) zu Studienbeginn 22.3 ± 2.5 Jahre alt (Range 19 bis 29 Jahre). Die Abiturnote lag bei 2.1 ± 0.4 (Range 1.2 bis 3.2), wobei Frauen eine bessere Abiturnote als Männer haben (1.96 ± 0.42 versus 2.15 ± 0.39 ; Mann-Whitney- $U=700$, $p=.026$, $d=.488$), nicht signifikant nach α -Korrektur. Vor Studienbeginn haben 36 Studierende (40.9%) ein Freiwilliges Soziales Jahr absolviert oder einen Beruf im Gesundheitsbereich erlernt, z. B. Gesundheits- und Krankenpflege, Physiotherapie oder Rettungssanitäter*in. Insgesamt sind 80% der Berufe in dieser Stichprobe im medizinischen Bereich angesiedelt. Ein Vorstudium, z. B. Biologie, Pflegewissenschaften oder Zahnmedizin, haben 15 Studierende (17%) begonnen. Zusätzlich zu dem vor Studienbeginn zu absolvierenden sechsmonatigen Pflegepraktikum konnten 67% ($N=59$) weitere Praktika vorweisen. Die durchschnittliche Selbstwirksamkeitserwartung der Studierenden entspricht mit 30.2 ± 3.4 dem Referenzmittelwert von 29 Punkten [37].

Demografische Merkmale und MEQ-Ergebnisse

Die Studierenden erreichten im formativen MEQ* bei maximal möglichen 116 Punkten eine Gesamtpunktzahl von 66.9 ± 13.9 (Jahrgang 45 im Wintersemester 2017/18) bzw. von 65.4 ± 13.1 (Jahrgang 46 im Sommersemester 2018). Dieser geringe Unterschied zwischen den beiden Jahrgängen ist nicht signifikant. Die im Peer-Review-Verfahren durch die Medizinstudierenden vergebenen Punkte im formativen MEQ* sind nicht signifikant höher als in der objektivierten Nachauswertung durch einen fachlich qualifizierten Autor (OB) (siehe Tabelle 2). Beide Gesamtpunktzahlen korrelieren hoch miteinander (Spearman: $r=.837$, $p=.000$). Die größte Übereinstimmung findet sich im Fach Anatomie ($r=.933$), gefolgt von Biochemie ($r=.904$), Physiologie ($r=.834$), Radiologie ($r=.775$) und zuletzt Klinischem Denken ($r=.463$) (Ergebnisse aller Korrelationsanalysen $p=.000$). Für weitere Analysen wurden die durch den fachlich qualifizierten Autor (OB) ermittelten Ergebnisse im formativen MEQ* verwendet. Im summativen MEQ-1 erreichte Jahrgang 45 im Durchschnitt 160.8 ± 29.3 Punkte bei 235 maximal möglichen Punkten (entspricht 68.4%) und Jahrgang 46 im Durchschnitt 193.8 ± 29.6 Punkte von 255 möglichen Punkten (entspricht 76%). In der erreichten Gesamtpunktzahl bestehen nach deren z-Transformation keine signifikanten Unterschiede zwischen den Jahrgängen ($U=909$, $p=.753$, $d=.105$).

Bei der Prüfung des möglichen Zusammenhangs zwischen dem Alter bei Studienbeginn und den MEQ-Ergebnissen zeigt sich eine moderat positive Korrelation ($r=.306$, $p=.004$) zum formativen MEQ* und im gleichen Ausmaß ein Semester später für den summativen MEQ-1 ($r=.307$, $p=.004$). Darüber hinaus besteht eine positive Assoziation zwischen dem Alter bei Studienbeginn sowie der Abiturnote ($r=.341$, $p=.001$), bedingt durch die Wartesemester bis zum Studienbeginn. Zwischen den Geschlechtern

Tabelle 2: Gegenüberstellung der Ergebnisse im formativen MEQ* (Mittelwert, Standardabweichung) durch Peer Review der Medizinstudierenden und fachliche Auswertung durch einen Autor (OB)

	MEQ* Peer review Jahrgang 45 (N=44)	Fachliche Auswertung Autor (OB)	MEQ* Peer review Jahrgang 46 (N=44)	Fachliche Auswertung Autor (OB)
	M ± SD	M ± SD	M ± SD	M ± SD
Anatomie	31.48 ± 6.72	29.78 ± 6.88	31.06 ± 6.27	29.73 ± 6.40
Physiologie	21.48 ± 7.09	20.07 ± 6.98	20.99 ± 6.29	19.05 ± 6.13
Klinisches Denken	10.14 ± 1.90	9.57 ± 1.75	10.81 ± 1.89	10.61 ± 1.72
Biochemie	3.99 ± 1.74	3.91 ± 1.71	3.07 ± 1.52	2.80 ± 1.36
Radiologie	1.36 ± .55	1.26 ± .55	1.41 ± .52	1.29 ± .55
Summenwert	70.87 ± 13.65	66.89 ± 13.92	69.77 ± 12.44	65.37 ± 13.11

Tabelle 3: Gruppenvergleich zum beruflichen Vorwissen (Vorstudium oder Beruf) und Ergebnissen im formativen MEQ* sowie summativen MEQ-1 (Mittelwert, Standardabweichung)

	MEQ* Wintersemester 2017/18 oder Sommersemester 2018	MEQ-1 Jahrgang 45 Wintersemester 2017/18	MEQ-1 Jahrgang 46 Sommersemester 2018
	M ± SD	M ± SD	M ± SD
Vorstudium: ja (N=15)	73.41 ± 9.46	176.56 ± 13.82	201.38 ± 15.99
nein (N=72)	64.51 ± 13.71	157.16 ± 30.77	192.10 ± 31.82
Beruf: ja (N=36)	70.32 ± 10.54	168.87 ± 22.98	203.40 ± 17.52
nein (N=52)	63.23 ± 14.54	156.43 ± 31.67	185.00 ± 35.64

Tabelle 4: Prädiktion der Ergebnisse im formativen MEQ* (Einflussvariablen in Regressionsanalyse: Alter, Abiturnote, Selbstwirksamkeitserwartung, Berufsausbildung und Vorstudium) (N=86) (Regressionskoeffizient Beta, Prüfwert T, Signifikanz)

	Beta	T	p
(Konstante)		3.649	.000
Alter	.230	2.008	.048
Abiturnote	-.232	-2.264	.026
SWE	-.117	-1.186	.239
Beruf	.223	2.008	.048
Studium	.216	2.069	.042

bestehen keine statistisch signifikanten Unterschiede in der erreichten Gesamtpunktzahl im MEQ* und MEQ-1. Abiturnote und Selbstwirksamkeitserwartung weisen keinen signifikanten Zusammenhang zur Punktzahl im formativen MEQ* oder summativen MEQ-1 auf (Ergebnisse aller Korrelationsanalysen $p > .050$).

Die Berücksichtigung der Variablen zu den beruflichen Vorerfahrungen der Studierenden ergab für die geleisteten Praktika oder ein Freiwilliges Soziales Jahr keine signifikanten Unterschiede bezüglich der Gesamtpunktzahl im formativen MEQ* und summativen MEQ-1. Wenn jedoch ein Vorstudium (N=15) abgeleistet worden war, dann waren die erreichten Punkte im formativen MEQ* höher als ohne Vorstudium ($U=363$, $p=.021$, $d=.437$). Allerdings ist dieser Unterschied nicht signifikant, ebenso wenig wie ein Semester später bei den Punktzahlen im summativen MEQ-1. Studierende mit Berufsausbildung (N=36) erreichten sowohl im MEQ* mehr Punkte ($U=685$, $p=.033$, $d=.466$), als auch ein Semester später mit signi-

fikantem Unterschied im MEQ-1 ($U=615$, $p=.009$, $d=.607$) (siehe Tabelle 3).

Prädiktoren der MEQ-Ergebnisse

In multiplen Regressionsanalysen wurden alle Einflussfaktoren gemeinsam als mögliche Prädiktoren für die erreichte Punktzahl im formativen MEQ* und im summativen MEQ-1 (jeweils als abhängige Variable) berücksichtigt. Bei der Untersuchung der Prädiktoren für den MEQ* zeigte sich, dass die im Aufnahmeprozess erfassten Variablen Alter, Abiturnote, Berufsausbildung und Vorstudium signifikante Prädiktoren für das Abschneiden im MEQ* sind, aber nicht die Selbstwirksamkeitserwartung (siehe Tabelle 4). Dieses Modell erklärt jedoch insgesamt nur 23.1% der Varianz, so dass weitere, nicht bekannte Variablen für die Leistung in Frage kommen. Die gemeinsamen Varianzanteile der Variablen Alter, Berufsausbildung und Vorstudium wurden in einer schrittweisen Regressionsanalyse ausdifferenziert. Dabei wurden allein

Tabelle 5: Prädiktion der Ergebnisse im summativen MEQ-1 (Einflussvariablen in Regressionsanalyse: Alter, Abiturnote, Selbstwirksamkeitserwartung, Berufsausbildung, Vorstudium und formativer MEQ*) (N=88) (Regressionskoeffizient Beta, Prüfwert T, Signifikanz)

	Beta	T	p
(Konstante)		.335	.739
Alter	.005	.051	.960
Abiturnote	.081	.888	.377
SWE	.128	1.478	.143
Beruf	.110	1.120	.266
Studium	.019	.202	.841
MEQ*	.636	6.627	.000

Tabelle 6: Prädiktion der Ergebnisse im summativen MEQ-1 (Einflussvariablen in Regressionsanalyse: Fächer aus dem formativen MEQ*) (N=88) (Regressionskoeffizient Beta, Prüfwert T, Signifikanz)

	Beta	T	p
(Konstante)		1.870	.065
MEQ* - Anatomie	.442	4.704	.000
MEQ* - Physiologie	.199	2.201	.031
MEQ* - Klinisches Denken	.321	4.048	.000
MEQ* - Biochemie	-.179	-2.056	.043
MEQ* - Radiologie	.155	1.840	.069

mit dem Alter ($\beta=1.74$, $T=3.23$, $p=.002$) 10.8% der Varianz aufgeklärt.

Die Berücksichtigung von Alter, Abiturnote, Berufsausbildung, Vorstudium und Selbstwirksamkeitserwartung in der Regressionsanalyse ergab, dass diese keine signifikanten Prädiktoren für die Gesamtpunktzahl im MEQ-1 darstellen. Mit dem Ergebnis aus dem formativen MEQ* konnten dagegen allein schon 40.5% der Varianz ($\beta=1.58$, $T=7.61$, $p=.000$) aufgeklärt werden, im gemeinsamen Modell mit den vorgenannten Variablen 44.4% (siehe Tabelle 5). Der signifikante korrelative Zusammenhang von positiver Proportionalität zwischen dem formativen MEQ* und dem summativen MEQ-1 im Jahrgang 45 ($r=.769$, $p=.001$) und Jahrgang 46 ($r=.684$, $p=.001$) belegt dessen Inhaltsvalidität [42].

Zur differenzierten Betrachtung werden die fünf medizinischen Fächer, die sowohl im formativen MEQ* als auch im summativen MEQ-1 explizit ausgewiesen werden (vgl. Tabelle 1), in den Regressionsanalysen als unabhängige Variablen berücksichtigt. Durch die Punktzahl in den Fächern Anatomie, Physiologie, Klinisches Denken, Biochemie und Radiologie im MEQ* ließen sich 53.5% der Varianz im MEQ-1 erklären (siehe Tabelle 6). Bis auf die Radiologie erwiesen sich alle als signifikante Prädiktoren. Durch die Ergebnisse im Fach Anatomie alleine sind 36.2% der Varianz aufgeklärt und durch Hinzunahme des Klinischen Denkens weitere 11.2% (insgesamt 47.4%). Physiologie und Biochemie haben dagegen nur einen schwachen prädiktiven Effekt.

Diskussion

In dieser Studie wurde untersucht, ob eine formative Prüfung wie der MEQ* im Modellstudiengang Humanme-

dizin der Universität Witten/Herdecke als Prädiktor für die summative Prüfung MEQ-1 ein Semester später dienen kann. Zudem wurden als mögliche Einflussvariablen Abiturnote, berufliches Vorwissen, Selbstwirksamkeitserwartung, Alter und Geschlecht berücksichtigt.

Geschlecht, Abiturnote und Selbstwirksamkeitserwartung

Frauen wiesen sowohl im MEQ* als auch im MEQ-1 gegenüber Männern keine besseren Leistungen auf, trotz eines geschlechtsbezogenen Unterschiedes bei der Abiturnote, der knapp die statistische Signifikanz verfehlte. Dies ist ähnlich zu der Beobachtung unter Medizinstudierenden an der Universität Heidelberg [32]. Auch im zweiten Staatsexamen Herbst 2018 zeigten sich für alle Medizinstudierenden in Deutschland vergleichbare Prüfungsleistungen bei Frauen mit 79.0% gegenüber Männern mit 79.2% der Gesamtpunkte [42]. Trapmann et al. [30] berichten in ihrer Metaanalyse zur Vorhersage des Studienerfolgs die Beobachtung, dass in den Studiengängen Zahn- und Veterinärmedizin die Validität der Abiturnote für die Vorklinik höher sei als für die klinischen Semester. Gleiches gilt für die Humanmedizin mit etwa 23% Aufklärung der Leistungsvarianz durch frühere akademische Leistungen zu Studienbeginn und insgesamt mit etwa 9%, wie Ferguson et al. [43] in einem systematischen Literaturreview belegen konnten. Frühere akademische Leistungen sind demnach ein guter, aber nicht perfekter Prädiktor für die Leistung in der medizinischen Ausbildung. Zudem waren Leistungsunterschiede zwischen den Geschlechtern eher gering und erreichten nur in großen Kohorten statistische Signifikanz. So zeigte sich auch in der vorliegenden Studie bei mäßiger Prädik-

tion für den formativen MEQ* kein bedeutsamer Zusammenhang zwischen der Abiturnote und dem Abschneiden im formativen MEQ* oder im summativen MEQ-1. Gleiches gilt für die Selbstwirksamkeitserwartung der Studierenden [38], womit die Befunde von Klassen & Klassen [28] und insbesondere von Demirören et al. [29] zur Wirkung im Kontext des problemorientierten Lernens nicht bestätigt werden konnten.

Alter und berufliche Vorbildung

Die Gesamtpunktzahl im formativen MEQ* weist für den Jahrgang 46 (Sommersemester 2018) einen moderat positiven Zusammenhang mit dem Alter auf, der für den Jahrgang 45 (Wintersemester 2017/18) zwar auch positiv ist, aber die statistische Signifikanz verfehlt. Der positive Alterseffekt ist ein Semester später beim MEQ-1 ebenfalls noch nachweisbar. Studierende mit einem höheren Alter haben in ihren Wartesemestern berufliche Vorerfahrungen gesammelt in Form eines Vorstudiums, einer Berufsausbildung im Gesundheitsbereich, spezifischen Praktika oder eines Freiwilliges Sozialen Jahres (FSJ). Allerdings hatten Praktika und FSJ keinen relevanten Einfluss auf die Ergebnisse im formativen MEQ* oder summativen MEQ-1, da scheinbar die erworbenen praktischen Kenntnisse ohne ein strukturierendes theoretisches Fundament keinen Vorteil für die in den Prüfungen geforderten akademischen Leistungen darstellen. Dagegen erreichten die Studierenden mit Vorstudium (17% der Stichprobe) im Gruppenvergleich bessere Leistungen im formativen MEQ* und hatten im Fach Physiologie erhöhte Punktzahlen. Die Studierenden mit beruflicher Vorbildung (40.9% der Stichprobe) erzielten ebenfalls höhere Punktzahlen im formativen MEQ*, speziell im Fach Anatomie. Die Medizinstudierenden scheinen aus beiden Arten der Vorbildung profitieren zu können. Einerseits die strukturelle, fachunabhängige Vorbildung eines Vorstudiums, in dem eigenverantwortliches universitäres Lernen erlernt wird, und andererseits die theoretisch-praktische Vorbildung einer Berufsausbildung im Gesundheitsbereich (80% der Stichprobe), die inhaltliches Wissen aus der Berufsschule und praktische Erfahrungen mit sich bringt. Dieser Vorteil scheint jedoch nur bis zum Ende des ersten Semesters für den formativen MEQ* zu gelten, nicht bis zum summativen MEQ-1, der ein Semester später geschrieben wird. Dies korrespondiert mit den Ergebnissen von Ferguson et al. [43], bezogen auf frühere akademische Leistungen. Parallelen zeigen sich ebenfalls in einer Studie von Grendel et al. [34], in der die Effekte berufspraktischer Vorerfahrung von beruflich qualifizierten Studierenden untersucht wurden. Dauer und Relevanz der beruflichen Vorerfahrungen scheinen dabei einen bedeutsamen Einfluss auf den Studienerfolg zu haben.

Formativer MEQ* und Peer-Review-Verfahren

Die prädiktive Bedeutung der Punktzahl des formativen MEQ* für die spätere Leistung im summativen MEQ-1 konnte im Regressionsmodell mit 40.5% Varianzaufklärung

bestätigt werden. Zudem zeigte sich in Korrelationsanalysen eine hohe inhaltliche Validität. Hinsichtlich der Bedeutung spezifischer Fächer im formativen MEQ* wird die besondere Wichtigkeit der Anatomie und des Klinischen Denkens deutlich. Die Anatomie stellt traditionell beim Start in das Medizinstudium vom Lernumfang her die größte Herausforderung dar, so auch als übergeordneter Themenbereich Bewegungsapparat im 1. Semester des Modellstudiengangs Humanmedizin der Universität Witten/Herdecke. Das Klinische Denken wird durch das Problemorientierte Lernen (POL) geschult und kann entsprechend des Constructive Alignment am besten mit MEQ-Freitextfragen geprüft werden [17], [18]. Dabei sind die im formativen MEQ* im Peer-Review-Verfahren vergebenen Punkte nicht signifikant erhöht im Vergleich zur Nachkorrektur durch den fachlich qualifizierten Autor (OB). Somit scheint das studentische Peer-Review-Verfahren, trotz eher günstiger Bewertung durch die Kommiliton*innen, eine effiziente Methode zu sein, um in einer formativen Prüfung ein schnelles Feedback zu geben. Die hohe Korrelation zwischen der studentischen und der objektivierten Nachauswertung belegt die Reliabilität des MEQ*, die am höchsten für die Anatomie und – wie zu erwarten – am niedrigsten für Klinisches Denken ist [14], [15], [44]. Problematisch ist jedoch, dass die Studierenden sich z. T. nicht auf die formative Prüfung vorbereiten. Wenn bei dem formativen MEQ* eine Bestehensgrenze von 60% [45] angelegt werden würde, wie es beim ersten Staatsexamen nach § 14 der Approbationsordnung für Ärzte (2002) der Fall ist, dann hätten mehr als die Hälfte der Studierenden nicht bestanden. Der durchschnittlich erzielte Punktwert entspricht 57% und liegt damit zwar deutlich oberhalb der durchschnittlich 30% erzielten Punkte von Psychologiestudierenden der UW/H im formativen Progress Test Psychologie [46], zeigt aber deutlich, dass die Studierenden das Prinzip und den Nutzen formativer Tests als Feedback über den eigenen Wissensstand, zur Lernmotivation und Reduktion der Prüfungsangst noch nicht ausreichend realisieren [47].

Limitationen

Hier sollte zunächst die kleine Fallzahl von 88 Medizinstudierenden aus zwei Semestern in Betracht gezogen werden, die zudem eine heterogene Gruppe bzgl. ihres Alters und der beruflichen Vorerfahrungen darstellen. Insofern ist eine Generalisierbarkeit der Ergebnisse wahrscheinlich nicht gegeben, da es sich um zufällige, kohortenspezifische Effekte handeln könnte. Die im Peer-Review-Verfahren ermittelten Ergebnisse des formativen MEQ* waren in der Nachkorrektur durch den fachlich qualifizierten Autor (OB) zwar nicht signifikant unterschiedlich, aber die Nachkorrektur gestaltete sich dahingehend als schwierig, dass die Antworten der Studierenden häufig nicht exakt im Antworthorizont abgebildet waren. So lag es im Ermessensspielraum der Nachkorrektur, wie die gegebenen Antworten gewertet werden. Hier besteht das Risiko, dass verschiedene Prüfer*innen zu unterschiedli-

chen Bewertungen der Antworten kommen und somit die Reliabilität des formativen MEQ* gemindert wird. Dies ist ein grundsätzliches Problem der Freitextformate und erfordert einen relativ hohen Aufwand, sowohl bei der Formulierung der Antworthorizonte, als auch im Reviewprozess und der Korrektur [14], [15].

Schlussfolgerungen

In der vorliegenden Untersuchung konnte gezeigt werden, dass es im Verlauf des 2. Semesters zu einem deutlichen Wissenszuwachs gekommen ist, da der Jahrgang 45 (Wintersemester 2017/18) im summativen MEQ-1 einen durchschnittlichen Punktwert von 68% und der Jahrgang 46 (Sommersemester 2018) von 76% erzielte. Die Leistungssteigerung im Vergleich zum Ergebnis im formativen MEQ* ist wohl auf gezielte Prüfungsvorbereitung zurückzuführen, bei annähernd gleichem Lernkonzept der POL-Tutorien während des 1. und 2. Semesters. Der MEQ* als formative Testung scheint den Studierenden eine Rückmeldung zu ihrem aktuellen Wissensstand zu geben. Das Nicht-Erreichen der Bestehensgrenze könnte für viele Studierende ein „Weckruf“ sein, das Lernbemühen im Hinblick auf den summativen MEQ-1 zu intensivieren. Parallelen können hier mit der Interview-Studie von Heeneman [48] gezogen werden, in der Studierende bestätigen, dass sie die Ergebnisanalyse des formativen Progress Test zur Anpassung ihrer Lernstrategien nutzen. Zudem ist denkbar, dass die nach dem formativen MEQ* vorhandene Kenntnis über das Freitextformat den Studierenden gemäß des Constructive Alignment hilft, ihren Lernstil und die Prüfungsvorbereitung anzupassen. Somit wäre dieses Ziel einer formativen Prüfung durch den MEQ* erreicht und der mit der Entwicklung und Durchführung dieser Art der formativen Testung verbundene Aufwand für die Dozierenden und das Studiendekanat der Universität Witten/Herdecke gerechtfertigt.

Ausblick

Nachdem gezeigt werden konnte, dass das Ergebnis des formativen MEQ* am Ende des 1. Semesters im Modellstudiengang Humanmedizin der Universität Witten/Herdecke ein bedeutsamer Prädiktor für das Ergebnis des summativen MEQ-1 am Ende des 2. Semesters ist, wäre die zielführende Fragestellung einer weitergehenden Studie, für welchen Zeitraum im Studienverlauf der MEQ* seine Funktion als Prädiktor behält. Sagt das Abschneiden in der formativen MEQ*-Prüfung auch etwas über die gesamte summative Äquivalenzprüfung (M1, Physikum) aus? Bemerkenswert ist außerdem, dass hinsichtlich der geprüften Fächer im MEQ* die Anatomie und das klinische Denken von herausragender Bedeutung für das Ergebnis im summativen MEQ-1 sind. Da das Abschneiden im Fach des Klinischen Denkens hier in keinem Zusammenhang mit den Gesamtpunkten der anderen Fächer zu stehen scheint, kommt dem Unterrichtskonzept des

Problemorientierten Lernens eine besondere Rolle zu diesem Kompetenzerwerb zu. Aus der niedrigen Durchschnittspunktzahl und den einzelnen Studierendenaussagen heraus lässt sich ableiten, dass sich die Studierenden wohl nur z. T. gezielt auf den formativen MEQ* vorbereiten. Um den Nutzen dieser formativen Prüfung als Feedback zu steigern, wäre eine sorgfältige Prüfungsvorbereitung von Seiten der Studierenden wünschenswert. In weitergehenden Studien wäre es wichtig zu klären, in wie weit das Feedback aus dem formativen MEQ* bei den Studierenden tatsächlich zu einer Änderung der Lernstrategien genutzt wird.

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Literatur

1. Biggs J. Enhancing teaching through constructive alignment. *High Educ.* 1996;32:347-364. DOI: 10.1007/BF00138871
2. Tang C, Biggs J. Teaching for Quality Learning at University. 4th Ed. Buckingham: Open University Press/McGraw Hill; 2011.
3. Wildt B, Wildt J. Lernprozessorientiertes Prüfen im "Constructive Alignment". HRK Projekt nexus. Berlin, Bonn: HRK; 2011. Zugänglich unter/available from: <https://www.hrk-nexus.de/fileadmin/redaktion/hrk-nexus/07-Downloads/07-03-Material/pruefen.pdf>
4. Maier U. Vergleichsarbeiten im Spannungsfeld zwischen formativer und summativer Leistungsmessung. *Dtsch Schule.* 2010;102:60-69.
5. Black P, William D. Assessment and Classroom Learning. *Assessment in Education: Principles. Policy Pract.* 1998;5:7-74. DOI: 10.1080/0969595980050102
6. Ben-David MF. The role of assessment in expanding professional horizons. *Med Teach.* 2000;22(5):472-477. DOI: 10.1080/01421590050110731
7. Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract.* 2010;15(5):695-715. DOI: 10.1007/s10459-010-9232-9
8. Epstein RM. Assessment in medical education. *New Engl J Med.* 2007;356(4):387-396. DOI: 10.1056/NEJMr054784
9. Schuwirth LW, van der Vleuten CM. Current Assessment in Medical Education: Programmatic Assessment. *J Appl Test Technol.* 2019;Vol 20(S2):2-10.
10. Thiessen N, Fischer MR, Huwendiek S. Assessment methods in medical specialist assessments in the DACH region - overview, critical examination and recommendations for further development. *GMS J Med Educ.* 2019;36(6):Doc78. DOI: 10.3205/zma001286
11. Hift RJ. Should essays and other open-ended-Type questions retain a place in written summative assessment in clinical medicine? *BMC Med Educ.* 2014;14:249. DOI: 10.1186/s12909-014-0249-2
12. Knox JD. What is a Modified Essay Question? *Med Teach.* 1989;1:51-s7. DOI: 10.3109/01421598909146276

13. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? *Research paper. BMC Med Educ.* 2007;7:49. DOI: 10.1186/1472-6920-7-49
14. Feletti GI, Smith EK. Modified Essay Questions: Are they Worth the Effort? *Med Educ.* 1986;20:126-132. DOI: 10.1111/j.1365-2923.1986.tb01059.x
15. Nendaz M, Tekian A. Assessment in problem-based learning medical schools: A literature review. *Teach Learn Med.* 1999;11(4):232-243. DOI: 10.1207/S15328015TLM110408
16. Butzlaff M, Hofmann M, Brunk C, Edelhäuser F, Lutz G, Reißweber J, Scheffer C, Tauschel D, Thiele S, Wirth D, Zupanec M. Der Modellstudiengang Medizin an der Universität Witten/Herdecke - auf dem Weg zur lebenslang lernfähigen Arztpersönlichkeit. In: Fuhrmann M, Güdler J, Kohler J, Schmidt U, editors. *Handbuch Qualität in Studium, Lehre und Forschung.* Berlin: DUZ Verlags- und Medienhaus GmbH; 2014. p.65-104.
17. Wild D, Rützler M, Haarhaus M, Peters K. The Modified Essay Question (MEQ) Test at the Medical School of the University Witten/Herdecke. *Gesundheitswesen (Suppl Med Ausbild).* 1998;15(Suppl2):65-69.
18. Anderson L, Krathwohl D. A taxonomy for teaching, learning and assessing: A revision of Bloom's taxonomy of educational objectives. London: Pearson Education; 2000.
19. Albanese MA, Mitchel S. Problem-based Learning: A Review of Literature on its Outcomes and Implementation Issues. *Acad Med.* 1993;68(1):52-81. DOI: 10.1097/00001888-199301000-00012
20. Colliver JA. Effectiveness of Problem-based Learning Curricula: Re-search and Theory. *Acad Med.* 2000;75(3):259-266. DOI: 10.1097/00001888-200003000-00017
21. Schlett CL, Doll H, Dahmen J, Polacsek O, Federkeil G, Fischer MR, Bamberg F, Butzlaff M. Job requirements compared to medical school education: differences between graduates from problem-based learning and conventional curricula. *BMC Med Educ.* 2010;10:1. DOI: 10.1186/1472-6920-10-1
22. Koh GC, Khoo HE, Wong ML, Koh D. The effects of problem-based learning during medical school on physician competency: a systematic review. *CMAJ.* 2008;178(1):34-41. DOI: 10.1503/cmaj.070565
23. Universität Witten/Herdecke. Prüfungsordnung der Universität Witten/Herdecke für den Modellstudiengang Medizin. Witten: Universität Witten/Herdecke; 2018. Zugänglich unter/available from: https://www.uni-wh.de/fileadmin/user_upload/03_G/07_Humanmedizin/Studiengang_HM/links_und_downloads/Pruefungsordnung_MSG.pdf
24. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of Clinical Competence using Objective Structured Examination. *Br Med J.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
25. Zupanec M, Ehlers S, Bauer D, Hofmann M, Ehlers JP. Was prüft der Modified Essay Questions Test an der Universität Witten/Herdecke? In: *Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA).* Wien, 19.-22.09.2018. Düsseldorf: German Medical Science GMS Publishing House; 2018. Doc28.2. DOI: 10.3205/18gma116
26. Bandura A. Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychol Rev.* 1977;84:191-215. DOI: 10.1037/0033-295X.84.2.191
27. Bandura A. Social cognitive theory: an agentic perspective. *Annu Rev Psychol.* 2001;52:1-26. DOI: 10.1146/annurev.psych.52.1.1
28. Klassen RM, Klassen JRL. Self-efficacy beliefs of medical students: a critical review. *Perspect Med Educ.* 2018;7(2):76-82. DOI: 10.1007/s40037-018-0411-3
29. Demirören M, Turan S, Öztuna D. Medical students' self-efficacy in problem-based learning and its relationship with self-regulated learning. *Med Educ Online.* 2016;21:30049. DOI: 10.3402/meo.v21.30049
30. Trapmann S, Hell B, Weigand S, Schuler H. Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Z Päd Psychol.* 2007;21:11-27. DOI: 10.1024/1010-0652.21.1.11
31. McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *BMJ.* 2003;327(7407):139-142. DOI: 10.1136/bmj.327.7407.139
32. Kadmon G, Resch F, Duelli R, Kadmon M. Predictive Value of the School-leaving Grade and Prognosis of Different Admission Groups for Academic Performance and Continuity in the Medical Course - a Longitudinal Study. *GMS Z Med Ausbild.* 2014;31(2):Doc21. DOI: 10.3205/zma000913
33. Schwarzer R, Jerusalem M. Das Konzept der Selbstwirksamkeit. In: Jerusalem M, Hopf D, editors. *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen.* Weinheim: Beltz; 2002. p.28-53.
34. Grendel T, Lübke H, Haußmann I. Effekte der Dauer und der Qualität berufspraktischer Vorerfahrungen auf den Studienerfolg beruflich Qualifizierter. *Beitr Hochschulforsch.* 2014;36:40-63.
35. WMA. Deklaration von Helsinki, Version 2008. Ferney-Voltaire: WMA; 2008. Zugänglich unter/available from: https://www.bundesärztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/International/Deklaration_von_Helsinki_2013_20190905.pdf
36. Gößling S. Systematische Übersichtsarbeit zur Anwendung schriftlicher Formate, Wissen kompetenzorientiert abzufragen. Dissertation. Witten: Universität Witten/Herdecke, Fakultät für Gesundheit; 2017.
37. Schwarzer R, Jerusalem M. Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Frei Universität Berlin; 1999.
38. Hinz A, Schumacher J, Albani C, Schmid G, Brähler E. Bevölkerungsrepräsentative Normierung der Skala zur Allgemeinen Selbstwirksamkeitserwartung. *Diagnostika.* 2006;52:26-32. DOI: 10.1026/0012-1924.52.1.26
39. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. *J Experim Psychol Gen.* 2012;141(1):30. DOI: 10.1037/a0026092
40. Lenhard W, Lenhard A. Berechnung von Effektstärken. Dettelbach: Psychometrica; 2016. Zugänglich unter/available from: <https://www.psychometrica.de/effektstaerke.html>
41. Victor A, Elsässer A, Hommel G, Blettner M. Wie bewertet man die p-Wert-Flut? Hinweise zum Umgang mit dem multiplen Testen. Teil 10 der Serie zur Bewertung wissenschaftlicher Publikationen. *Dtsch Arztebl Int.* 2010;107(4):50-56. DOI: 10.3238/arztebl.2010.0050
42. Institut für medizinische und pharmazeutische Prüfungsfragen. Ergebnisse des Zweiten Abschnitts der Ärztlichen Prüfung - Herbst 2018. Mainz: IMPP; 2018. Zugänglich unter/available from: <https://www.impp.de/informationen/berichte/ergebnisberichte/medizin.html>
43. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ.* 2002;324(7343):952-957. DOI: 10.1136/bmj.324.7343.952

44. von Bergmann H, Dalrymple KR, Wong S, Shuler CF. Investigating the Relationship between PBL Process Grades and Content Acquisition Performance in a PBL Dental Program. *J Dent Educ.* 2007;71(9):1160-1170. DOI: 10.1002/j.0022-0337.2007.71.9.tb04380.x
45. Möltner A, Schellberg D, Jünger J. Grundlegende quantitative Analysen medizinischer Prüfungen. *GMS Z Med Ausbild.* 2006;23(3):Doc 53. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000272.shtml>
46. Zupanic M, Troche SJ, Ehlers JP. Absolvierendenniveau im formativen Progress Test Psychologie: Anspruch oder Wirklichkeit? In: Krämer M, Preiser S, Brusdeylins K, editors. *Psychologiedidaktik und Evaluation XII*. Aachen: Shaker; 2018. p.359-368.
47. Schüttelpelz-Brauns K, Karay Y, Arias J, Gehlhar K, Zupanic M. Comparison of the evaluation of formative assessment at two medical faculties with different conditions of undergraduate training, assessment and feedback. *GMS J Med Educ.* 2020;37(4):Doc41. DOI: 10.3205/zma001334
48. Heeneman S, Schut S, Donkers J, Vleuten van der C, Muijtjens A. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Med Teach.* 2017;39(1):44-52. DOI: 10.1080/0142159X.2016.1230183
49. Moosbrugger H, Kelava A, editors. *Testtheorie und Fragebogenkonstruktion*. 2. Aufl. Berlin: Springer Verlag; 2012. DOI: 10.1007/978-3-642-20072-4

Korrespondenzadresse:

Jun.-Prof. Dr. phil. Michaela Zupanic
Universität Witten/Herdecke, Fakultät für Gesundheit,
Interprofessionelle und kollaborative Didaktik in Medizin-
und Gesundheitsberufen, Witten, Deutschland
Michaela.Zupanic@uni-wh.de

Bitte zitieren als

Büssing O, Ehlers JP, Zupanic M. The prognostic validity of the formative for the summative MEQ (Modified Essay Questions) . *GMS J Med Educ.* 2021;38(6):Doc99.
DOI: 10.3205/zma001495, URN: urn:nbn:de:0183-zma0014958

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/zma/2021-38/zma001495.shtml>

Eingereicht: 01.03.2021

Überarbeitet: 14.06.2021

Angenommen: 28.06.2021

Veröffentlicht: 15.09.2021

Copyright

©2021 Büssing 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/>.